UNIVERSITY of CALIFORNIA PRESS | Collabra: Psychology

Eben, C., Vermeylen, L., Chen, Z., Notebaert, W., Ivanchei, I., & Verbruggen, F. (2023). When Response Selection Becomes Gambling: Post-error Slowing and Speeding in Self-paced Colour Discrimination Tasks. *Collabra: Psychology*, *9*(1). https://doi.org/10.1525/collabra.73052

Cognitive Psychology

When Response Selection Becomes Gambling: Post-error Slowing and Speeding in Self-paced Colour Discrimination Tasks

Charlotte Eben[1][a], Luc Vermeylen[1], Zhang Chen[1], Wim Notebaert[1], Ivan Ivanchei[1], Frederick Verbruggen[1]

[1] Department of Experimental Psychology, Ghent University, Ghent, Belgium

Keywords: post-error slowing, post-error speeding, action control, self-pace, behavioral adjustments

https://doi.org/10.1525/collabra.73052

Collabra: Psychology

Vol. 9, Issue 1, 2023

People tend to slow down after committing an error in many tasks. However, some studies failed to observe such post-error slowing. Furthermore, recent work found speeding after another type of sub-optimal outcomes: people often speed up after losses in gambling situations. What features determine whether people slow down or speed up after sub-optimal outcomes (error vs. loss)? To answer this question, we focused on the role of task characteristics and control over the outcome, by making a task where we previously observed post-error slowing more like tasks where we previously observed post-loss speeding. First, we made a color-discrimination task completely self-paced (Experiment 1A) and added reward/punishment (Experiment 1B). In both experiments, post-error slowing was observed, without modulation by reward/punishment. We then manipulated task difficulty to investigate the influence of control over the outcome. Consistent with our predictions, control over the outcome modulated post-error adjustments, as participants slowed down after controllable errors, but sped up after uncontrollable errors (Experiment 3). Importantly, this effect was global as post-error speeding was observed when controllable and 'uncontrollable' errors were intermixed (Experiment 2), suggesting an influence of overall task context. Thus, responses to sub-optimal outcomes might depend on the control over the outcome.

Introduction

In psychology, the finding that participants slow down after committing an error in cognitive tasks has a long standing history (Rabbitt & Rodgers, 1977). Additionally, post-error slowing has received much attention in more clinically-oriented and applied fields as the phenomenon might provide a unique window into how individuals or groups respond differently to errors or sub-optimal outcomes (e.g., losses, outcomes that are worse than the best possible outcome, or outcomes that come with serious costs). Based on such work, it has been argued that failures to adjust behavior after errors or sub-optimal outcomes might contribute to the development of a variety of clinical and behavioral problems, such as behavioral and substance addictions (Garavan & Stout, 2005; Sullivan et al., 2019). However, a recent set of studies, using a variety of tasks, suggest that post-error slowing might not be a ubiquitous phenomenon after all (Damaso et al., 2020; Eben et al., 2020; Verbruggen et al., 2017; Williams et al., 2016). The aim of this study was to investigate why seemingly similar sub-optimal outcomes influence subsequent behavior differently.

To slow or not to slow: That's the question

One of the most common observations in cognitive psychology is that participants slow down after committing an error (post-error slowing). Different accounts have been proposed to explain this phenomenon. Cognitive control accounts (e.g. Botvinick et al., 2001; Dutilh, Vandekerckhove, et al., 2012) assume that a monitoring system evaluates actions and action outcomes. If the action outcome is sub-optimal, which is the case for errors, participants adjust the "task set" or "task parameters" (Logan & Gordon, 2001), such as the amount of information that is required to make a decision, to avoid subsequent errors. Such adjustments might persist for multiple trials (e.g., Forster & Cho, 2014). Cognitive control accounts attribute the slowing after errors to such changes in response threshold (e.g., Dutilh, Vandekerckhove, et al., 2012). Thus, post-error slowing is assumed to be an adaptive process which leads to increased accuracy after errors.

a Correspondence concerning this article should be addressed to Charlotte Eben, Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, B-9000 Gent, Belgium. Email: ebencharlotte01@gmail.com

However, this reduced error rate is not always observed (e.g. Hajcak & Simons, 2008; Rabbitt & Rodgers, 1977). Therefore, non-strategic accounts of post-error slowing have been proposed as well. For example, previous research has shown that unexpected/infrequent events produce an orienting response, leading to slower responses on the subsequent trial (Barcelo et al., 2006; Rabbitt & Phillips, 1967). Notebaert et al. (2009) proposed that this might also explain why responses are slower after errors, as errors are typically infrequent events. Consistent with this idea, they found that participants slowed down and also made more errors after infrequent errors; importantly however, participants no longer slowed down if the error rate was high, suggesting that frequency indeed matters.

Since there has been support for both strategic (control) and non-strategic (orienting) accounts, some have tried to integrate the two (Wessel, 2018). For example, Wessel (2018) proposed that all unexpected action outcomes will first produce an orienting response (slowing responses). In case of errors, this orienting response can be followed by more controlled adjustments of behavior, and as such, increased task accuracy. But whether or not such adjustments can be made would depend highly on the pace of the task: If the task is slow paced and the next stimulus does not appear before task settings have been adjusted, post-error slowing would be associated with increased accuracy (i.e., it would be adaptive). However, if the task is fast paced and there is not sufficient time to adjust task settings, post-error slowing would be associated with reduced task accuracy (i.e., it would be maladaptive).

Regardless of the theoretical differences between these accounts, they all seem to assume that people slow down after (infrequent) errors. Nevertheless, a few studies did not observe slowing after errors, and sometimes even observed speeding (Damaso et al., 2020; Fievez et al., 2021). Such inconsistencies have also been (at least partly) attributed to the pace of the task. For example, Yeung and Summerfield (2012) assumed that errors and post-error slowing mainly occur in fast-paced decision making tasks in which participants are aware of their own errors but simply did not sample enough information to make the correct informed choice. In line with this, Williams et al. (2016) found that (unpaid) participants in a more slow-paced task of several seconds per trial did not slow down but rather sped up after errors, which they explained as a motivational effect: participants seemed to be discouraged if they could not control the errors and therefore became more 'reckless' whereas successes might have encouraged the participants to try harder.

The speeding after errors is also consistent with a recent series of findings in the gambling literature. Specifically, several studies indicate that participants speed up after losses compared to wins and non-gambling trials (Chen et al., 2022; Corr & Thompson, 2014; Dixon et al., 2013; Eben et al., 2020; Verbruggen et al., 2017). This has been explained within the framework of appraisal accounts, which assume that behavior results from a discrepancy between a current (i.e., loss) and a desired (i.e., win) state. The bigger the discrepancy is, the bigger is the urge to act, and the more subsequent behavior becomes invigorated (e.g. faster responses, Frijda, 2010; Moors et al., 2013). Frijda (2010) called such actions 'impulsive'. Others have linked this increased vigor after a failure to obtain a reward to (primary) 'frustration' (see e.g., the seminal work by Amsel, 1958). But regardless the use of different theoretical concepts (e.g., 'impulsivity' vs. 'frustration'), the general idea is that behaviour might also become invigorated after unsuccessful events or sub-optimal outcomes (i.e., increased vigor instead of restraint).

To have control or not to have control: That might be the answer?

