

LEARNING FROM INFORMATIVE LOSSES BOOSTS THE SENSE OF AGENCY

Bartosz Majchrowicz^{1*}, Eugenia Kulakova^{2,3}, Steven Di Costa², Patrick Haggard²

- 1) Consciousness Lab, Institute of Psychology, Jagiellonian University, Krakow, Poland
- 2) Institute of Cognitive Neuroscience, University College London, United Kingdom
- 3) Charité Universitätsmedizin Berlin, Klinik für Psychiatrie und Psychotherapie, Campus Benjamin Franklin, Berlin, Germany

* corresponding author at: majchrowicz.b@gmail.com
Institute of Psychology, Jagiellonian University, ul. Ingardena 6, 30-060 Krakow, Poland.

AUTHOR FINAL COPY

Cite as:

Majchrowicz, B., Kulakova, E., Di Costa, S., & Haggard, P. (2020). Learning from informative losses boosts the sense of agency. *Quarterly Journal of Experimental Psychology*, 1747021820958258.

DOI:

<https://doi.org/10.1177/1747021820958258>

Supplemental files:

<https://journals.sagepub.com/doi/suppl/10.1177/1747021820958258>

Abstract

Sense of agency, the feeling of having control over one's actions, is modulated by whether one's choices lead to desired or undesired outcomes. Learning similarly depends on outcome values from previous experience. In the current study we evaluate a possible link between the sense of agency and learning, by investigating how intentional binding, an implicit measure of agency, changes during a probabilistic learning task. In two experiments we show increased intentional binding in trials that follow losses, compared to trials that follow wins. Experiment 1 demonstrated that this post-error agency boost (PEAB) effect is rule-specific, since it did not occur if the trial following an error involved different action-outcome contingencies. Furthermore, PEAB was not modulated by the type of outcome presentation (monetary vs. affective). Experiment 2 showed that the PEAB effect can also occur when the current action involves a forced (as opposed to free) choice, but only when the *previous*, loss-provoking action was chosen freely. Thus, PEAB occurs when current actions are informed by outcomes of one's own previous action choices. EEG data linked these effects to two event-related potential components, namely the Feedback Related Negativity and the P300. Taken together, these results support the notion that PEAB reflects an adaptive property of human sense of agency, facilitating effective learning about the action-outcome structure of a specific task, in order to optimize future performance. By clarifying the conditions for enhancing the sense of agency through learning, this work adds to our understanding of human learning and agency.

Keywords

intentional binding; sense of agency; probabilistic learning; action selection; decision making; volitions

INTRODUCTION

Learning from errors is a crucial skill allowing humans to optimize their behaviour and decrease errors in future action (Metcalf, 2017). Learning should direct future voluntary action control, since behaviours can be adapted according to what has been learnt and aligned to current goals. Volition refers to the ability to choose and guide one's own actions, based on one's goals, reasons and beliefs. It is clear from both developmental and social models (Fried, Haggard, He, & Schurger, 2017; Haggard, 2017) that volition is modulated by learning. Nevertheless, action-outcome reinforcement learning and volition have generally been studied in separate literatures.

Importantly, volitional action is often accompanied by a subjective sense of agency (SoA), i.e., a feeling that one is in control over one's actions and their external consequences (Gallagher, 2012; Haggard, 2019). In the current work we consider whether the subjective sense of agency might provide an important bridge between volition and outcome-based learning. Here we investigate how the process of learning, and the specific features of the learning task, might influence subjective SoA, and, through this, guide the future behaviour of an agent.

Research on SoA has employed both explicit measures (based on subjective reports) and implicit measures of agency experience (Dewey & Knoblich, 2014), with the latter gaining a major interest in the past years. The measures can be seen as reflecting two levels of the experience, i.e. an explicit/declarative judgement and an implicit/pre-reflective feeling of agency, respectively (Synofzik, Vosgerau, & Newen, 2008). One implicit measure of SoA is intentional binding (IB), also referred to as temporal binding (Haggard, Clark, & Kalogeras, 2002). In IB, an action-outcome interval is perceived to be shorter when outcomes result from volitional action, compared to when the same outcomes result from TMS-evoked or passive movements, or from forced action choices – in other words, freely-chosen intentional actions and their effects are subjectively bound together in time (for a review, see Moore & Obhi, 2012).

It has previously been shown that IB can be influenced by the valence of action outcomes. In one study (Takahata et al., 2012), auditory outcomes of self-initiated button presses were either neutral, or associated with monetary rewards or losses. IB was reduced on trials where participants lost money, rather than gained it. The authors interpreted this as a manifestation of self-serving bias (Mezulis, Abramson, Hyde, & Hankin, 2004): participants readily attributed positive events to themselves, unlike negative ones which were less often attributed as their own in order to prevent aversive psychological states. Similar results were reported by Yoshie & Haggard (2013) who employed the IB paradigm with positive and negative human

vocalizations serving as action outcomes. The authors found that negative outcomes, compared to positive outcomes, reduced IB. Thus, both the economic and affective value of the outcome has been shown to modulate IB. Valence effects such as self-efficacy and self-serving bias therefore operate not only on the explicit, narrative level of experience, but also on more implicit level of SoA captured by IB (cf. Gallagher, 2000).

Volitional actions typically aim at positive outcomes. But even when an undesired outcome, i.e. an error, does occur we can still use the error to learn how to achieve a desired outcome with a subsequent action. Long-term behavioural success may be increased by unsuccessful action attempts that produce negative outcomes, but nonetheless allow for testing predictions regarding the structure of the task, rather than by constantly avoiding and/or discounting errors, without learning from them (Huelser & Metcalfe, 2012; Kornell, Hays, & Bjork, 2009). Knowing how and why errors arise can help us to better avoid them in the future, making strategies that involve some errors more beneficial in the long run. In such accounts, learning is understood as a process starting immediately from registering undesired outcomes, aimed at optimizing subsequent behaviour by guiding future action choices. Such arguments view error as functional for long-term performance. In contrast, the implications of errors for SoA have barely been considered, even though SoA may be strongly linked to motivation and ultimately to performance (Metcalfe & Greene, 2007).

A recent study by Di Costa, Théro, Chambon, & Haggard (2017) used the probabilistic reversal learning paradigm (PRL) to address the link between learning an agency. The task required participants to choose between pressing one of two keys, each mapped to a different monetary reward probability (80:20). This key mapping changed after a few rewarded trials, so participants had to continuously monitor the action-outcome mapping to optimize their *next* action choice based on the received feedback. IB was combined with PRL in order to assess how binding would change as a function of feedback from the previous trial. IB was increased on the next trial following a loss – an effect the authors called “post-error agency boost”, or PEAB¹. Crucially, the effect was absent in a random feedback condition, i.e. when learning was not possible due to 50:50 reward scheduling. Using reinforcement learning modeling (Niv, Edlund, Dayan, & O’Doherty, 2012), the authors also found that learning rates were higher for rewarded than non-rewarded action effects and also that a strong PEAB was pronounced in

¹ More specifically, however, the effect was driven not so much by incorrect actions, as by negative feedback. Because of the probabilistic nature of the task, these two concepts were dissociable. IB was elevated by preceding losing outcomes, not wrong responses, so the effect might plausibly also be called post-*loss* agency boost. However, for the sake of consistency with this previous paper (Di Costa et al., 2017), we keep the term PEAB.

individuals' readiness to learn from errors (rather than wins), which speaks for a dominant role of loss feedback in the observed effect. The fact that the PEAB effects occurred only when it was possible to *learn* from feedback seems to undermine explanations based on the merely arousing value of errors. An account based on arousal alone would predict that PEAB should also have been observed in the random feedback condition. This suggests that the SoA is only influenced by errors that can inform and optimize subsequent behaviour, and thus have an adaptive function. However, to support such claims of adaptivity a more direct investigation of the generality of the PEAB effect is needed. The present study aims to characterize the relation between preceding loss/error and sense of agency in more detail to further specify the requirements of error-induced agency changes during learning.

In a first experiment, we investigated whether the PEAB reflected a non-specific effect of negative feedback or was rather linked to learning a specific action-outcome contingency. In natural environments, agents rarely deal with only one problem or learning task at a time, and instead must handle multiple, interspersed action-outcome contingencies. Does an error on one task alter sense of agency on a subsequent, different task? To investigate this issue, we employed a task-switching procedure requiring participants to engage in two reversal-learning tasks alternating in AABB fashion. This design engenders between-task interference effects, because of the cognitive difficulty of switching between task sets (Monsell, 2003). Notably, mixing tasks within blocks generally results in poorer performance than performing the tasks in separate blocks (Allport, Styles, & Hsieh, 1994; Rogers & Monsell, 1995), although such mixing costs may also arise from task ambiguity (Rubin & Meiran, 2005). By employing alternating runs of the tasks (as opposed to task-cueing) we made tasks switches predictable and therefore easy to monitor (cf. Koch, 2005). Moreover, previous research has shown that separating response sets between tasks helps to further reduce switch-dependent interference effects (Gade & Koch, 2007; Yeung & Monsell, 2003).

We were particularly interested whether the PEAB effect would occur if the trial following a loss would be part of a different task set than the task in which the original error was produced. The transfer of an error-dependent effect between tasks would be consistent with phasic arousal elicited by losses and suggest PEAB to have a globally motivational, incentive character – although still requiring a context that allows for learning (cf. Di Costa et al., 2017). In other words, if committing an informative error in one task could produce PEAB the subsequent trial of another task the effect would propagate between tasks and thus be *task-general*. Alternatively, the PEAB effect might be *specific* to each task, and not transfer to a different, interspersed task. Such a result might suggest that PEAB plays a role in motivating persisting

goal-directed action in situations where people still need to learn what to do. PEAB would then be intimately related to the feelings of engagement in learning, and of mastery once actions have eventually been learned. To investigate the generality vs specificity of PEAB we measured intentional binding in a PRL paradigm that involved two alternating tasks. Furthermore, in the case of a task-specific PEAB effect, one might ask which affective and cognitive factors influence it. Our first experiment investigated these factors by varying the meaning of the outcome stimuli produced by participants' actions. In the second experiment, we investigated the role of free action choices on PEAB task-specificity, and also examined electrophysiological components corresponding to the observed behavioural effects.

EXPERIMENT 1

In Experiment 1, we employed a task switching PRL paradigm with two alternating tasks to investigate whether PEAB would propagate across tasks. Moreover, we tested whether the added affective value of the outcomes would influence the PEAB effect beyond monetary value alone. Thus, action outcomes valence was either presented as monetary outcome (Di Costa et al. 2017), or through human facial expressions. Using two different types of outcomes also ensured that both tasks were visually sufficiently different, providing a clear visual marker of the two distinct tasks and action-outcome contingencies. This arrangement excluded alternative explanations of any transfer of PEAB from one task to another in terms of confusion between tasks.

Method

Participants

30 volunteers (7 males) participated in the study, with a mean age of 25.1 ($SD = 4.2$). The sample size was chosen to approximately match or surpass the power of previous studies reporting intentional binding modulations due to outcome valence (Takahata et al., 2012; Yoshie & Haggard, 2017). All participants gave informed consent and were compensated at the rate of £7.50 per hour plus an additional bonus depending on their performance. All had normal or corrected-to-normal vision and hearing. The research was approved by the UCL Research Ethics Committee.

Stimuli and procedure

Participants were examined one at a time in a quiet room. The procedure was run on a laptop computer with a 15.6 inch display using Matlab with Psychophysics Toolbox (Brainard,

1997). The procedure consisted of two PRL tasks alternating in AABB order, i.e. 2 trials of one task were predictably followed by 2 trials of another task.

Both tasks were equivalent probabilistic reversal learning tasks (Figure 1). Two keys were available to press in each task: Z/X and N/M as possible left and right response keys, giving 4 possible responses, made with the left middle, left index, right index and right middle fingers respectively. Importantly, each task was performed with just one hand, thus, references to left and right responses refer to the external spatial position of the key pressed within each task, and not to the hand used to respond. Thus, a typical participant might choose between pressing Z or X in one task, and between pressing N or M in the other task. Key set and task mapping were counterbalanced across subjects, as were green and blue coloured backgrounds associated with each task. The fixed AABB order, background colour cueing, and non-overlapping response sets helped to minimise ambiguity about which task was being performed on any trial, and to reduce any potential task-switching interference effects.