The brief overview above suggests that people do not always slow down after errors or other types of sub-optimal outcomes. This raises the question what makes people slow down or speed up if something goes wrong. As already mentioned above, work by Fievez et al. (2021), Wessel (2018), and Yeung and Summerfield (2012) suggest that the task pace and time course might influence behavioral responses after errors or other sub-optimal outcomes. But this cannot be the full story as, for instance, some studies failed to observe post-error slowing and an increase in task accuracy even when there should have been sufficient time to make control adjustments to the task set (e.g. Jentzsch & Dudschig, 2009).

In particular, it seems that control over the outcome might determine to what extent participants *slow down*, or by contrast, *speed up* when something goes wrong. For example, Damaso et al. (2020) investigated changes in response speed after different types of errors. They distinguished between 'response speed errors' and 'evidence quality errors'. Similar to Yeung and Summerfield (2012), they defined 'response speed errors' as errors in which participants simply responded too quickly (i.e., they did not sample enough information to make a correct choice; thus, their decision threshold was too low). On the other hand, 'evidence quality errors' were defined as errors committed with very slow responses due to the very poor quality of the stimulus (i.e., there was insufficient evidence for either response; for a similar distinction see also Beatty et al., 2018). Damaso et al. (2020) found that participants slowed down after 'response speed errors' but not after 'evidence quality errors' and they attributed these findings to the controllablity of the errors. That is, participants could have avoided 'response speed errors' if they had slowed down and sampled more information (i.e., they had control over the outcome); by contrast, 'evidence quality errors' could not have been avoided by slowing down and sampling more information (i.e., participants did not have control over these outcomes). Similarly, Steinhauser and Kiesel (2011) found that participants slowed down if they committed the error (internally caused error) but did not slow down when they thought the apparatus was malfunctioning (externally caused error). Comparably, de Bruijn et al. (2004) found no performance adjustments (i.e., post-error slowing) after false feedback when the task was easy enough for participants to detect that their response was correct.

Controllability over the outcome also seems to play a role in games of chance or gambling-like tasks. For example, Dyson et al. (2018) found in a computerized 'rock, paper, scissors' game that participants sped up after losses when they were playing against an unexploitable (i.e., random) computer opponent, but they slowed down when they were playing against an exploitable opponent (e.g., the computer responded with the option that would have beaten the player's previous response). In line with the findings of Damaso et al., this study also suggests that participants who feel in control over the outcome slow down after sub-optimal outcomes, whereas those who do not feel in control over the outcome do not slow down and might even speed up. This would explain the consistent speeding in previous gambling studies (Corr & Thompson, 2014; Dixon et al., 2013; Eben et al., 2020; Verbruggen et al., 2017), as participants did not have control over the outcome in these tasks.

A recent study further tested this idea in a gambling task by manipulating the *subjective feeling* of control (without changing the chance-determined nature of gambling tasks). Specifically, outcome sequences were manipulated at the beginning of the experiment to induce an illusion of control (building on the seminal work by Langer & Roth, 1975). The self-report data indicated that the manipulation was successful. Participants in the experimental group indeed felt more in control compared with those in the control groups. However, the post-loss speeding was not influenced by this illusion-of-control manipulation. Arguably, the manipulation might have been too weak and obscured by a (stronger) effect of overall task context. Consistent with the latter idea, the self-report data indicated that even the participants in the experimental group still seemed to realise that they were playing a game of chance (Eben et al., 2022).

The Present Study

Thus, to further investigate the role of overall task-context, in the present study, we investigated whether the *objective* controllability over the outcome indeed determines whether participants speed up or slow down after sub-optimal outcomes. In contrast to Dyson et al. (2018), we did not test this in a gambling-related context, but in a task that typically allows participants to control the outcome. But before we could test this hypothesis, we first needed to rule out other task features that could influence speeding vs. slowing. First of all, gambling tasks are often self-paced, which means participants press a key to start the next trial. However as mentioned above, the post-error slowing literature suggests that the pace of a task also influences the extent to which people slow down (and even speed up) after errors, regardless of controllability over the outcome. Second, in gambling tasks participants can usually win or lose points (or money). Thus, negative outcomes in gambling tasks might be more motivationally salient compared to errors in most choice reaction time experiments, in which no reward or punishment is delivered after every outcome. The effect of time (or pace) and reward/punishment might even interact in unexpected ways. For example, some studies found that in fast-paced tasks participants slowed down more after errors when they were rewarded for correct responses and punished for incorrect responses (Riesel et al., 2012; Stürmer, 2011). By contrast, in their slow-paced task, Williams et al. (2016) observed the most pronounced post-error effects in the unrewarded condition. Importantly though, Williams et al. (2016) observed speeding after errors rather than slowing. Given this pattern of results, we first attempted to replicate post-error slowing in a self-paced task (Experiment 1A), in which reward and punishment were delivered (Experiment 1B).

Encouraged by the findings of these first two experiments (in which we observed post-error slowing), we then focused on the main task feature of interest, namely controllability over the outcome. In Experiment 2, we manipulated controllability over the outcome within participants. We compared an 'easy' condition in which distinguishable stimuli were presented (i.e., participants had control over the outcome) with a 'hard' condition in which we predetermined the outcome as in gambling tasks and participants would respond 'correctly' on only half of the trials (i.e., they did not have control over the outcome). To our surprise, we observed post-error speeding in both conditions, suggesting that the reduced controllability over the outcome in the 'hard' condition might have a global influence on post-error performance. Therefore, we ran a third experiment in which we manipulated control over the outcome between groups.

Experiment 1A

Method

Transparency and Openness for all Experiments

All raw and processed data, code, and materials of all experiments can be found on OSF (https://osf.io/azjey/). The preregistrations for Experiments 1B-3 are also on OSF (Experiment 1B, Experiment 2, Experiment 3). We report (here and in our preregistrations) how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. Moreover, the series of studies conducted received approval of the local ethics committee.

Sample Size

Experiment 1A was exploratory in nature. We initially decided to test 50 participants to see whether we could observe post-error slowing in a self-paced choice task. Not all of the post-error slowing effects (see online supplementary material) showed strong evidence for either hypothesis (BF_{10} > 10 or BF_{10} < 1/10), therefore we added another 50 participants. Importantly, the Bayesian sequential testing procedure allows us to interpret Bayes Factors despite optional stopping (Rouder, 2014; Schönbrodt et al., 2017; Schönbrodt & Wagenmakers, 2018; but see also de Heide & Grünwald, 2021)

Participants

In total 100 participants (recruited via Prolific in two samples of 50 participants) completed the entire online experiment (42 females; age M = 27.37 years, SD = 8.19 years,

range = 18-55). Only participants who spoke English and never participated in an experiment by the first author were allowed to participate. Participants were paid at a rate of £6/hour; as this study took approximately 30 minutes to complete, they received £3. Participants agreed to the consent form before starting the experiment.

An additional 15 participants signed up for the experiment but never started or completed it. From our experience, this seems to be a common practice on Prolific in order to 'reserve' a spot. The first 50 participants were tested on the 10th June 2020 and the second 50 participants were tested on the 4th August 2020.