In each task, one key (the “correct” key) delivered rewards with 80% probability and errors with 20%, while the other key (the “wrong” key) had reversed probabilities. In both tasks the outcomes were associated with real rewards and penalties – receiving reward in each trial resulted in the addition of 1 penny to the final pay-out, whereas after an error 1 penny was subtracted from the bonus (but not from the basic hourly compensation). This ensured that both tasks were equivalent in their incentive, motivational quality, and only differed in the visual presentation of the outcomes (see below).

Each probability mapping was maintained until the participant had pressed the correct key for 5 to 7 times in a row (randomised), and then the mapping was reversed. Crucially, the two tasks were entirely independent, so each task had a distinct action-outcome contingency which determined the correct key to press at any given time. For example, if in task A the left key was the correct one, in task B the correct key could be right or left. Reversals were also independent between tasks, e.g. if participants pressed the correct key 6 times in a row in task A, then the mapping might reverse for task A, but not for task B. Instructions explicitly informed participants that such reversals would occur, and would be independent across tasks. Thus, participants were required to constantly and independently monitor both tasks in order to optimize their performance in each of them, to increase their final reward.

In task A (monetary reward), the rewarding outcome was a realistic picture of a £1 coin presented in the centre of the screen. If the outcome was erroneous, the presented coin was crossed with a grey, dim colour, ensuring that physical saliency of both positive and negative outcomes was equivalent. In task B (affective reward), photos of facial expressions were used

as outcomes (modified subset of The Karolinska Directed Emotional Faces; Lundqvist, Flykt, & Öhman, 1998). The set consisted of photos of 4 actors (2 males), each in positive (happy) and negative (sad) expression versions. Each participant saw pictures of only one randomly assigned actor throughout the procedure. In both tasks all visual outcomes were presented in the centre of the screen for 800 ms and were always accompanied by a sound delivered through speakers. The sound was a sinusoidal 600 Hz wave lasting 100 ms and was identical for both tasks and outcome types.

At the beginning of each trial a fixation cross was displayed, signalling to participants that they could press a key. The participant's task was thus to find out which key was correct in each of the tasks, and then keep pressing this key until they believed a reversal had occurred. After pressing a key, a 200, 500, or 800 ms delay followed, after which the outcome (picture + sound) was presented. After the picture disappeared from the screen, participants verbally reported the estimated duration of the delay, which constitutes the interval estimation procedure of IB. They were told that the delay could be any interval between 0 and 1000 ms and that their reports should be as accurate as possible. The estimate was entered by the experimenter using an external keyboard and displayed to the participant. No feedback was given about their accuracy of the estimate. Before the main procedure, participants underwent a training session, divided into three parts, due to the complexity of the main task. In part 1, participants were acquainted with temporal intervals covering the range of action-outcome intervals encountered in the task. In part 2, the basic PRL tasks (without interval estimation) were introduced, including explanation of scoring (reward) methods and practice with the key mappings used for each task. In part 3, the full task was presented and practiced, including verbal interval estimations. Practice sessions were terminated when participants reported they had understood the relevant task, and the experimenter had confirmed this. The main experiment consisted of 5 blocks of 120 trials, which took about 60 minutes to complete.

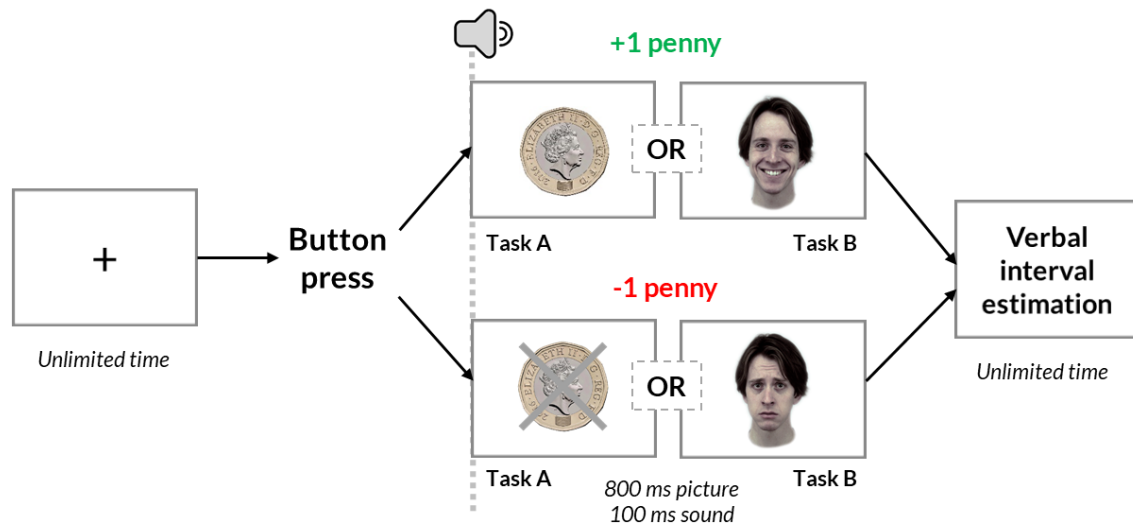


Figure 1. Experiment 1 procedure. Tasks followed in AABB order throughout the experiment.

Results

Two participants performed the PRL task at chance level (54% and 46%) and were thus excluded from further analyses. Trials with a temporal estimate error higher than 3 SD from each subject's mean were also removed from the dataset (0.27% of trials in total). Performance level (i.e. mean accuracy) did not differ between tasks: paired t-test $t(27) = -.58, p = .57, d = -.12$. Accuracy, reaction times, and specific means and standard errors in each cell of the experimental design can be found in Appendix.

The data were analysed using linear mixed-effects modelling using *lme4* (Bates, Mächler, Bolker, & Walker, 2015) and *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2017), and visualized using *ggplot2* (Wickham, 2016) packages for R (R Core Team, 2016). Error bars in all figures denote 95% CI of fixed effects uncertainty (Bolker, 2019). Linear mixed models were used because of the necessarily unbalanced numbers of win/loss outcomes, and in order to account for individual differences in both intercepts and slopes of specified effects in each subject. Parameter estimates β , together with 95% Wald confidence intervals, are reported to present the magnitude of effects, with t statistics and p values calculated using Satterthwaite approximation (Kuznetsova et al., 2017). More details regarding models can be found in the Supplementary materials.

The reported interval was modelled using the following fixed simple effects: actual interval (200, 500, 800 ms), previous outcome (win, loss), current outcome (win, loss), task type (monetary, affective), and type of trial transition (task-switch or task-repetition relative to previous trial). Fixed terms also included 2- and 3-way interactions between previous outcome,

task type, and type of trial transition, and between current outcome, task type, and type of trial transition (but not between current and previous outcomes, or between actual interval and any other factor). Finally, the effect of previous outcome was also modelled as a random effect varying across participants). In all models interval was treated as a metric predictor and mean-centered, while for all other dichotomous predictors effect coding (-0.5/0.5) was applied. False Discovery Rate (FDR) correction of p values has been used for follow-up of pairwise comparisons. We restrict our interpretations to the relevant comparisons only, i.e. those related to the hypotheses regarding effects of winning and losing outcomes. Accordingly, significant interactions are followed up in a way that allows to evaluate how error-related effects are influenced by additional factors, allowing a comprehensible interpretation of findings with respect to PEAB. Other interaction effects that do not involve a factor of outcome valence, as well as all remaining effects, are presented in the Supplementary materials. Figure 2 presents a schematic structure of trials and reported effects.

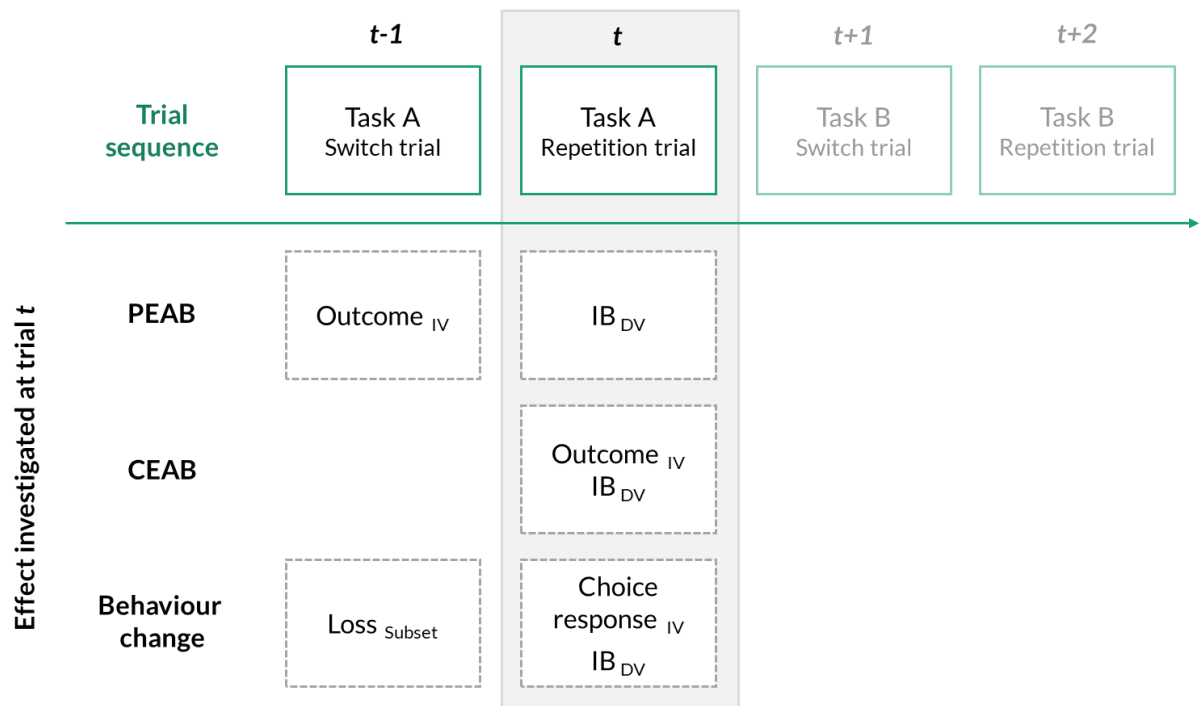


Figure 2. Scheme of trial types, and effects of interest. PEAB: post-error agency boost; CEAB: current-error agency boost; IB: intentional binding; DV: dependent variable; IV: independent variable. For simplicity, the scheme shows the definition of the effects of interest for trial t only. PEAB and CEAB effects can also be calculated when trial t is a switch trial, and for task B. However, our definition of behaviour change is limited to within-task response-switches following a loss, so applied onto when trial t is a repetition trial.

The model (Figure 3) revealed an unsurprising, significant effect of the actual interval: $\beta = 0.45$, $t = 105.98$, $p < 0.001$, 95% CI = [0.44 – 0.46]. All main effects were significant: previous outcome, $\beta = 15.10$, $t = 4.82$, $p < 0.001$, 95% CI = [8.96 – 21.23]; current outcome, $\beta = 17.21$, $t = 7.98$, $p < 0.001$, 95% CI = [12.98 – 21.44]; task type, $\beta = -9.14$, $t = -4.15$, $p < 0.001$, 95% CI = [-13.45 – -4.82]; and trial type, $\beta = -8.99$, $t = -4.08$, $p < 0.001$, 95% CI = [-13.30 – -4.67].

Crucially, there were also three significant interactions. First, a previous outcome by trial type interaction, $\beta = -27.26$, $t = -6.32$, $p < 0.001$, 95% CI = [-35.71 – -18.81] followed up with pairwise comparisons presented shorter temporal estimates following previous losses than wins in task-repetition trials $\beta_{\text{contrast}} = -28.73$, z ratio = -7.61, $p < 0.001$, 95% CI = [-36.12, -21.33], but no such effect of previous outcome in task-switch trials $\beta_{\text{contrast}} = -1.47$, z ratio = -0.38, $p = 0.702$, 95% CI = [-8.97, 6.04]. Further, resolving the current outcome by task type interaction, $\beta = 14.86$, $t = 3.45$, $p = 0.001$, 95% CI = [6.42 – 23.30] presented smaller estimates following current losses vs. wins in the monetary task, $\beta_{\text{contrast}} = -24.64$, z ratio = -8.08, $p < 0.001$, 95% CI = [-30.61, -18.66] and, to a noticeably smaller degree, also in the affective task, $\beta_{\text{contrast}} = -9.78$, z ratio = -3.21, $p = 0.002$, 95% CI = [-15.75, -3.81]. Finally, a current outcome by trial type interaction, $\beta = 12.83$, $t = 2.98$, $p = 0.003$, 95% CI = [4.38 – 21.28] showed smaller estimates following current losses vs. wins in task-switch trials $\beta_{\text{contrast}} = -23.64$, z ratio = -7.78, $p < 0.001$, 95% CI = [-29.58, -17.67] and, to a smaller degree, also in task-repetition trials $\beta_{\text{contrast}} = -10.79$, z ratio = -3.53, $p < 0.001$, 95% CI = [-16.79, -4.8].

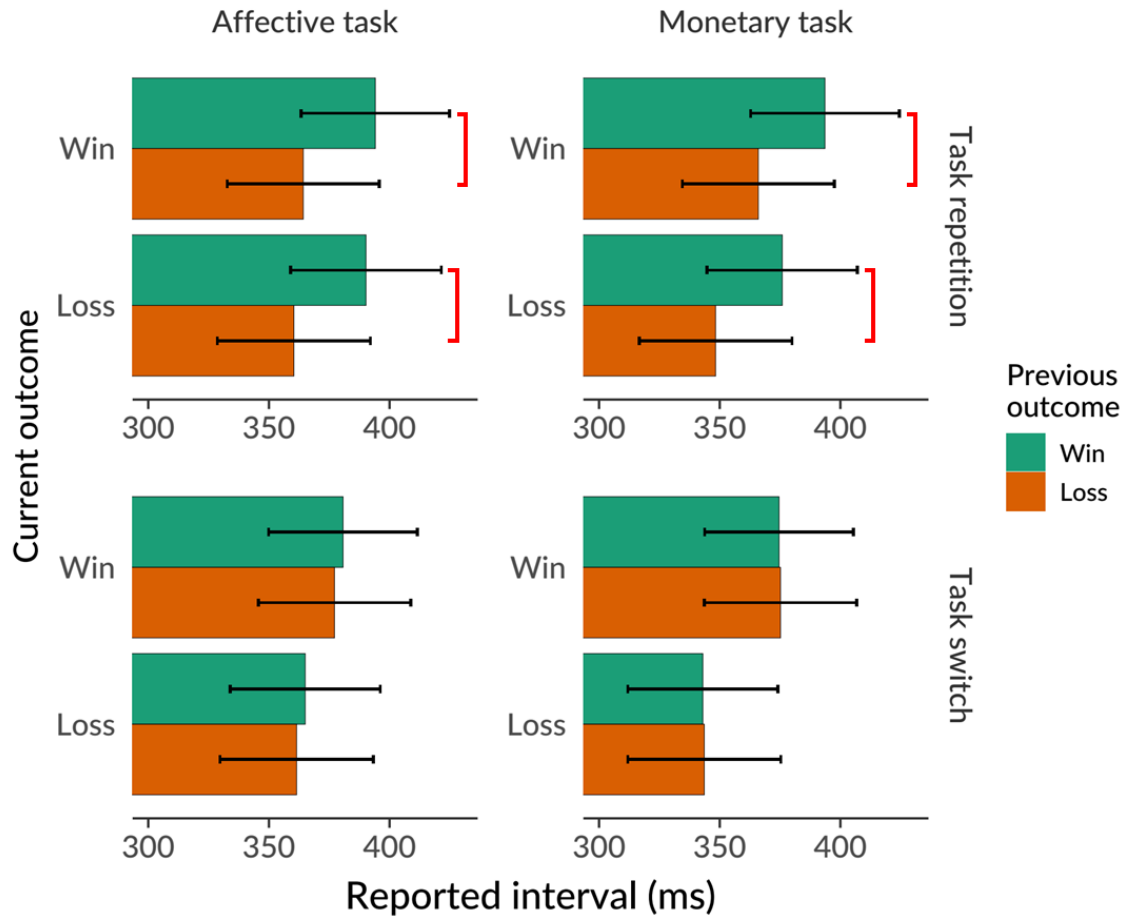


Figure 3. Effects of previous outcome, current outcome, task type, and trial type on interval estimates in Experiment 1. Lower interval estimates are interpreted as indicating a stronger sense of agency. PEAB (i.e., the difference between win and loss outcome in previous trials) is present for task repetition, but not for task switch trials. The contrast shown by the red bracket is the PEAB. Error bars denote 95% CI.

To assess whether PEAB was related to adaptive behaviour, we selected the subset of data that included loss on a previous trial and then repetition of the same task on the current trial. This comprised 19.4% of trials. A simple linear mixed-effects model was then fitted, predicting reported interval on the current trial based on the choice response (switched vs. stay compared to the preceding trial). This did not show a significant effect of the button press, $\beta = -9.77$, $t = -1.42$, $p = 0.156$, 95% CI = [-23.28 – 3.73]. As such, switching the response after a loss did not strongly influence the magnitude of the PEAB.

Discussion

In the first experiment we replicated the effect of increased intentional binding following losses using linear mixed-effects modelling. This PEAB was visible after receiving a negative outcome in the previous trial, but only when the current trial was a repetition of the same task that had produced the negative outcome. No transfer of the effect across tasks was observed: getting a negative outcome in one task did not lead to PEAB if the next trial involved another task. This result shows that PEAB is not a task-general low-level phenomenon related to attention, arousal or emotional reactions to losses. Instead, PEAB depends on features *specific* to the task at hand. This means that only opportunities for functional learning, that is to adjust behaviour based on mistakes made within the current task, influences SoA on that task. This finding is consistent with several possible cognitive theories relating learning and agency. One obvious account involves a continuous, implicit monitoring of the “agentic contrast” between the negative outcome of a previous action, and the potentially positive outcome of the next attempt, given that learning and behavioural adaptation may allow improved performance in subsequent attempts. In other words, PEAB might reflect an increased sense of agency when people have an opportunity, through their action, to make an adjustment that matters.

Follow-up analyses investigated whether PEAB was related to actual change in behaviour, as marked by a response switch compared to the preceding trial. No significant difference was observed. Losing in the previous trial ($t-1$) increased IB in the following trial (t), irrespective of whether losing also led to a switch of response strategy or not. This suggests that PEAB is not an artefact of loss-induced response switching but is indeed an effect of previous loss that persists into subsequent trials. However, this interpretation should be seen as tentative, since it is based on a null result and a small number of trials.

We also found a *current-error* (loss) agency boost (or CEAB), in addition to the previously-reported PEAB. This CEAB exhibited a different pattern of modulation than did PEAB: it was especially strong when a loss occurred in task-switch trials, and weaker in task-repetition trials. Moreover, CEAB was more pronounced in the monetary compared to the affective task. This suggests that IB of the current outcome is more influenced by visual and affective features of task presentation from which valence information has to be decoded. However, once outcomes have been identified as wins or losses, their influence on the subsequent task, in terms of PEAB, is not further affected by previously carried affective/visual features. This indicates that PEAB, compared to CEAB, may rely on a more abstract representation of outcome value which is relevant for guiding actions in a learning environment.

This differential pattern of PEAB and CEAB suggests that these outcome-dependent effects on intentional binding are each driven by different mechanisms. Since the CEAB shows a valence-dependent pattern that is opposite to the effects reported in previous studies, in which agency was commonly “boosted” after wins, it might be part of a specific error-based learning mechanism, possibly playing the role of a precursor or complementary part of the PEAB effect. After all, some process of error-processing has to take place before downstream effects on agency and behaviour can constitute the PEAB in the subsequent trial. While PEAB might rely more on predictive processing, CEAB might be driven by a retrospective inference triggered by the outcome of the current trial. This possibility is consistent with previous research (Di Costa et al., 2017; Moore & Haggard, 2008), and in general supports the notion of SoA resulting from the combination of multiple predictive and inferential signals (Sidarus, Vuorre, & Haggard, 2017b; Synofzik, Vosgerau, & Voss, 2013).

Taken together, Experiment 1 revealed the task-specificity of PEAB, showing that it only occurs between two trials of the same task. Further, the lack of difference between tasks indicated that the affective value of outcomes is not a crucial factor for evoking PEAB, but it does influence CEAB.

EXPERIMENT 2

A key factor in any learning process is the ability to use outcomes to discover the structure or rules of the task, which can in turn inform future decisions. Negative outcomes only have such ‘epistemic value’ if (a) they are informative about the rules of the task at hand, and (b) if the agent is able to use this information to guide subsequent action choices. Experiment 2 investigated this second aspect, namely the requirement of free action selection. We compared PEAB between trials where participants could freely choose between alternative actions, and a forced choice condition, where they were told which action to make. Free and forced trials followed an AABB sequence. We could thus test whether free choice influences the PEAB and whether free choice is required on the trial that resulted in the loss, on the subsequent trial (that might benefit from any error-driven learning), or on both. In order to make both tasks equally informative and promote learning, both free and instructed tasks shared the same action-outcome contingency (in contrast to Experiment 1). This allowed participants to use outcomes from one task to learn about the structure of the other task.

In Experiment 2, we additionally recorded EEG to preliminarily investigate electrophysiological components related to PEAB. We were particularly interested in topographically central components implicated in outcome monitoring, which have previously

been shown to play a role in coding (reward) prediction error and to be involved in outcome evaluation, i.e. the Feedback-Related Negativity (FRN), and Feedback P300 (P3) (Glazer, Kelley, Pornpattananankul, Mittal, & Nusslock, 2018). FRN is a fronto-central component peaking around 250 ms after outcomes onset, being the most negative after negative, unexpected feedback (San Martín, 2012). Previous work showing enhanced FRN related to the presence (vs absence) of response choice (Yeung, Holroyd, & Cohen, 2005) suggests that this component may present a differential pattern of modulation in free- and instructed-choice tasks employed in Experiment 2. The FRN has further been linked to post-decisional monitoring and sense of agency ratings over action effects (Sidarus, Vuorre, & Haggard, 2017a). P300, a positive, central deflection peaking between 300-500 ms following a stimulus feedback, reflects categorization of outcome-related information (Polich, 2007), and may differ depending on whether this information could be used to drive future actions. P300 has also been shown to be implicated in agency judgements (Kuhn et al., 2011). We aimed to explore the link between the above-mentioned components and loss-dependent enhancement of binding effects, expecting that both components would be differently modulated by PEAB.

Method

Participants

30 new volunteers (9 males) participated in the study, with a mean age of 24.4 ($SD = 4.1$). All participants gave informed consent and were compensated at the rate of £7.5 per hour plus an additional bonus depending on their performance. All had normal or corrected-to-normal vision and hearing. The research was approved by the UCL Research Ethics Committee.

Setup and procedure

Participants were examined one at a time in a quiet room. The procedure was run on a laptop computer with a 15.6-inch display using Matlab, with external keyboards used by the participant and experimenter. The EEG setup consisted of a customized 16+6 channels Biosemi system, with electrodes focused on midline scalp locations: F3, Fz, F4, FC3, FCz, FC4, C3, Cz, C4, CPz, P3, Pz, P4, POz, O1, and O2.

Two PRL tasks alternated in AABB order, with rewards delivered at 80:20 probability (Figure 4). Both tasks now shared a single action-outcome contingency, but two separate sets of keys for each task were still used. That is, two keys were available to press in each task: Z/X and N/M as possible left and right response keys. For example, task A used Z and X with the left middle and left index fingers, while task B used N and M with the right index and right

middle fingers. If, at a given moment, Z happened to be the more rewarding key for task A, then and only then, N was the more rewarding key for task B. This ensured that participants could learn the single action-outcome contingency (in this example “prefer the leftmost key”) from both tasks equally, while maintaining motor requirements comparable between the two experiments. Outcome probabilities of the correct/wrong cues in the instructed task were programmed to produce accuracy/error rates based on those we found in Experiment 1, i.e. the cue was correct in 67% of trials in this task. Key set and task mapping were counterbalanced across subjects, as was a green and blue coloured background associated with each task. The correct/wrong key mapping was maintained until the participant had pressed the correct key in 5 to 7 consecutive trials and after that the mapping was reversed. Reversals of the mapping always applied to both tasks at the same moment.