Apparatus and Stimuli

The experiment was programmed in jsPsych (version 6.0.5) and only ran on desktop computers and laptops in Chrome (de Leeuw, 2015). Keyboards were used to register responses. Participants had to perform a color discrimination task. On each trial, they saw a circle (size: 94.5 px) in one out of six possible shades of blue (RGB codes: 8,81,156; 49,130,189; 107,174,214; 158,202,225; 198,219,239; 239,243,255; going from dark to light blue). Participants had to identify the color of the circle by pressing the corresponding key (s, d, f, j, k or l; keys were arranged from dark – s - to the lightest blue option - j). Note that we did not counterbalance the stimulus-response mapping. All colors were presented equally often in a random order. The circles always appeared in the centre of the screen against a white background. After each trial, the word 'correct' or 'incorrect' (depending on their response) was presented as feedback. In the practice block, all possible colors with the corresponding keys were also presented below as a reminder.

Procedure

Each trial started with the message "Press the space bar to start the next trial.". After participants had pressed the space bar, the target circle was shown until they responded. Immediately after the response, participants got feedback about whether their response was correct or incorrect for 1000 ms. Then the next trial started with the message "Press the space bar to start the next trial." again. Thus, the entire experiment was self-paced.

The experiment started with a practice block of 36 trials, followed by eight experimental blocks of 72 trials (576 experimental trials in total). During the breaks after each block which were also self-paced, participants got a reminder about the stimulus-response mapping and they were told that they could take a short break if they wished to do so. At the end of the experiment, all participants filled in the short version of the UPPS-P impulsive behavior questionnaire (Cyders et al., 2014). Note that these questionnaire data were not used in this study but were part of a bigger individual differences study.

Analyses

All data processing and analyses were completed with R (R Core Team, 2013, version 4.0.2) using the packages optimx (Nash, 2014, version 4.2), lmerTest (Kuznetsova et al., 2017, version 3.1-3), car (Fox & Weisberg, 2019, version 3.0-10), ggeffects (Lüdecke, 2020, version 1.0.1), sjPlot (Lüdecke, 2021, version 2.8.7), cowplot (Wilke, 2020, version 1.1.1), BayesFactor (Morey & Rouder, 2018, version 0.9.12-4.2), emmeans (Lenth, 2021, version 1.5.4), brms (Bürkner, 2018, version 2.15.0), broom (Robinson et al., 2021, version 0.7.4), rtdists (Singmann et al., 2020, version 0.11-2) and tidyverse (Wickham, 2021, version 1.3.0).

For the analyses, we excluded all practice trials, trials in which the start RT or choice RT was > 5000 ms (assuming participants took a break here), and all trials with a choice RT < 150 ms. The analyses focused on the effect of the outcome of the previous trial; thus, we also excluded the first trial of each block as well as trials in which the previous outcome (error vs. correct) was not known (as we were testing online, it was possible that single trials were not recorded). In line with our previous online studies, we decided to exclude and replace (a) all participants with more than 5% missing trials and (b) all participants who started the experiment again. However, no participant met these two exclusion criteria in this experiment.

We report Bayes factors for statistical inferences. For the Bayesian analyses, we report the Bayes Factor BF_{10}, which quantifies the evidence for the alternative hypothesis against the null hypothesis. A Bayes Factor > 1 is in favor of the alternative hypothesis, whereas a Bayes Factor < 1 is in favor of the null hypothesis. A Bayes Factor around 1 yields inconclusive evidence. The size of the Bayes Factor determines whether the evidence is anecdotal (1/3 – 1; 1-3), moderate (1/3 – 1/10; 3-10), strong (1/10 – 1/30; 10-30), very strong (1/30 – 1/100; 30-100) or extreme (< 1/100; > 100) (Wagenmakers et al., 2018). We used the default prior width as used in R for BayesFactors. For Bayesian t-tests this prior width is 0.707, corresponding to a medium effect.

Here we report three Bayesian t-tests comparing the start RT, the choice RT and the error rate after correct and incorrect trials. Note that this is the traditional way of measuring post-error slowing. Originally, we also planned on using the more robust measure of post-error slowing proposed by Dutilh, van Ravenzwaaij, et al. (2012) as it accounts for global fluctuations in response latencies and takes pre-error speeding into account. But for consistency purposes (i.e., in our later studies we were not able to calculate the robust measure of post-error slowing) we decided to only report the traditional measure in the main text. The results for the robust measure (where possible) can be found in the online supplementary materials. Importantly, all results were consistent with the traditional measure. Please also note that we originally planned to investigate post-error slowing as a function of difficulty (i.e. stimulus color) as well (see preregistration). These analyses can also be found on OSF.

Results and Discussion

The descriptive statistics can be found in Figure 1. For the inferential statistics, see Table 1. The Bayesian t-test on the start RT showed strong evidence for a difference between post-correct and post-error trials (BF_{10} > 100), show-

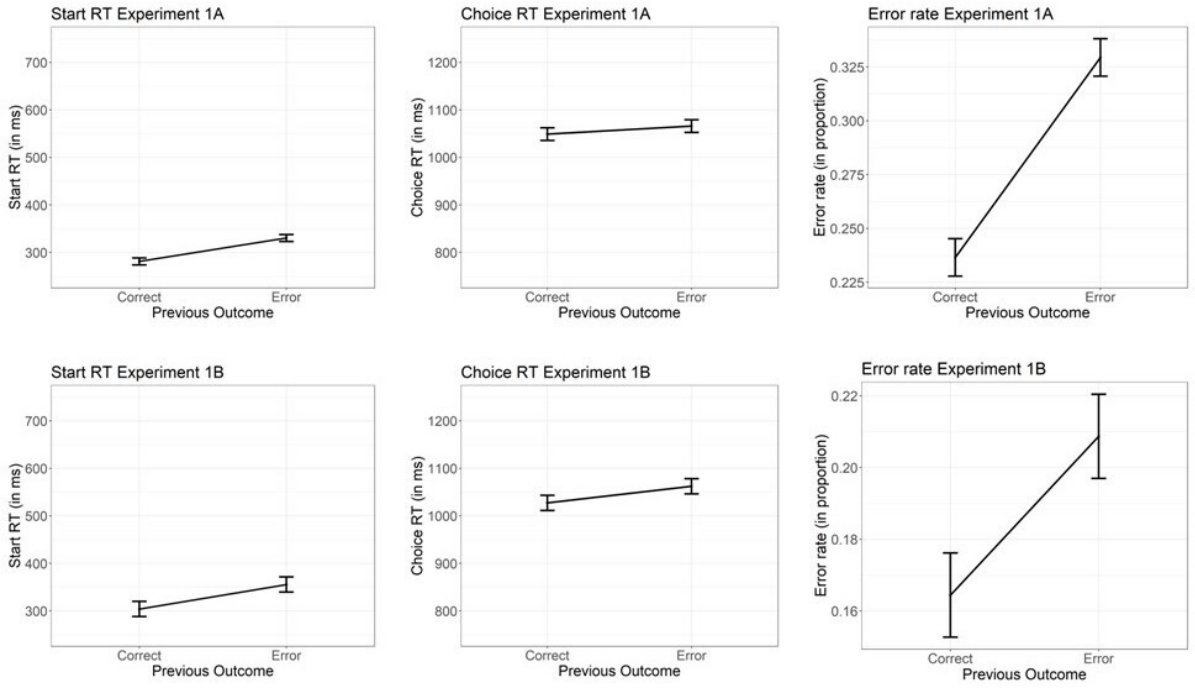

Figure 1. Start RT, choice RT and error rate as a function of previous outcome (traditional measure of post-error slowing) in Experiment 1A and Experiment 1B.