In both tasks, trials started with a black square presented on the screen for 800 ms. In task A (free choice), a grey cross was next presented atop of the square, signalling that the participant could now make an action of their choice by pressing the left or right key. In task B (instructed choice), either the left or right half of the cross was presented instead of a full cross, instructing participants which of two available keys they should press. After pressing a key, the outcome followed after a delay of 500, 800, or 1100 ms. We used longer delays than in Experiment 1, to allow for proper time-locking of the EEG data. In both tasks the outcomes were identical to the ones in the monetary task in Experiment 1, i.e. they consisted of either a full or a crossed coin picture accompanied by a sound. After the outcome was presented, participants verbally reported the temporal estimates of delays which were written down by the experimenter.

Instructions informed participants that they should press the key according to the instructional cue, regardless of whether they thought it was the correct key to press or not. It was stressed that they could learn from this task and use its outcomes to figure out which key they should choose in subsequent free choices. They were also told that the reported delay could be any interval between 100 and 1500 ms. The main part of the study consisted of 5 blocks of 100 trials each and was preceded by a training session similar to that of Experiment 1.

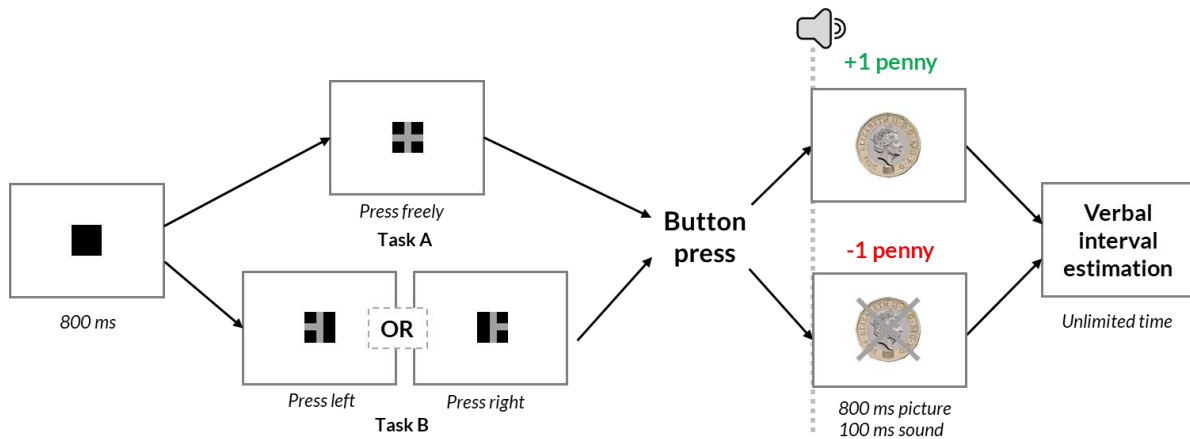


Figure 4. Experiment 2 procedure. Tasks followed in AABB order throughout the experiment.

Results

Behavioural data

Out of all trials 0.38% were removed as being more than ± 3 SDs from subject's mean temporal estimate. Performance in the free choice task was above chance level for all participants. Performance level differed significantly between tasks, with higher accuracy in the free ($M = 72.6\%$) than in the instructed task (where the programmed accuracy of instructions was set in advance, at $M = 67.5\%$): paired t-test $t(29) = 3.87$, $p < .001$, $d = 1.2$. Summary of data split by the design factors can be found in the Appendix. Data was analysed using linear mixed effects models, with fixed and random terms identical to those used in Experiment 1 analyses. More details regarding the models can be found in the Supplementary materials.

The model (Figure 5) revealed a significant effect of the actual interval, $\beta = 0.53$, $t = 103.43$, $p < 0.001$, 95% CI = [0.52 – 0.54]. All main effects were significant, previous outcome, $\beta = 19.33$, $t = 4.50$, $p < 0.001$, 95% CI = [10.91 – 27.76]; current outcome, $\beta = 28.64$, $t = 8.32$, $p < 0.001$, 95% CI = [21.89 – 35.40]; task type, $\beta = 15.45$, $t = 4.37$, $p < 0.001$, 95% CI = [8.53 – 22.38]; and trial type, $\beta = -11.77$, $t = -3.33$, $p = 0.001$, 95% CI = [-18.69 – -4.84].

Further interactions were also significant. A current outcome by task type interaction, $\beta = 19.44$, $t = 2.82$, $p = 0.005$, 95% CI = [5.92 – 32.96], presented that temporal estimates were shorter after current losses than wins in the free task $\beta_{\text{contrast}} = -38.36$, z ratio = -7.81 , $p < 0.001$, 95% CI = [-47.99, -28.74], but also to a smaller degree in the instructed task $\beta_{\text{contrast}} = -18.92$, z ratio = -3.91 , $p < 0.001$, 95% CI = [-28.41, -9.44]. Another significant interaction of current outcome and trial type, $\beta = 23.10$, $t = 3.35$, $p = 0.001$, 95% CI = [9.59 – 36.60], showed that estimates were shorter after current losses than wins in switch trials $\beta_{\text{contrast}} = -40.19$, z ratio =

-8.24, $p < 0.001$, 95% CI = [-49.75, -30.64], and also to a smaller degree after current losses than wins in repetition trials $\beta_{\text{contrast}} = -17.1$, z ratio = -3.51, $p = 0.001$, 95% CI = [-26.64, -7.56]. These interactions present that temporal estimates were shorter after current losses than wins in the free task, and after current losses in task-switch trials than in task-repetition trials.

More interestingly, a three-way interaction of previous outcome, task type, and trial type was also significant, $\beta = -51.21$, $t = -3.71$, $p < 0.001$, 95% CI = [-78.24 – -24.19], with 1) lower estimates following previous losses than wins in repetition trials of free task, but not in switch trials of free task, and also 2) lower estimates following previous losses than wins in switch trials of instructed task, but not in repetition trials of instructed task. In other words, in the free task temporal estimates were shorter after previous losses than wins only in task-repetition trials (i.e., when the previous, loss-generating trial was also free). Further, in the instructed task previous losses led to more negative estimates only in task-switch trials (i.e., when the previous, loss-generating trial was free). In short, PEAB effect was observed as long as the *previous*, loss-generating trial involved a free choice. Whether the current trial, where PEAB was expressed, was free or instructed did not matter.

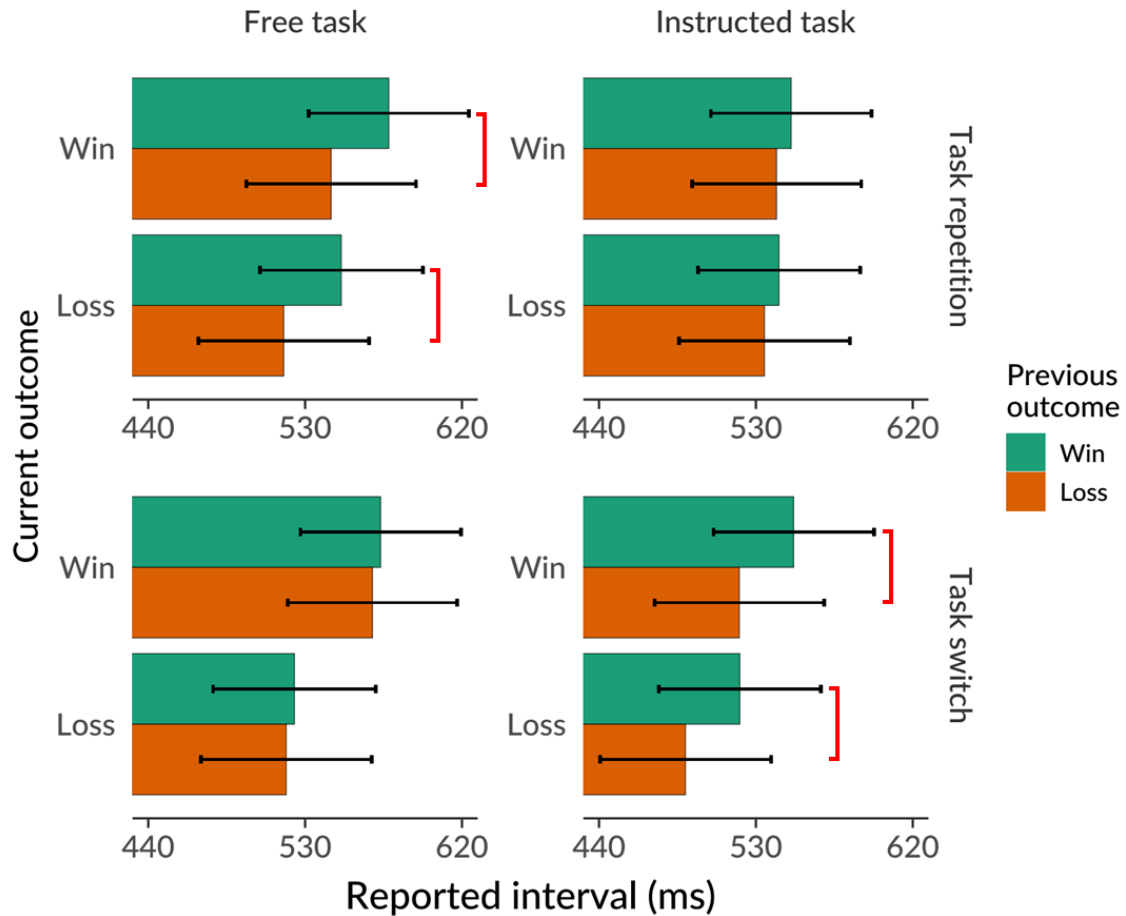


Figure 5. A previous outcome, current outcome, task type, and trial type influences on the time estimates in Experiment 2. PEAB and CEAB effects are the difference between win and loss outcome in previous and current outcome condition, respectively. The contrast shown by the red bracket is the PEAB. Error bars denote 95% CI.

We next tested whether these changes in IB were related to behaviour change. Only trials involving a previous loss and a task repetition were used (19.2% of trials). A simple linear mixed-effects model predicting reported interval based on the choice response (switched vs. stay compared to the preceding trial) was fitted. The effect of the choice response was not significant, $\beta = -3.02$, $t = -0.29$, $p = 0.772$, 95% CI = [-23.43 – 17.39]. Therefore, the PEAB did not appear simply to be artefact of loss-induced decisions to switch or stay, but rather a direct result of the previous loss itself.

EEG data

Two participants were lost from the EEG dataset for technical reasons: one due to equipment failure during recording, and another due to failure to record correct trigger signals.

Pre-processing of the data was done in EEGLab for Matlab (Delorme & Makeig, 2004). Raw EEG recordings were re-referenced to the scalp average, filtered using low-pass 40 Hz, high-pass 0.5 Hz, and notch 50 Hz filters, and downsampled to 256 Hz. Continuous data was manually cleaned from block breaks and major movements, and subjected to Independent Component Analysis used to remove eyeblinks and other typical signal artefacts. Data was then epoched from -0.5 to 1.5 s relative to outcome presentation with the baseline (average amplitude between -200 to 0 ms) removed. Epochs were cleared using automatic EEGLab algorithms and average amplitudes per subject and condition were exported to R. Outcome-locked time-courses of the signal at Cz and CPz electrodes are shown in Figure 6.

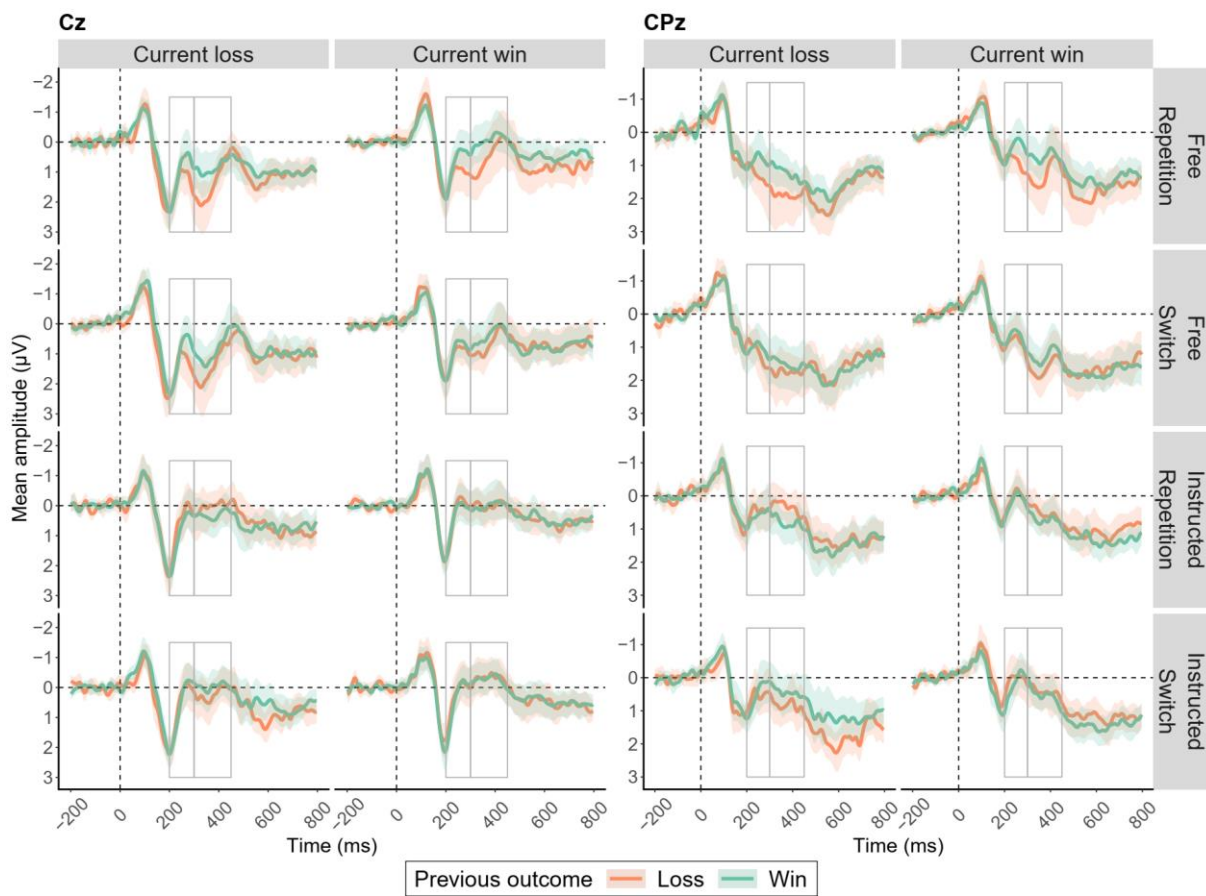


Figure 6. Outcome-locked time-course of EEG signal at Cz and CPz electrodes, with FRN and P3 time windows marked by grey boxes. Error ribbons denote 95% CI.

Data were subjected to a linear mixed effects model closely based on that used for modelling the behavioural data: the fixed effects term consisted of previous outcome, current outcome, task type, and trial type together with key interactions (identical to those used in behavioural data models of both experiments). The action-outcome interval was not included due to time-locking data to the onset of the outcome (i.e. the action-outcome interval had

already elapsed by the time of the event of interest). Due to better convergence of the model the random term included current (rather than previous) outcome, varying by subjects. As subject averages were used instead of single trial data, the number of trials per condition was added to the model's weights argument, thus accounting for necessarily unbalanced numbers of win/loss outcomes.

FRN component. The average amplitude of the Feedback Related Negativity (FRN) was determined in a 200 to 300 ms outcome-locked time window, and pooled over electrodes Fz, FCz, Cz, and CPz.

Modelling revealed a significant main effect of task, $\beta = -0.35$, $t = -6.69$, $p < 0.001$, 95% CI = [-0.45 – -0.25] and a previous outcome by task interaction, $\beta = 0.26$, $t = 2.51$, $p = 0.012$, 95% CI = [0.06 – 0.46], showing that in free task errors were followed by a more positive (weaker) FRN compared to wins $\beta_{\text{contrast}} = 0.16$, $t \text{ ratio} = 2.18$, $p = 0.039$, 95% CI = [0.02, 0.3], which was not the case in the instructed task $\beta_{\text{contrast}} = -0.1$, $t \text{ ratio} = -1.37$, $p = 0.171$, 95% CI = [-0.24, 0.04], where the pattern was somewhat reversed (Figure 7).

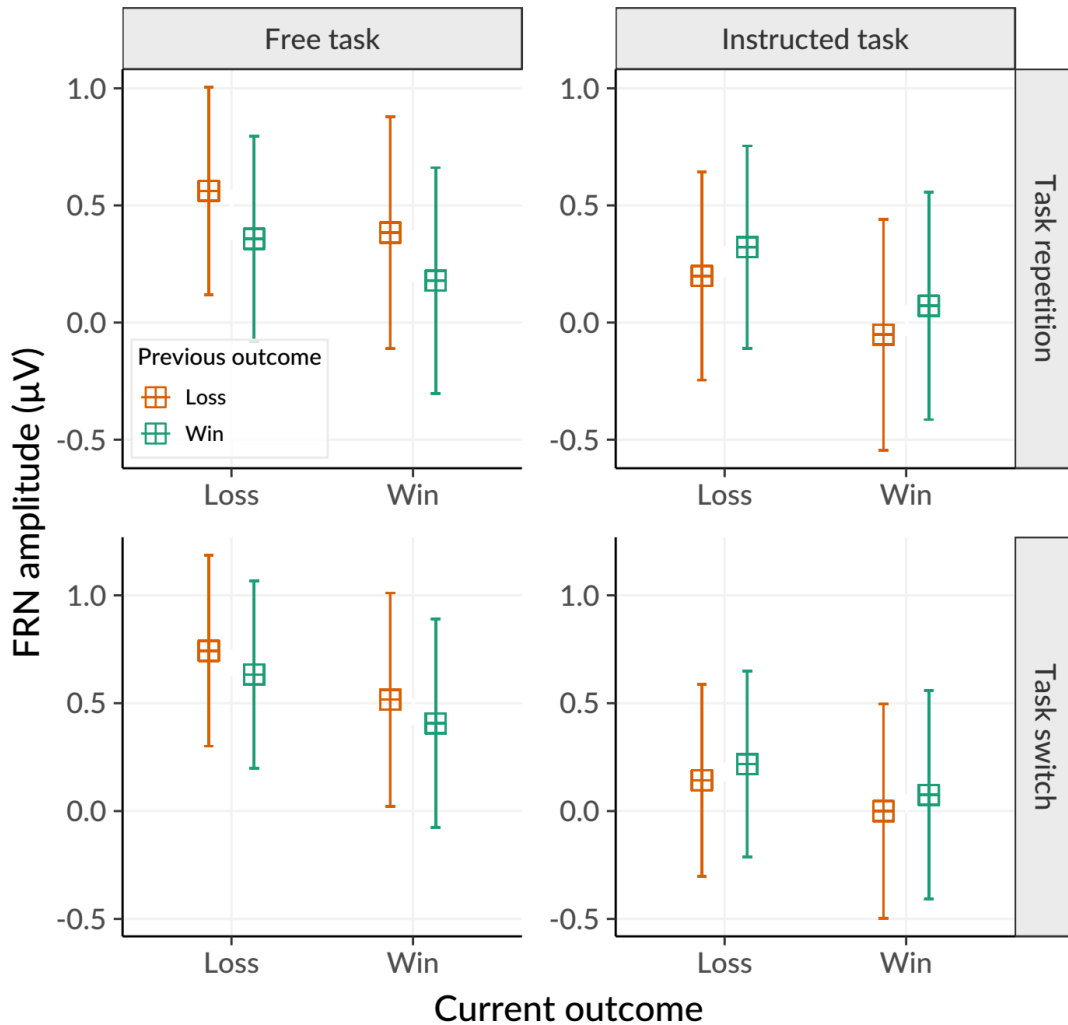


Figure 7. A previous outcome, current outcome, task type, and trial type influences on the amplitude of the FRN component. Error bars denote 95% CI.

P3 component. The average amplitude of the P3 component was computed across a 300 to 450 ms outcome-locked time window, and pooled over electrodes Fz, FCz, Cz, CPz, and Pz.

Modelling showed significant main effects of previous outcome, $\beta = -0.16$, $t = -2.99$, $p = 0.003$, 95% CI = [-0.26 – -0.05], current outcome, $\beta = -0.33$, $t = -3.67$, $p < 0.001$, 95% CI = [-0.51 – -0.16], and task type, $\beta = -0.63$, $t = -11.64$, $p < 0.001$, 95% CI = [-0.73 – -0.52].

Further, a previous outcome by task type interaction, $\beta = 0.37$, $t = 3.49$, $p < 0.001$, 95% CI = [0.16 – 0.57], presented higher amplitudes following previous losses than wins in free $\beta_{\text{contrast}} = 0.34$, t ratio = 4.6, $p < 0.001$, 95% CI = [0.19, 0.48], but not instructed task $\beta_{\text{contrast}} = -0.03$, t ratio = 0.72, $p = 0.72$, 95% CI = [-0.17, 0.12]. A current outcome by task type interaction, $\beta = 0.31$, $t = 2.92$, $p = 0.003$, 95% CI = [0.10, 0.51], also showed higher amplitude following current losses than wins in free $\beta_{\text{contrast}} = 0.49$, t ratio = 4.58, $p < 0.001$, 95% CI = [0.28, 0.7],

but not instructed task $\beta_{\text{contrast}} = 0.18$, t ratio = 1.72, $p = 0.086$, 95% CI = [-0.03, 0.39]. This shows that the P3 amplitude changed similarly due to previous and current outcomes and was stronger following losses than wins in free task only (Figure 8).

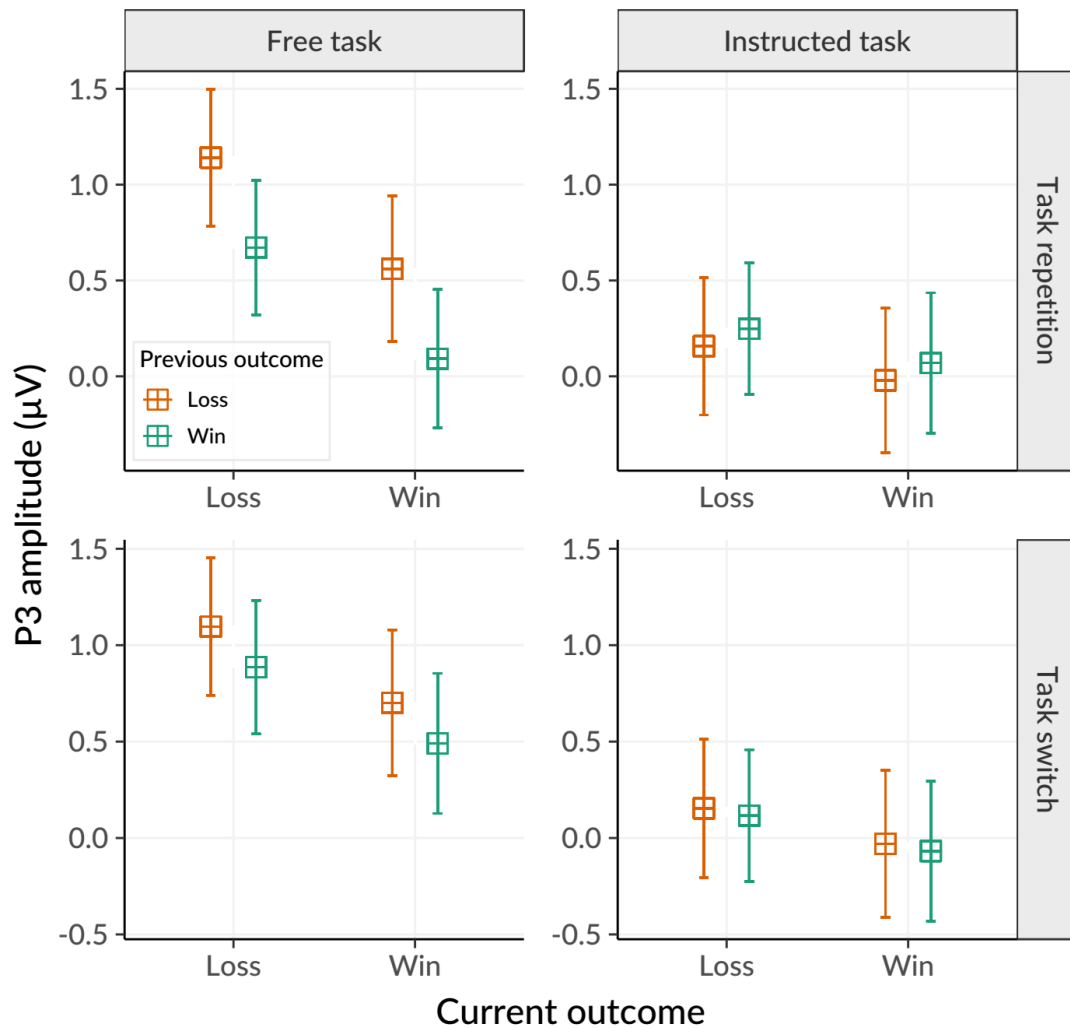


Figure 8. A previous outcome, current outcome, task type, and trial type influences on the amplitude of the P3 component. Error bars denote 95% CI.