The error bars reflect the 95% confidence intervals.

Table 1. Pairwise comparisons of the post-error trials with post-correct trials on the start RT, choice RT and error rate in Experiment 1A and Experiment 1B.

	diff	Lower CI	Upper CI	df	*t*	*p*	BF_{10}	*gav*
Experiment 1A								
start RT	49.34	38.83	59.85	99	9.31	< .001	> 100	0.49
choice RT	18.29	0.84	37.41	99	1.90	.061	0.62	0.10
error rate	0.09	0.08	0.11	99	14.97	< .001	> 100	0.92
Experiment 1B								
start RT	50.35	27.89	72.81	59	4.49	< .001	> 100	0.44
choice RT	35.19	12.62	57.75	59	3.12	.003	10.70	0.16
error rate	0.04	0.03	0.06	59	5.34	< .001	> 100	0.51

Note: diff = difference; CI = confidence interval (95%); BF_{10} = Bayes Factor 10; g_{av} = Hedge's average *g*; To determine statistical significance, we used an alpha of .05.

ing a post-error slowing effect of 49 ms. Numerically, we also found a post-error slowing effect in the choice RT (18 ms) but the evidence for this effect was (inconclusive) in favor of the null hypothesis (BF_{10} = 0.619). Note that for both robust measures of post-error slowing, the Bayes factors showed strong evidence in favor of the alternative hypothesis (BF_{10} > 100 in the start RT and BF_{10} = 23 in the choice RT). Overall, this pattern of results indicates that making the task self-paced does not make participants speed up (rather than slow down) after an error. We also found that participants made more mistakes following an error compared to following a correct trial (difference = 9.3%; BF_{10} > 100), which supports the idea that slowing might not always be an adaptive process.

Experiment 1B

In Experiment 1B, we examined reward and punishment elements as an alternative explanation for slowing after errors and speeding after losses. Therefore, in addition to the self-pace element, we rewarded participants for every correct response and punished them for every incorrect response.

Method

Sample Size

As described in our preregistration, we set a minimum sample size of 50 participants and a maximum sample size of 100 participants. To determine the final sample size, we

used sequential Bayesian hypothesis testing by increasing the sample in steps of 10 participants until strong evidence was obtained for either hypothesis (BF_{10} > 10 or < 1/10) for either post-error effect (start RT or choice RT; Schönbrodt et al., 2017). We only wanted to know whether we could observe slowing or speeding after errors when simply adding reward/punishment. As the effects on both start RT and choice RT pointed in the same direction in Experiment 1A and as our cut-off value for the Bayes Factor was quite conservative, we decided that strong evidence for one of the dependent variables (i.e., start RT or choice RT) would be sufficient. Eventually, both effects reached the crucial BF (for details on the Bayesian t-tests see online supplementary material).

Participants

60 participants (recruited via Prolific) completed the entire online experiment (22 females, *M* = 25.88 years, *SD* = 7.58 years, *range* = 18-50). Only participants who spoke English and never participated in any experiment by the first author were allowed to participate. The participation requirements and the payment structure were the same as in Experiment 1A.

An additional 4 participants did the entire experiment and got paid for their participation but were excluded in accordance with our preregistered exclusion criteria: one participant had more than 5 % missing trials and three other participants started the experiment again, possibly because they missed the completion code for Prolific. An additional nine participants signed up for the experiment but never started or completed it. All data were acquired on 15th March 2021.

Apparatus, Stimuli and Procedure

We used the same procedure as in Experiment 1A. The only difference was the feedback presented to the participants: instead of the message "correct" or "incorrect", participants received one point for every correct response and lost one point for every incorrect response ("+1" or "-1" presented for 500 ms after every trial). At the end of the experiment, these points would be converted into real money (i.e., for every 100 points, participants would get 1£ extra). The maximum extra payout was set at 3£ and participants were informed about this payout structure in advance (average reward: *M* = 2.88 £, *range*= 1.34 - 3.00 £).

Analyses

For consistency purposes and to further improve the clarity of this manuscript, we decided to deviate from our preregistered analyses. Specifically, our main processing, and analysis pipeline as well as the inference criteria were the same as for Experiment 1A (and the other experiments reported below). For transparency purposes, we report all preregistered analyses in the online supplementary material. Note that all conclusions reported here are fully supported by the preregistered analyses as well.

Results and Discussion

The descriptive statistics can be found in Figure 1. The Bayesian t-test on the start RT showed that the difference between post-correct and post-error trials yielded extreme evidence in favor of the alternative hypothesis (BF_{10} > 100), showing a post-error slowing effect of 50 ms. We also found a post-error slowing effect in the choice RT (35 ms; BF_{10} = 10.7), as well as more errors on post-error trials compared to post-correct trials (difference = 4.4%; BF_{10} > 100). For more detailed inferential statistics, see Table 1.

Again, we were able to find post-error slowing as well as more errors following error trials than following correct trials. In Experiment 1B, we also added reward and punishment as a possible task feature that could influence post-error slowing. However, a Bayesian ANOVA with experiment as a between-subject factor indicated that there was no difference in the amount of slowing between the experiment without reward/punishment (Experiment 1A; for further information see OSF) and the experiment with reward/punishment (Experiment 1B), as the ANOVA yielded moderate evidence for the null hypothesis (BF_{10} = 0.0169) in the start RT and anecdotal evidence (BF_{10} = 0.308) in the choice RT for the interaction between experiment and previous outcome. Thus, the delivery of reward and punishment could not explain the discrepancies between studies, or more generally, why people slow down or speed up after errors or sub-optimal outcomes. For this reason, we also omitted the reward/punishment manipulation from the remaining experiments.

Experiment 2

In this experiment, we tested whether having *objective* control over the outcome determines how people respond to sub-optimal outcomes. Specifically, we aimed to test the idea that people would slow down after errors in situations with control over the outcome, but not in situations without control over the outcome. In the latter case, we predicted that they would even speed up, in line with our previous gambling results. To test this, we again used a perceptual decision-making task. Compared to the first two experiments, participants only had to distinguish between dark and light gray (i.e. a two-choice task instead of a six-choice task). Importantly, we introduced two different blocks: 'easy' and 'hard'. In the 'easy' blocks, the dark and light stimuli were relatively easy to distinguish. In the 'hard' blocks, participants were told that it was very hard to distinguish between the two, but they did not know that in fact, we always presented the same stimulus and predetermined the outcome (50% correct and 50% incorrect trials). These 'hard' blocks were therefore like gambling trials (i.e., the outcome was not controllable), except that participants could still have thought that it was possible to detect a difference (at least initially) based on the set-up of the experiment and the task instructions at the beginning. If slowing vs. speeding was indeed determined by participants' having control over the outcome, we expected post-error slowing in the 'easy' blocks and in the first 'hard' blocks. We also expected that as the experiment progressed, partici-

pants would realize that they were not in control over the outcome in the 'hard' blocks. Based on previous literature (Damaso et al., 2020) we predicted that not having control over the outcome in the hard blocks would result in post-error speeding (in contrast to the "easy" blocks, in which we still predicted post-error slowing). In order to test our manipulation, we asked participants how much they felt in control over the outcome at the end of the experiment. Note that post-hoc analyses of these 'illusion-of-control' questions at least suggested that participants felt more in control in the 'easy' compared to the 'hard' blocks (see supplementary material on OSF).