Discussion

In Experiment 2 we again replicated the general effect of stronger intentional binding following losses. Similar to Experiment 1, PEAB was pronounced in the task-repetition trials of the free task. Intriguingly, PEAB also occurred in task-switch trials of the instructed task. That is, binding was stronger on instructed trials when a previous free choice had produced a loss, compared to when a previous free choice had produced a win. This clear differential pattern shows that: (1) free choice in the preceding loss-producing trial is necessary to elicit

the PEAB effect, and (2) the epistemic value of the previous freely-produced outcomes can lead to PEAB even when the current action is instructed, rather than freely-chosen.

We also again found a CEAB effect, as binding was stronger on trials which produced a loss than on trials which produced a win. CEAB was especially pronounced in the free task, although still observable in the instructed task, and was especially strong in switch trials, but still significant in repetition trials. This result implies that CEAB is a rather general effect, although free choice over actions seems to enhance it.

EEG data showed that two well-known ERP components can be linked to these effects. First, the FRN component on the current trial was decreased after a loss on the previous trial in the free task, but not after a loss in the previous trial in the instructed task. In contrast, the outcome of the current trial did not influence FRN. This suggests that the FRN can be linked to the PEAB effect, and reflects a mechanism representing downstream effects of previous, but not current, losses. Second, we also observed P3 effects – the component was increased if the previous trial was a loss vs. a win, but only if the current task involved a free as opposed to an instructed action. The P3 was also increased when a negative vs. positive outcome was received on the current trial, and this effect was also limited to the free-choice task. These parallel patterns of PEAB and CEAB regarding the P3 component suggest that the P3 does not distinguish between these two behavioural effects. Instead, it could be related to a more general mechanism of prediction error coding that accompanies outcome valence processing, common to both PEAB and CEAB.

GENERAL DISCUSSION

It is common to learn from our mistakes. If we learn to adjust our actions to achieve a desired outcome, our experience of both actions and outcomes changes and we typically experience an increased sense of agency (Metcalfe and Green, 2007). Despite these straightforward intuitions, the exact relationship between agency and learning from errors has hardly been studied. Intentional binding, one implicit measure of SoA, is stronger in the trial following a loss, than in a trial following a win, constituting a post error agency boost, or PEAB effect (Di Costa et al., 2017). Losses can convey important information regarding the structure of the task (as in probabilistic learning setting), and thus can inform future decisions by helping agents to choose subsequent actions which maximize their rewards. Thus, PEAB has been argued to reflect an important mechanism supporting adaptive behaviour. In two experiments, we have further investigated in what conditions PEAB is most pronounced, and through this managed to clarify the mechanisms and functional relevance of this effect.

First of all, we established that PEAB is a task-specific phenomenon. That is, when action leads to a negative outcome in a specific task, and the next trial involves a different, new task, then PEAB is abolished, and binding is not increased on the new task. That is, the PEAB effect does not transfer or propagate across tasks. Thus, domain-general processes such as attention, affect, arousal, motivation and global, phasic dopaminergic activation, which have previously been shown to be implicated in appetitive behaviour linked to learning (Wanat, Willuhn, Clark, & Phillips, 2009) can be ruled out as a mechanism underlying PEAB. The observed task specificity speaks for a certain flexibility of the PEAB effect, as the underlying mechanism must be selective to the specific action-outcome relation that is learned. A continuity between the outcome of the previous trial and the next trial allows feedback information to optimize behaviour. As such, the PEAB effect reflects a high-level agentic process devoted to the fine-grained optimizations dependent on and constrained by the content of the current task, which helps to enhance future performance in this very task.

It seems that PEAB can develop only when continuity between previous outcome and current trial is sustained. Previous research (Di Costa et al., 2017) described PEAB only for the action binding component of intentional binding, which is thought to be driven primarily by predictive processes (Moore & Haggard, 2008). We found that the continuity across trials can be interrupted in the case of a switch to a different action-outcome contingency, thereby preventing the PEAB effect.

We also found an agency boost due to losses in the current trial, which we referred to as CEAB (current-error agency boost). Since the CEAB was only partially affected by the continuity from previous trials, we speculate that it may primarily rely on inferential processing based on just-received feedback that a loss has just occurred. This might include the binding of outcomes back towards the actions that caused them (again in line with Di Costa et al., 2017). We add that these remarks remain speculative, because the present study measured IB using interval estimation, which cannot separate the binding of actions towards outcomes from the binding of outcomes towards actions.

Learning and adaptive behaviour

With task-specificity of PEAB established in Experiment 1, we further aimed to answer the question which specific contents, or parameters of the task, influence the PEAB effect. Experiment 1 also included a manipulation of outcome content (affective, monetary), which had no influence on PEAB (while affecting CEAB). Taken together, this suggests that PEAB is not affected by sensory or affective outcome specifics, but rather reflects a higher-level

process that integrates a previously registered outcome valence for subsequent behavioural adaptation. Accordingly, Experiment 2 showed that the informational or epistemic value of outcomes indeed does play a major role for PEAB to occur. Here, participants were confronted with only one action-outcome contingency at a time – thus, there was a single ‘task’ or ‘rule’ in the sense of Experiment 1. Instead, in Experiment 2, free-choice and instructed actions alternated in an AABB design. In Experiment 2, PEAB occurred in repetition trials of the free-choice condition, but also in instructed trials which were preceded by a free-choice trial. Thus, the informational value of previous freely produced losses was sufficient for a PEAB to occur on the subsequent trial, even if at that point the agent no longer had a free choice of what action to make. To put it differently, PEAB occurred when the agent had free control over what action to make, when this action led to an error, and this error was epistemically informative and could potentially be used for optimizing behaviour – even if in fact there was no actual possibility to use error information on the current trial. In sum, Experiment 2 this established three necessary conditions for PEAB: (1) Loss encoding (last trial led to a perceivable loss); (2) The action that produced the loss was freely chosen (3) Loss was epistemically relevant (last trial was informative about the tasks structure, enabling learning). As such, the task specificity of the PEAB effect observed in Experiment 1 can be taken as a special case of epistemic relevance, since an error made in previous task A did not provide information about the rule of current task B if a task switch had occurred, thus lacking the continuous informational relevance required for the PEAB effect to occur.

Importantly, PEAB was present also when the current trial did not involve any choice of action. For example, when a loss on a free-choice trial was followed by an instructed trial, we found stronger IB on the instructed trial, compared to a win on a free-choice trial followed by an instructed trial. This shows that sense of agency strongly depends on the *possibility to learn*, or *test what has just been learned*, even if this learning cannot immediately guide action choices. PEAB on instructed trials following free-choice trials is all the more surprising, because several previous studies showed that IB is generally reduced on instructed trials compared to free-choice trials (Barlas, Hockley, & Obhi, 2017; Borhani, Beck, & Haggard, 2017; Caspar, Christensen, Cleeremans, & Haggard, 2016). Our results show that the informativeness of a previous negative outcome can boost IB, even when the information in question cannot guide immediate behaviour. As such, the PEAB is less indicative of a general feeling of agency occurring during a sequence of unrelated button-presses, but is rather related to an ongoing process that spans multiple trials and monitors a changing rule in the context of a learning environment.

Exploratory follow-up analyses tested whether PEAB was related to an actual change in behaviour, such as a “switch after loss” strategy. In both experiments, non-significant results were found. That is, we did not find evidence linking the PEAB to the specific response choices participants made after a loss. While caution is required in interpreting these null results, they may indicate that PEAB reflects an adaptation of the experience of agency, rather than adaptation of behaviour.

PEAB cannot readily be explained by confounding factors such as the attentional or motoric demands of a novel response. Instead, PEAB appears to represent a delayed epistemic effect related to a previously encountered loss. Nevertheless, we note that a numerical trend pointed toward PEAB being related to a response switch (especially so in Experiment 1). Further, these analyses were based on just a small subset of the data, so may be prone to type II errors. Our design was not optimized to detect possible adaptation effects that either depend on accumulation over multiple losses or are expressed by changes in future behaviour beyond the immediate impending action choice. Clearly, “switch after loss” is just one example of a behavioural adaptation, and we cannot exclude the possibility that PEAB is associated with other, more subtle behavioural adaptations. Thus, we cannot rule out some association between PEAB and behavioural adaptation. Further studies need to clarify this issue with designs optimized to sequential analyses of this kind (see Future directions section).

Alternative explanations and confounds

As previously mentioned, some of the effects observed in our study are inconsistent with previous findings showing diminished binding when actions produce losses or negative outcomes (Takahata et al., 2012; Yoshie & Haggard, 2013). In particular, our CEAB appears directly opposed to the effects of negative outcomes in those studies. However, those studies did not involve adaptive guidance of action choices and did not enable reward-based learning. In our view, these are critical factors allowing agents to adapt their performance and thus form the basis for PEAB-like effects. Furthermore, valence-dependent modulations of IB were not found at all in a series of 4 experiments conducted by Moreton, Callan, & Hughes (2017), despite observing standard IB – the study, however, also lacked the element of dynamic reward-based learning. Some other studies, however, did report increased binding for negative, as opposed to positive, outcomes. For example, Moretto, Walsh, & Haggard, (2011) reported enhanced binding following negative moral outcomes in the choice dilemma procedure, which was especially pronounced in severe, as compared to moderate, outcomes. In this task participants were able to choose whether they wanted to stay or change the current setting of

dilemma, so such enhanced binding might perhaps reflect salient information about one's own affective response to severe outcomes, which could adaptively guide moral decisions. This resonates with findings that IB increases with the number of salient alternatives, both in terms of available actions and obtainable outcomes (Kulakova, Khalighinejad, & Haggard, 2017). Saliency or predictability of the outcomes in itself may also play a role here, as enhanced binding has been reported following unpredictable oddball-like outcomes (Majchrowicz & Wierchoń, 2018), also interacting with their valence (Christensen, Yoshie, Di Costa, & Haggard, 2016).

Could the effects observed in this study simply reflect differences in outcome expectation? Stronger binding is found when an outcome is predictable, compared to when it is not (Beck, Di Costa, & Haggard, 2017; Hughes, Desantis, & Waszak, 2013). Duration estimation in the absence of action shows similar effects (Pariyadath & Eagleman, 2007). Thus, outcome predictability should increase binding. In the current task, positive outcomes were generally more frequent than negative ones, presumably leading to an overall expectation of wins, rather than losses. Nevertheless, we found CEAB when this expectation was violated by a loss. Similarly, on the trial after a loss, expectations were most likely reduced even further (due to increased uncertainty about whether the loss signalled a rule change). In fact, however, this produced an increase in binding, i.e. PEAB. Thus, the boosts in agency that we have found are more likely linked to factors such as expectation violation, or surprise, rather than expectation alone. Further, in our Experiment 2, outcomes in both free and instructed task were similarly predictable (due to generally high accuracy and the single underlying action-outcome rule), yet task differences in binding were observed with stronger binding in free vs. instructed trials. These differences are more readily explained by the epistemic value of the error and/or how it could be used by the subjects in their subsequent action, than by mere expectation.

This raises the question of whether PEAB is an automatic consequence of having previously made an error, or rather it is related to how an agent interprets that error, and its role in guiding future behaviour. Is it more related to an ongoing, background process of monitoring errors, or to the use of error information to regulate the current action (cf. Botvinick, 2007)? We found that losses on instructed trials could not generate PEAB, while losses on free trials could. The epistemic value of errors appeared in a context that allowed learning, suggesting a likely function in adapting future behaviour. Crucially, PEAB was present on instructed trials following a loss on a free trial. This is a situation where learning can occur, but its current behavioural expression is not possible. Our result provides an intriguing suggestion that learning may boost the subjective experience of agency, even when

factual agency, understood as the ability to control outcomes, is decreased, rather than increased. PEAB reflects a boost in sense of agency based on the future potential of optimising action choices, even if the better choices cannot currently be implemented.

Alternative account of the current results could refer to response caution driven by error registration, as in post-error slowing effect (Danielmeier & Ullsperger, 2011; Laming, 1968). In post-error slowing significant increases in reaction times can be observed after negative feedback, reflecting enhanced response caution following losses (Dutilh, Vandekerckhove, et al., 2012). In such a view, PEAB could be seen as a result of boosted attentional processing and perhaps more cautious action selection, leading to an increased feeling of control and agency. Interestingly, a recent study employing a task-switching paradigm (Steinhauser, Ernst, & Ibal, 2017) has shown a post-error slowing effect not only on the same trial, but also on subsequent trial of the given task, which could be linked to the PEAB effect in repetition trials of our experiments. However, analyses of reaction times did not show significant decreases due to losses in preceding trials: in fact, the numerical effect was in the opposite direction. That is, we found a trend toward post-error *acceleration* of reactions (in analyses based on mixed-effects and PSE_{robust} method (Dutilh, Van Ravenzwaaij, et al. 2012), see Supplementary materials). The current design enforcing continuous switching between tasks could also disturb attentional and/or response-related processing, and thus influence results particularly dependent on the comparison of repetition and switch trials, but by employing alternating runs of the tasks (with supporting cueing) and also separated response sets we minimized between-tasks interference effects (Gade & Koch, 2007; Yeung & Monsell, 2003).

Event-related potentials

EEG allowed us to investigate two ERP effects related to outcome monitoring which co-occurred with the behavioural effects. The first component was the FRN, which is a commonly observed negative deflection following negative feedback in reinforcement learning tasks, and has been interpreted to reflect reward prediction errors signalling (Hauser et al., 2014; Walsh & Anderson, 2012). The component was decreased in trials which followed losses vs. wins in free-choice, but not in instructed tasks. Interestingly, current outcome did not influence this component. This shows that FRN may be related to PEAB but not CEAB. However, most studies report higher amplitudes of FRN after negative feedback (Gehring, Liu, Orr, & Carp, 2012), especially when choices are made freely (Yeung et al., 2005), and our results show the opposite pattern, with more negative amplitudes following wins in a free-choice task. Some studies did show more pronounced FRN effects after unexpected compared to expected

outcomes (Pfabigan, Alexopoulos, Bauer, & Sailer, 2011), so one potential interpretation would be that the predictability of outcomes outweighed the effect of their valence. This also supports the notion of the behavioural effect being related to expectation violation. Recent accounts of FRN suggest that it may not only code reward prediction errors, but that it is also sensitive to “unsigned” prediction errors or salience indexed by the anterior cingulate cortex (ACC) (Sambrook & Goslin, 2015), used as a reinforcement learning signal conveyed to the posterior medial frontal cortex to supervise action selection (Holroyd & Coles, 2002). This would concord with our findings and may suggest that PEAB is its behavioural, adaptive expression observable in volitional action. Moreover, larger outcome-related FRN amplitudes have also been associated with lower SoA in action-priming paradigms (Sidarus et al., 2017a). Similarly, in our probabilistic reversal learning task, FRN was reduced (less negative) following losses than following wins, which could be linked to increased sense of agency following losses, as in the PEAB.

Our results also demonstrate a similarity between PEAB and CEAB effects regarding the P3 component. This component was increased in free-choice trials if there was an error vs. win on the previous trial (PEAB) but was also increased when a current free-choice trial produced a loss vs. a win (CEAB). This seems to show that P3 is not specific to PEAB or CEAB but captures the outcome valence of either past or current outcomes, related to a more general mechanism of prediction error coding, which must be at play in both PEAB and CEAB. For example, both PEAB and CEAB involve a mismatch between predicted feedback (in all our tasks, participants can typically expect a win following each action) and actual feedback signalling a loss. P3 could reflect an unspecific mechanism signalling prediction error relating either to current or past losses. As both error-related effects were associated with a P3 effect in free choice tasks but not in instructed tasks, this component may be considered as a signature of intentionally caused, but surprisingly erroneous outcomes – a form of negative prediction error.

The rather frontal topography and rapid increase of the observed P3 effect resembles those of P3a component, or novelty P3 (Polich, 2007). It has been suggested that difficult tasks and unexpected outcomes (Hagen, Gatherwright, Lopez, & Polich, 2006; Verleger, Jaskowski, & Wauschkuhn, 1994) are more likely to produce this type of P3, and so the less likely error outcomes of the demanding PRL task employed in the current study could elicit similar responses. Motivational significance or task relevance are further crucial factors in influencing the P3 amplitude (Nieuwenhuis, Aston-Jones, & Cohen, 2005), and in the current task negative outcomes were more important for optimizing behaviour and increasing reward.

Future directions and conclusions

The effects investigated in this study can have implications for the understanding of some psychopathological conditions. For example, it may be the case that PEAB plays a role in disturbances of probabilistic learning, perhaps mediated by inadequate predictive processing, as seen e.g. in schizophrenia (Voss et al., 2010; Waltz & Gold, 2007). Potentially, PEAB – or its absence – could be linked to passivity and learned helplessness, as observed in depression. We suggest that increased sense of agency after loss may be an important element in persistence, and in keeping on trying despite frustration. These are important cognitive elements of general psychological resilience. The present study adds three key pieces of information about such agency persistence. First, it is linked to the specific task challenge, and is thus distinct from general motivation. Second, it is triggered by self-generated losses, rather than losses per se: increased agency reflects the epistemic value of *one's own* mistakes. Third, the effect does not depend on the immediate expression of learning through freedom of choice for the subsequent action. Rather an instructed task-related action is sufficient.

Our study includes a number of limitations. The first is related to our implicit measure of agency. Future studies should further investigate whether the PEAB would be observed using different implicit measures of SoA, but also to verify whether similar processes can be observed on a more conscious, declarative level of experience (i.e. judgement of agency; Synofzik, Vosgerau, & Newen, 2008). For the purposes of the present study, we have taken intentional binding as a suitable proxy for sense of agency, and have then used it in a reverse-inferential manner. However, the value of this approach, like other reverse inference approaches, could be questioned. Another important research question is the relation between PEAB and adaptive behaviour. Future studies might investigate the relationship between PEAB strength and PRL task difficulty, and trial-by-trial effects of losing in behavioural and ERP measures. Such studies should use designs optimized for sequential analyses, with longer runs of a single task.

In conclusion, we showed that sense of agency, as measured by IB, is increased following trials in which agents are free to choose what action to take, when this action leads to unwanted outcomes (loss), and when the agent is able to learn from these outcomes by actively or passively engaging in the same task. In other words, PEAB occurs when agents can learn (as already shown by Di Costa et al., 2017), when information about outcomes has been generated by their previous free action choices (as found in Experiment 2), and when they persist in the task that supports learning (as opposed to switching between tasks, as in Experiment 1). This

effect can be considered as a mechanism which preserves the feeling of agency even when success is elusive. It presumably helps agents to avoid learning helplessness when actions repeatedly result in unwanted outcomes (Maier & Seligman, 2016; Soral, Kofta, & Bukowski, 2020), and thus fosters further attempts to try and effectively learn about a specific task at hand, and achieve mastery. We suggest that sense of agency, at least as measured by intentional binding, is boosted when there is motivation to improve performance by learning from errors. The PEAB provides a novel experimental approach to the three-way relation between volitional choice, sense of agency, and learning.

ACKNOWLEDGMENTS

The authors would like to thank Nura Sidarus for her helpful comments on the previous version of the manuscript, Elisabeth Parés-Pujolràs for her contribution to the pre-processing of EEG data, and Roland Pfister for his helpful review of the first submission. BM was financed by the Polish Ministry of Science and Higher Education under the “Diamond Grant” programme (0097/DIA/2014/43), and also supported by a Study Visit Grant from the Experimental Psychology Society. EK and PH were supported by ERC Advanced Grant HUMVOL to PH (grant agreement 323943).

APPENDIX

Table A.1. Means (and Standard Errors) of reaction times (ms), accuracy (percent of correct button presses), and time reports (ms) in Experiment 1.

Previous outcome	Current outcome	Task type	Trial type	RT	Accuracy	Report
Loss	Loss	Monetary	Switch	863 (119.2)	33.9 (1.8)	334 (29.8)
			Repetition	519.6 (64.6)	32.5 (1.9)	355.5 (33.2)
		Affective	Switch	826.4 (99.2)	33.3 (2.8)	356.3 (32.4)
			Repetition	545.2 (83.5)	27.4 (1.8)	357.7 (30.4)
	Win	Monetary	Switch	782.5 (74.9)	88.7 (1.3)	383.2 (30.3)
			Repetition	515.9 (55.2)	86.7 (1.3)	362.3 (27.9)
		Affective	Switch	810 (75.9)	89 (1.1)	382.5 (28.4)
			Repetition	536.1 (60)	86.7 (1.4)	367.6 (28.5)
Win	Loss	Monetary	Switch	793.2 (85.9)	35.2 (2.1)	352.3 (27.9)
			Repetition	555.6 (55.3)	47.1 (1.5)	372.7 (32.3)
		Affective	Switch	820.1 (86.3)	37.2 (1.6)	363.8 (26.2)
			Repetition	561.7 (62.4)	45.4 (2)	395.1 (32.1)
	Win	Monetary	Switch	745 (57.3)	89.9 (0.9)	373 (22.5)
			Repetition	559.4 (45.5)	94 (0.6)	393.9 (24.2)
		Affective	Switch	743 (50.5)	89.6 (0.9)	381.7 (21.9)
			Repetition	556 (40.1)	93.4 (0.6)	390.6 (22.6)

Table A.2. Means (and Standard Errors) of reaction times (ms), accuracy (percent of correct button presses), time reports (ms), and amplitudes of FRN and P3 (μV) in Experiment 2.