Method

Sample Size

To determine our sample size, we focused on the three-way interaction between previous trials (error vs. correct), condition ('easy' vs. 'hard') and part of the experiment (first half vs. second half). Note that the proper experiment only started after the individual difficulty level was determined for every participant. Thus, our staircase procedure blocks were not included in the analyses. Specifically, we predicted that post-error slowing would be observed throughout the whole experiment in the 'easy' condition, whereas it would be substantially reduced (and even reversed) in the second half of the experiment for the 'hard' condition. To test this hypothesis, we conducted a Bayesian repeated measures ANOVA with these three factors and determined the BF_{10} for the three-way interaction. Our minimal sample size was 50 participants. However, this interaction did not show a strong evidence for either hypothesis ($BF_{10} > 10$ or $BF_{10} < 1/10$), therefore, we increased the sample size in steps of 10 until we reached our maximum sample size of 100. Note that we fixed the random seed in order to have the same BF analyses during the Bayesian sampling.

Participants

In total 100 participants (recruited via Prolific) completed the entire online experiment (38 females, 4 non-binary, $M = 28.16$ years, $SD = 9.62$ years, *range* = 18-64). The participation requirements and the payment structure were the same as in Experiment 1A. Participants agreed to the consent form before starting the experiment.

In addition to the 100 participants included in the data analyses, an additional 27 participants did the entire experiment and got paid for their participation, but were excluded in accordance with our preregistered exclusion criteria: three participants started the experiment again, 23 participants did not have enough trials left for the analyses (see Analyses section) and one participant reported they had problems with running the study. Moreover, one participant was rejected (in accordance with the Prolific guidelines) because they never provided a data set or completion code. An additional 27 participants signed up for the experiment but never started or completed it. All data were acquired between the 7th and the 12th October 2021.

Apparatus and Stimuli

In this study, participants performed a different color discrimination task than in Experiment 1A and 1B. They had to indicate whether the square (size 94.5 px) presented in the center of the screen was dark gray or light gray. They had to press the left arrow key if they thought the stimulus was dark, and the right arrow key if they thought it was light. In the 'easy' blocks, we presented on each trial one of two stimuli that were distinguishable from each other (see next paragraph), whereas in the 'hard' blocks, we always presented the same stimulus (exactly between dark and light; i.e., RGB 127, 127, 127) without telling participants (in the instruction they were presented with two very similar, yet different stimuli as an example).

For the 'easy' blocks we used a three-up/one-down staircase procedure to determine the individual difficulty for every participant in order to make sure that there were sufficient trials for the post-error analysis. Specifically, after every three correct trials, the RGB difference between dark (e.g. RGB 135, 135, 135) and light (e.g. RGB 119, 119, 119) was reduced by one RGB step (RGB 134, 134, 134 and RGB 120, 120, 120, respectively); but after every error, the RGB difference was increased by one step. Like this, we aimed for an average accuracy of about 80% (exact accuracy was 81.39%) for the first two blocks. The RGB values obtained at the end of the second block were used for the actual experimental procedure. Note that we piloted this procedure with 10 participants before the registration to make sure we had approximately 20% error trials per participant. These data were not included in the final analyses but are made available as well.

Procedure

Each trial started with the message "Press the space bar to start the next trial." After subjects pressed the space bar, the target was shown until the participant gave a response. After each trial, the word 'correct' or 'incorrect' (depending on their response) was presented as feedback for 500 ms, after which the start-up message for the next trial was shown.

The experiment started with ten 'easy' practice trials with the fixed RGB values of 87, 87, 87 (dark) and 167, 167, 167 (light) for the stimuli. The stimulus-response mapping remained at the bottom of the screen, to allow participants to practice the mapping. These trials were followed by two 'easy' practice blocks of 32 trials each (without the reminder about stimulus-response mappings). In these two blocks, we used the staircase procedure to determine the individual difficulty (i.e., the individual RGB values for the rest of the experiment) for every participant. This was followed by 16 short experimental blocks (eight 'easy' and eight 'hard' blocks) of 32 trials each. 'Easy' and 'hard' blocks alternated, always starting with the 'easy' blocks. Hence, the task consisted of 586 trials in total.

At the end of the experiment, the participants answered five questions about their estimated ability to do the task in both the 'easy' and the 'hard' blocks, with a visual slider response (like Langer & Roth, 1975, in their illusion-of-con-

trol experiment). These questions were included to explore the relationship between slowing/speeding and the feeling of control. Analysis revealed that paricipants felt more in control in the 'easy' compared to the 'hard' condition. However, we did not find any relationship between slowing/speeding and the feeling of control. Further details on these analyses can be found in the online supplementary material on OSF. Participants again filled in the UPPS-P questionnaire, but we did not use these data here (see above).

Analyses

Trial and participant exclusion criteria were the same as in Experiment 1A and 1B. Additionally, we replaced participants if they had fewer than 28 trials (5% of all trials) for the post-error slowing estimate (i.e. correct trials that follow an error) in the 'easy' condition after these exclusions.

We conducted Bayesian repeated measures ANOVAs with the factors condition (hard vs. easy), previous outcome (error vs. correct) and part of the experiment (first half vs. second half). For the Bayes factors, we used the default prior widths as used in R and JASP. For Bayesian ANOVAs the prior width is 0.5. We conducted the Bayesian ANOVA with JASP, using the Bayesian repeated measures ANOVA calculating the BF with effects compared to a null model. Specifically, we report the inclusion Bayes factors across matched models. The preregistered follow-up t-tests can be found in the online supplementary material.

Results

The descriptive statistics can be found in Figure 2. First, the evidence for an effect of previous outcome was anecdotal (BF_{10} = 0.48), showing no overall difference between trials following errors (M = 312 ms; SD = 185 ms) and trials following correct trials (M = 302 ms; SD = 170 ms). Second, we found moderate evidence (BF_{10} = 3.48) for an effect of condition, indicating faster start latencies in the easy condition (M = 301 ms; SD = 169 ms) compared with the hard condition (M = 315 ms; SD = 185 ms). Third, we found extreme evidence (BF_{10} > 100) for the effect of part of the experiment, showing slower start latencies in the first half of the experiment (M = 330 ms; SD = 193 ms) compared with the second half (M = 286 ms; SD = 155 ms). Fourth, the interaction between condition and part of the experiment yielded anecdotal (BF_{10} = 0.38) evidence. Finally, the interaction between previous outcome and part of the experiment (BF_{10} = 0.13), the interaction between previous outcome and condition (BF_{10} = 0.23) and the three-way interaction (BF_{10} = 0.18) yielded moderate evidence for the null hypothesis. For detailed inferential statistics, see Table 2.