Previous outcome	Current outcome	Task type	Trial type	RT	Accuracy	Report	P3	FRN
Loss	Loss	Free	Switch	805.9 (100.8)	29.9 (2.1)	528.7 (52)	1.1482 (0.1384)	0.7867 (0.1715)
			Repetition	646.8 (94)	32.8 (2.3)	529.9 (53.9)	1.1199 (0.1415)	0.4848 (0.1799)
		Instructed	Switch	565.1 (48.5)	32.3 (2.1)	492 (56.3)	0.2841 (0.1137)	0.2113 (0.1637)
			Repetition	532.7 (41.8)	38.4 (2.2)	556.5 (56.9)	0.0606 (0.1147)	0.1137 (0.1612)
	Win	Free	Switch	835.1 (100.2)	88.6 (1)	580 (49)	0.6646 (0.1294)	0.4583 (0.1314)
			Repetition	638 (78.5)	88.3 (1.4)	547.3 (48)	0.5922 (0.1337)	0.4421 (0.1396)
		Instructed	Switch	621.3 (70.1)	89.3 (1)	514.3 (47.6)	-0.0945 (0.1197)	-0.0435 (0.1356)
			Repetition	530.6 (34.1)	90.9 (0.9)	547.5 (46.3)	-0.0039 (0.0977)	-0.0414 (0.1264)
Win	Loss	Free	Switch	820.1 (93.8)	40.2 (2.3)	527.5 (51.8)	0.8132 (0.1359)	0.5688 (0.1656)
			Repetition	644.1 (70.1)	50.8 (2.8)	534.2 (52)	0.6952 (0.1123)	0.4069 (0.1629)
		Instructed	Switch	608.7 (46.5)	33.2 (1.6)	523.9 (46.2)	0.026 (0.1014)	0.1697 (0.1516)
			Repetition	547 (40.9)	29.3 (1.7)	547.7 (48.9)	0.2852 (0.1045)	0.392 (0.1371)
	Win	Free	Switch	749.2 (64)	92.1 (0.9)	568.4 (36.8)	0.5308 (0.109)	0.4412 (0.1201)
			Repetition	633.9 (49.8)	95.5 (0.4)	574.8 (35.8)	0.1055 (0.1101)	0.1613 (0.1235)
		Instructed	Switch	605.7 (40.9)	90 (0.8)	539 (36.9)	-0.0182 (0.0947)	0.1157 (0.1275)
			Repetition	543.6 (30.4)	89.5 (0.8)	544.5 (38.8)	0.0674 (0.0994)	0.051 (0.1239)

REFERENCES

- Allport, A., Styles, E., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. U. & M. Moscovitch (Ed.), *Attention and performance XV* (pp. 421–452). Hillsdale, NJ: Erlbaum.
- Barlas, Z., Hockley, W. E., & Obhi, S. S. (2017). Effects of free choice and outcome valence on the sense of agency: evidence from measures of intentional binding and feelings of control. *Experimental Brain Research*, 236(1), 129–139.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, 67(1).
- Beck, B., Di Costa, S., & Haggard, P. (2017). Having control over the external world increases the implicit sense of agency. *Cognition*, 162, 54–60.
- Bolker, B. (2019). GLMM FAQ (retrieved from: <https://bbolker.github.io/mixedmodels-misc/glmmFAQ.html>).
- Borhani, K., Beck, B., & Haggard, P. (2017). Choosing, Doing, and Controlling: Implicit Sense of Agency Over Somatosensory Events. *Psychological Science*, 28(7), 882–893.
- Botvinick, M. (2007). Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive, Affective, & Behavioral Neuroscience*, 7(4), 356–366.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436.
- Caspar, E. A., Christensen, J. F., Cleeremans, A., & Haggard, P. (2016). Coercion Changes the Sense of Agency in the Human Brain. *Current Biology*, 26, 1–8.
- Christensen, J. F., Yoshie, M., Di Costa, S., & Haggard, P. (2016). Emotional valence, sense of agency and responsibility: A study using intentional binding. *Consciousness and Cognition*, 43, 1–10.
- Danielmeier, C., & Ullsperger, M. (2011). Post-error adjustments. *Frontiers in Psychology*, 2(SEP), 1–10.
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21.
- Dewey, J. A., & Knoblich, G. (2014). Do Implicit and Explicit Measures of the Sense of Agency Measure the Same Thing? *PLoS ONE*, 9(10), e110118.
- Di Costa, S., Théro, H., Chambon, V., & Haggard, P. (2017). Try and try again: Post-error boost of an implicit measure of agency. *The Quarterly Journal of Experimental Psychology*, 71(7), 1584–1595.
- Dutilh, G., Van Ravenzwaaij, D., Nieuwenhuis, S., Van der Maas, H. L. J., Forstmann, B. U., & Wagenmakers, E. J. (2012). How to measure post-error slowing: A confound and a simple solution. *Journal of Mathematical Psychology*, 56(3), 208–216.
- Dutilh, G., Vandekerckhove, J., Forstmann, B. U., Keuleers, E., Brysbaert, M., & Wagenmakers, E. J. (2012). Testing theories of post-error slowing. *Attention, Perception, and Psychophysics*, 74(2), 454–465.
- Fried, I., Haggard, P., He, B. J., & Schurger, A. (2017). Volition and Action in the Human Brain: Processes, Pathologies, and Reasons. *Journal of Neuroscience*, 37(45), 10842–10847.
- Gade, M., & Koch, I. (2007). The influence of overlapping response sets. *Memory & Cognition*, 35(4), 603–609.
- Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends in Cognitive Sciences*, 4(1), 14–21.

- Gallagher, S. (2012). Multiple aspects in the sense of agency. *New Ideas in Psychology*, 30(1), 15–31.
- Gehring, W. J., Liu, Y., Orr, J. M., & Carp, J. (2012). The Error-Related Negativity (ERN/Ne). In Emily S. Kappenman & Steven J. Luck (Eds.), *The Oxford Handbook of Event-Related Potential Components*.
- Glazer, J. E., Kelley, N. J., Pornpattananangkul, N., Mittal, V. A., & Nusslock, R. (2018). Beyond the FRN: Broadening the time-course of EEG and ERP components implicated in reward processing. *International Journal of Psychophysiology*, 132(Part B), 184–202.
- Hagen, G. F., Gatherwright, J. R., Lopez, B. A., & Polich, J. (2006). P3a from visual stimuli: Task difficulty effects. *International Journal of Psychophysiology*, 59, 8–14.
- Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, 18(4), 1–12.
- Haggard, P. (2019). The Neurocognitive Bases of Human Volition. *Annual Review of Psychology*, 70, 1–20.
- Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, 5(4), 382–385.
- Hauser, T. U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., & Brem, S. (2014). The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization. *NeuroImage*, 84, 159–168.
- Holroyd, C. B., & Coles, M. G. H. (2002). The Neural Basis of Human Error Processing: Reinforcement Learning, Dopamine, and the Error-Related Negativity. *Psychological Review*, 109(4), 679–709.
- Huelser, B. J., & Metcalfe, J. (2012). Making related errors facilitates learning, but learners do not know it. *Memory and Cognition*, 40(4), 514–527.
- 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), 133–51.
- Koch, I. (2005). Sequential task predictability in task switching. *Psychonomic Bulletin & Review*, 12(1), 107–112.
- Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful Retrieval Attempts Enhance Subsequent Learning. *Journal of Experimental Psychology: Learning Memory and Cognition*, 35(4), 989–998.
- Kuhn, S., Nenchev, I., Haggard, P., Brass, M., Gallinat, J., & Voss, M. (2011). Whodunnit? electrophysiological correlates of agency judgements. *PLoS ONE*, 6(12), 1–6.
- Kulakova, E., Khalighinejad, N., & Haggard, P. (2017). I could have done otherwise: Availability of counterfactual comparisons informs the sense of agency. *Consciousness and Cognition*, 49, 237–244.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13).
- Laming, D. (1968). *Information theory of choice-reaction times*. London, UK: Academic Press.
- Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces (KDEF). *CD ROM from Department of Clinical Neuroscience, Psychology Section, Karolinska Institutet*.
- Maier, S. F., & Seligman, M. E. P. (2016). Learned helplessness at fifty: Insights from neuroscience. *Psychological Review*, 123(4), 1–19.

- Majchrowicz, B., & Wierzchoń, M. (2018). Unexpected action outcomes produce enhanced temporal binding but diminished judgement of agency. *Consciousness and Cognition*, *65*, 310–324.
- Metcalf, J. (2017). Learning from Errors. *Annual Review of Psychology*, *68*, 465–489.
- Metcalf, J., & Greene, M. J. (2007). Metacognition of agency. *Journal of Experimental Psychology: General*, *136*(2), 184–199.
- Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological Bulletin*, *130*(5), 711–747.
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*(3), 134–140.
- Moore, J. W., & Haggard, P. (2008). Awareness of action: Inference and prediction. *Consciousness and Cognition*, *17*(1), 136–44.
- Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: a review. *Consciousness and Cognition*, *21*(1), 546–61.
- Moreton, J., Callan, M. J., & Hughes, G. (2017). How much does emotional valence of action outcomes affect temporal binding? *Consciousness and Cognition*, *49*, 25–34.
- Moretto, G., Walsh, E., & Haggard, P. (2011). Experience of agency and sense of responsibility. *Consciousness and Cognition*, *20*(4), 1847–1854.
- Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin*, *131*(4), 510–532.
- Niv, Y., Edlund, J. A., Dayan, P., & O’Doherty, J. P. (2012). Neural Prediction Errors Reveal a Risk-Sensitive Reinforcement-Learning Process in the Human Brain. *Journal of Neuroscience*, *32*(2), 551–562.
- Pariyadath, V., & Eagleman, D. (2007). The effect of predictability on subjective duration. *PLoS ONE*, *2*(11), 2–7.
- Pfabigan, D. M., Alexopoulos, J., Bauer, H., & Sailer, U. T. A. (2011). Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. *Psychophysiology*, *48*, 656–664.
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148.
- R Core Team. (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rogers, R., & Monsell, S. (1995). Costs of a Predictable Switch Between Simple Cognitive Tasks, (September 2015).
- Rubin, O., & Meiran, N. (2005). On the Origins of the Task Mixing Cost in the Cuing Task-Switching Paradigm, *31*(6), 1477–1491.
- Sambrook, T. D., & Goslin, J. (2015). A Neural Reward Prediction Error Revealed by a Meta-Analysis of ERPs Using Great Grand Averages. *Psychological Bulletin*, *141*(January), 213–235.
- San Martín, R. (2012). Event-related potential studies of outcome processing and feedback-guided learning. *Frontiers in Human Neuroscience*, *6*.
- Sidarus, N., Vuorre, M., & Haggard, P. (2017a). How action selection influences the sense of agency: An ERP study. *NeuroImage*, *150*(February), 1–13.
- Sidarus, N., Vuorre, M., & Haggard, P. (2017b). Integrating prospective and retrospective cues to the sense of agency: a multi-study investigation. *Neuroscience of Consciousness*, *3*(1), 1–10.

- Soral, W., Kofta, M., & Bukowski, M. (2020). Helplessness experience and intentional (un-)binding: Control deprivation disrupts the implicit sense of agency. *Journal of Experimental Psychology: General*.
- Steinhauser, M., Ernst, B., & Ibal, K. W. (2017). Isolating component processes of posterror slowing with the psychological refractory period paradigm. *Journal of Experimental Psychology: Learning Memory and Cognition*, 43(4), 653–659.
- Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: a multifactorial two-step account of agency. *Consciousness and Cognition*, 17(1), 219–39.
- Synofzik, M., Vosgerau, G., & Voss, M. (2013). The experience of agency: an interplay between prediction and postdiction. *Frontiers in Psychology*, 4(March), 1–8.
- Takahata, K., Takahashi, H., Maeda, T., Umeda, S., Suhara, T., Mimura, M., & Kato, M. (2012). It's Not My Fault: Postdictive Modulation of Intentional Binding by Monetary Gains and Losses. *PLoS ONE*, 7(12), 1–8.
- Verleger, R., Jaskowski, P., & Wauschkuhn, B. (1994). Suspense and surprise: On the relationship between expectancies and P3. *Psychophysiology*, 31(4), 359–369.
- Voss, M., Moore, J. W., Hauser, M., Gallinat, J., Heinz, A., & Haggard, P. (2010). Altered awareness of action in schizophrenia: a specific deficit in predicting action consequences. *Brain: A Journal of Neurology*, 133(10), 3104–12.
- Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience and Biobehavioral Reviews*, 36(8), 1870–1884.
- Waltz, J. A., & Gold, J. M. (2007). Probabilistic reversal learning impairments in schizophrenia: Further evidence of orbitofrontal dysfunction. *Schizophrenia Research*, 93(1–3), 296–303.
- Wanat, M. J., Willuhn, I., Clark, J. J., & Phillips, P. E. M. (2009). Phasic Dopamine Release in Appetitive Behaviors and Drug Addiction. *Current Drug Abuse Reviews*, 2, 195–213.
- Wickham, H. (2016). *ggplot2: elegant graphics for data analysis*. Springer.
- Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP Correlates of Feedback and Reward Processing in the Presence and Absence of Response Choice. *Cerebral Cortex*, 15(5), 535–544.
- Yeung, N., & Monsell, S. (2003). Switching Between Tasks of Unequal Familiarity: The Role of Stimulus-Attribute and Response-Set Selection. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2), 455–469.
- Yoshie, M., & Haggard, P. (2013). Negative Emotional Outcomes Attenuate Sense of Agency over Voluntary Actions. *Current Biology*, 23(20), 2028–2032.
- Yoshie, M., & Haggard, P. (2017). Effects of emotional valence on sense of agency require a predictive model. *Scientific Reports*, 7(1), 8733.