In the choice RT the pattern was a bit different. First, we found moderate evidence (BF_{10} = 3.35) for an effect of previous outcome, suggesting *faster* response times for trials following error trials (M = 701 ms; SD = 304 ms) compared to trials following correct trials (M = 722 ms; SD = 292 ms). For an effect of condition, we found moderate evidence (BF_{10} = 0.28) in favor of the null hypothesis, suggesting no difference between easy (M = 717 ms; SD = 262 ms) and hard trials (M = 705 ms; SD = 331 ms). Third, evidence for an effect of part of the experiment was extreme (BF_{10} > 100), showing the same pattern as in the start RT: participants were slower in the first part of the experiment (M = 760 ms; SD = 285 ms) compared with the second part of the experiment (M = 662 ms; SD = 286 ms). Fourth, the interaction between condition and part of the experiment yielded anecdotal evidence (BF_{10} = 1.42). Finally, all other interactions yielded moderate evidence in favor of the null hypothesis. For detailed inferential statistics, see Table 2.

Discussion

Contrary to our predictions, we did not find a difference between the 'easy' and the 'hard' conditions. Instead, we found that (at least for the choice RT data) there seemed to be a general shift towards post-error *speeding* in the later stage of the experiment. Thus, it seems that introducing a condition in which participants did not have control over the outcome (i.e., the 'hard' blocks) also influenced the condition in which they did have control (i.e., the 'easy' blocks) when these conditions were intermixed. This idea is supported by the fact that we found post-error *slowing* (start RT: 170 ms; choice RT: 59 ms; both BF_{10} were bigger than 10) in the 'easy' training blocks, so before the hard blocks were introduced. These post-hoc analyses can be found in the online supplementary materials. To further investigate the influence of a hard condition on post-error slowing/speeding, we conducted a third experiment.

Experiment 3

In Experiment 3, we used a between-subject design. The first group exclusively performed the 'easy' blocks of Experiment 2. The second group exclusively performed the 'hard' (but 'supposedly doable') blocks of Experiment 2. Additionally, to further clarify the role of task context and to what extent this might be mediated by task instructions, we added a third group of participants. These participants were also presented with the stimuli from the 'hard' blocks of Experiment 2, but they were told that it was impossible for the human eye to see the difference between them. In other words, they were told they had to guess. Thus, in Experiment 3, we compared the conditions 'easy', 'hard but doable' and 'impossible' with each other.

Based on our earlier studies, we assumed that the three groups would differ in the size and direction of the post-error effects. For the 'easy' group, we expected a post-error slowing effect in both the start RT and the choice RT (as in Experiment 1A and 1B). For the 'hard but doable' group, we expected that participants would initially slow down after an error but that this post-error slowing would disappear throughout the experiment (and possibly, leading to post-error speeding towards the end). Finally, for the 'impossible' group, we expected participants to speed up after suboptimal outcomes (in this case errors) from the beginning of the experiment.

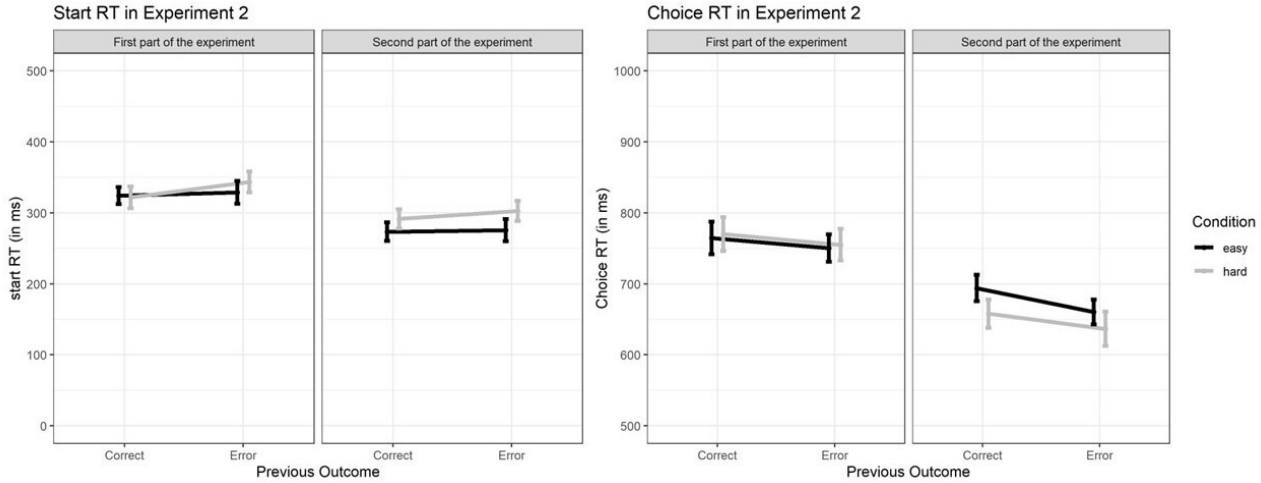

Figure 2. Start RT and choice RT as a function of previous outcome, condition and part of the experiment in Experiment 2.

The error bars reflect the 95% confidence intervals.

Table 2. Inferential statistics of the latency data in Experiment 2. Here we conducted an ANOVA with the factors previous outcome, condition and part of the experiment.

Effect	DFn	DFd	MSE	*F*	*p*	ges	BF_{10}
Start RT							
previous outcome	1	99	5749.59	3.42	.07	0.00	0.48
condition	1	99	5954.52	6.93	.01	0.00	3.48
part of the experiment	1	99	13770.53	27.87	< .001	0.03	> 100
previous outcome x condition	1	99	3145.98	2.68	.11	0.00	0.23
previous outcome x part of the experiment	1	99	2458.19	0.80	.37	0.00	0.13
condition x part of the experiment	1	99	3462.51	3.99	.05	0.00	0.38
previous outcome x condition x part of the experiment	1	99	2401.71	0.31	.58	0.00	0.18
Choice RT							
previous outcome	1	99	17253.70	5.17	.03	0.00	3.35
condition	1	99	18410.76	1.66	.20	0.00	0.28
part of the experiment	1	99	20257.63	94.50	< .001	0.05	> 100
previous outcome x condition	1	99	5460.04	0.30	.58	0.00	0.12
previous outcome x part of the experiment	1	99	7454.79	8.26	.01	0.00	0.15
condition x part of the experiment	1	99	6494.40	1.25	.27	0.00	1.42
previous outcome x condition x part of the experiment	1	99	4597.98	0.45	.51	0.00	0.19

Note: ges = Generalized Eta-Squared measure of effect size; BF_{10} = Bayes Factor 10; To determine statistical significance, we used an alpha of .05.

Method

Sample Size

We changed our procedure to determine sample size as we used a between-subjects design in Experiment 3. For each group, we first calculated the post-error difference scores in start RT (i.e., mean start RT after correct responses – mean start RT after errors) for each participant. The difference scores were then compared between groups, using three two-tailed independent t-tests ('easy' vs. 'hard but doable', 'hard but doable' vs. 'impossible', and 'easy' vs. 'impossible'). We started with 50 participants per group (150 in total), and set the maximum (preregistered) sample size to 200 per group (600 in total). All three tests needed to show a BF_{10} of > 10 or < 1/6, otherwise we increased the sample size in steps of 25 participants per group (75 in total), until reaching the crucial Bayes Factors or the maximum sample size (N = 600). All BFs reached the thresholds at 525 participants.

We used asymmetrical BF boundaries (i.e., BF_{10} > 10 for the alternative hypothesis, but BF_{10} < 1/6 for the null hypothesis). Like this, we increased the efficiency of the experiment by reducing the required sample size because ob-

taining strong evidence for the null hypothesis (e.g., BF_{10} < 1/10) requires a relatively large sample size (Brysbaert, 2019; Schönbrodt & Wagenmakers, 2018).

Participants

In total 525 participants (175 per group; recruited via Prolific) completed the entire online experiment (293 females, 10 non-binary; *M* = 30.49 years, *SD* = 11.56 years, *range*= 18-83; note: one participant indicated 369 as their age, so they were excluded from the age analyses). The participation requirements and the payment structure were the same as in Experiment 1A. Participants were paid 1.50 £ as the experiment lasted 15 minutes. Participants agreed to the consent form before starting the experiment.

In addition to the 525 participants included in the analyses, an additional 15 participants did the entire experiment and got paid but were excluded in accordance with preregistered exclusion criteria: four participants started the experiment again, ten participants did not have enough trials left for the analyses and one participant had more than 5% missing trials. Moreover, one participants' submission was rejected (in accordance with Prolific guidelines) because they never provided a data set or completion code. An additional 90 participants signed up for the experiment but never started or completed it. All data were acquired between the 17th and the 23rd November 2021.

Apparatus and Stimuli

We used the same apparatus and stimuli as in Experiment 2. Importantly, we presented the different conditions to different groups of participants in a between-subject design: in the 'easy' condition, we presented on each trial one of two stimuli that were easily distinguishable from each other, whereas in the two hard conditions, we always presented the same stimulus (exactly between dark and light; i.e., RGB 127, 127, 127) without telling participants. In the latter two (hard) conditions, the outcome of each trial ('correct' or 'incorrect', with an equal probability) was predetermined .

Procedure

Participants were randomly allocated to one of the three conditions: 'easy', 'hard but doable' and 'impossible'. For all three groups, the experiment started with two training blocks consisting of 32 trials each. In the 'easy' group we used a staircase procedure in these first two blocks to determine the individual difficulty for every participant (as in Experiment 2). In the 'hard but doable' and 'impossible' groups, participants immediately got the actual 'hard' stimulus in training blocks. For all groups, the training blocks were followed by 8 short 'experimental' blocks of 32 trials each.

The trial procedure was the same as in Experiment 2. Importantly, the 'hard but doable' and 'impossible' groups did the exact same task, but were presented with different instructions at the beginning of the experiment. The 'hard but doable' group was instructed that the task was hard but doable, whereas the 'impossible' group was instructed that the difference between the stimuli was not detectable for the human eye; therefore, the task was like a coin-flip and they had to guess.

At the end of the experiment, participants answered five questions about their estimated ability to do the task (like Langer and Roth (1975) in their illusion-of-control experiment) with a visual slider. Again, we did not find any relationship between the feeling of control and the amount of speeding/slowing, nor conclusive evidence for a difference between the 'hard but doable' and 'impossible' groups. However, participants felt more in control in the easy group compared to the two hard groups (e.g., BF_{10} > 100). Further details can be found in the online supplementary materials on OSF. Finally, participants filled in the UPPS-P questionnaire. Again, we did not use the UPPS-P data in this study.

Analyses

The inference criteria and the trial exclusion criteria were the same as in Experiment 2. In contrast to Experiment 2, we included the first two blocks in the analyses (even though these blocks were used to determine the individual difficulty in the 'easy' condition) as the first two blocks seemed to be crucial in the 'hard but doable' condition.

We performed three two-tailed independent t-tests on the start RT difference scores. Moreover, we explored the evolution of the post-error latency difference score over the course of the experiment for each group separately. On the basis of the orienting account one would predict that in the easy condition, post-error slowing decreases when errors accumulate. In the hard but doable condition, we assumed that participants might realize at some point that they do not have control over the outcome and that this would alter how they respond to errors, a pattern we did not expect in the impossible condition due to the explicit instruction. First, we calculated the post-correct mean start RT for each block separately. Note that due to the short blocks it was unlikely that participants sped up significantly during the block. Then we took the start RT after an error and subtracted this from the post-correct mean start RT of the corresponding block. Like this, we had individual post-error slowing scores for each error. We also assigned an error count to each error depending on its occurrence during the experiment. The first error got number 1, the second error got number 2 etc. This was then analysed using a linear mixed model. As participants in the easy condition did not have an equal amount of errors, we used the z-score of the error count for these analyses.

Results

The descriptive statistics for the start RT and choice RT can be found in Figure 3. In the start RT, we found extreme evidence for post-error *slowing* in the easy group (mean difference between post-error and post-correct trials = 29 ms, BF_{10} > 100). By contrast, we found extreme evidence for post-error *speeding* in the 'hard but doable' group (difference = 30 ms, BF_{10} > 100) and in the 'impossible' group (dif-

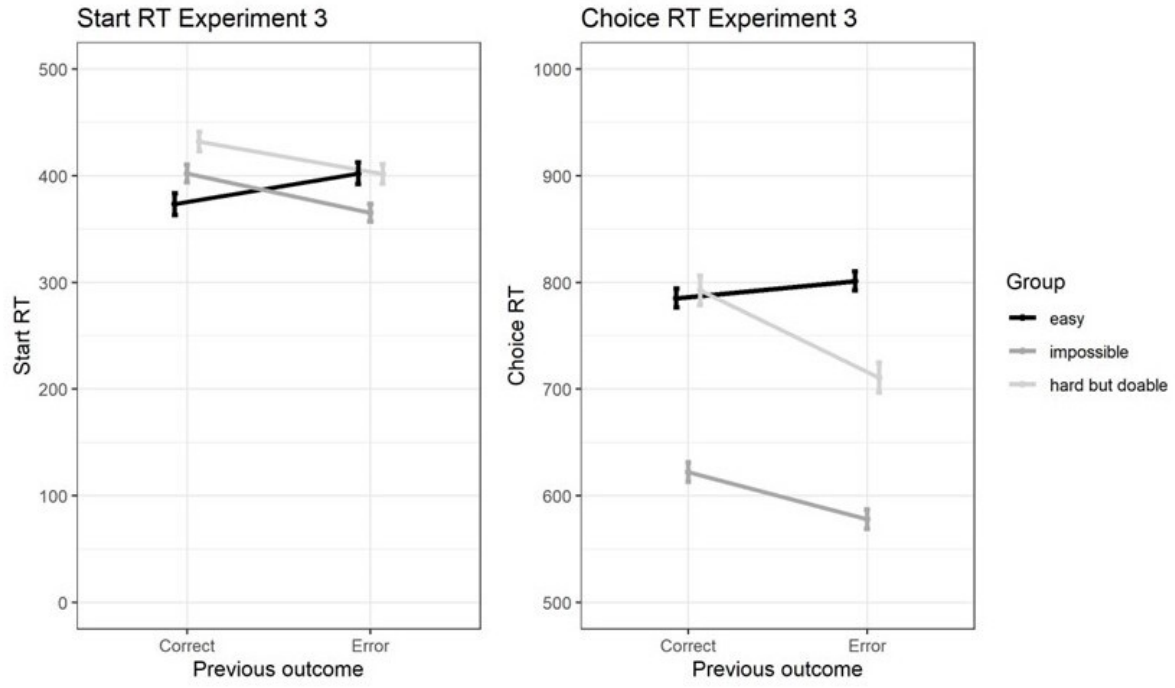

Figure 3. Start RT and choice RT as a function of previous outcome (traditional measure of post-error slowing) and group in Experiment 3.

The error bars reflect the 95% confidence intervals.

Table 3. Pairwise comparisons of the post-error slowing difference score between the three groups.

	diff	Lower CI	Upper CI	df	*t*	*p*	BF_{10}	*gav*
start RT								
'impossible' vs. 'hard but doable'	-6.23	-23.44	10.98	344.84	-0.71	.48	0.15	.08
'easy' vs. 'impossible'	65.54	47.04	84.04	332.33	6.97	< .001	> 100	.74
'easy' vs. 'hard but doable'	59.31	40.06	78.56	342.73	6.06	< .001	> 100	.65
choice RT 'impossible' vs. 'hard but doable'	37.34	13.82	60.85	300.59	3.13	.00	12.26	.33
'impossible' vs. 'easy'	60.3	42.44	78.16	347.340	6.64	< .001	> 100	.71
'easy' vs. 'hard but doable'	97.64	74.42	120.85	293.00	8.28	< .001	> 100	.88

Note: diff = difference; CI = confidence interval (95%); BF_{10} = Bayes Factor 10; g_{av} = Hedge's average *g*; To determine statistical significance, we used an alpha of .05.

ference = 36 ms, BF_{10} > 100). The Bayesian independent t-tests revealed extreme evidence (BF_{10} > 100) for a difference in post-error scores for the easy group on the one hand, and the two hard groups on the other hand (see Table 3). However, there was strong evidence (BF_{10} = 0.15) for no difference between the 'hard but doable' and 'impossible' groups.

In the choice RT, we found anecdotal evidence for post-error *slowing* in the easy group (difference between post-error and post-correct trials = 16 ms, BF_{10} = 1.997) but we found strong evidence for post-error *speeding* in the 'hard but doable' group (difference = 81 ms, BF_{10} > 100) and in the 'impossible' group (difference = 44 ms, BF_{10} > 100). Thus, here we found strong to extreme evidence for a difference between all three groups. For detailed inferential statistics see Table 3.

In order to explore post-error slowing and speeding effects as errors accumulate, we measured the difference in start RT between post-error and post-correct trials by error count per group. In order for the model to converge, we used z-scores of the error counts. In the start RT in the 'easy' group, we found an effect of error count, $\chi^2(1) = 7.09$, $p = .007$, showing that participants *slowed down* after errors in the beginning of the experiment but less so when errors accumulated (b = -13.12). Surprisingly, participants in the 'hard but doable' group *sped up* after error trials in the beginning but this effect was also reduced towards the end of experiment (b = 16.73), $\chi^2(1) = 10.86$, $p < .001$. The 'impossible' group showed a similar numerical pattern as the 'hard but doable' group (b = 8.98), which was however not significant, $\chi^2(1) = 2.70$, $p = .100$. In the choice RT, the nu-

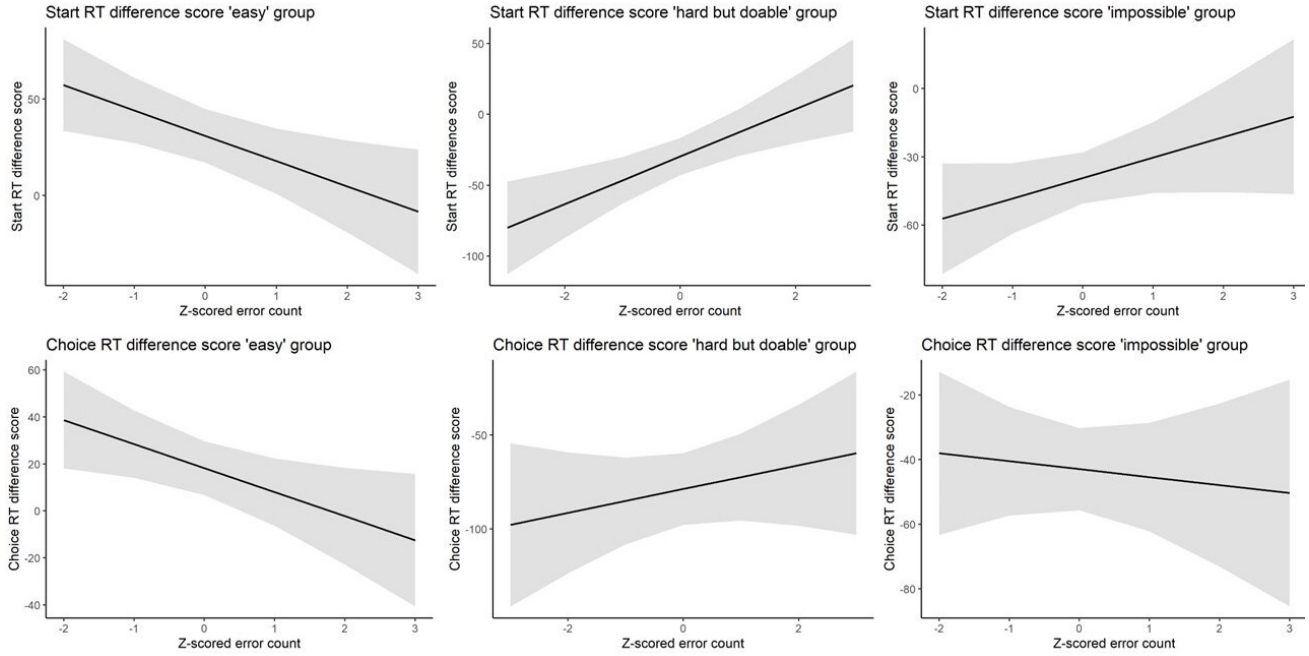

Figure 4. Start RT and choice RT difference score predicted by the z-scored error count in Experiment 3.

merical pattern was the same as in the start RT but the effect of error count was only significant in the 'easy' group (b = -10.22), $\chi^2(1) = 5.47$, $p = .019$, but not for the 'hard but doable' group (b = 6.31), $\chi^2(1) = .92$, $p = .337$, or the 'impossible' group (b = -2.45), $\chi^2(1) = 0.19$, $p = .659$.

Discussion

Our between-group comparison revealed that participants who had control over the outcome (i.e., the 'easy' group) slowed down after committing an error. However, participants who did not have control over the outcome (i.e., the 'hard but doable' and 'impossible' groups), sped up after errors. Our results suggest that control over the outcome is indeed a crucial factor in determining whether participants speed up or slow down after sub-optimal outcomes. This difference between conditions further supports the conclusion of Experiment 2 that introducing trials in which participants did not have control over the outcome also influenced performance in the condition in which they did have control over the outcome. After all, when controllability over the outcome was manipulated within participants (Experiment 2), we observed post-error speeding in both conditions (i.e., there was no effect of condition); but when controllability over the outcome was manipulated between participants (Experiment 3), post-error speeding/slowing did depend on the condition. In other words, the overall task context also seems to play a role. Follow-up linear mixed effect models on the difference between post-error and post-correct trials showed that post-error *slowing* decreased over time, as predicted by the orienting account. However, post-error *speeding* also dissipated over time. Moreover the choice RT data suggests that instructions ('hard but doable' vs. 'impossible') had an influence as well: the speeding in choice RTs was more pronounced in the 'hard but doable' group compared with the 'impossible' group, which could indicate that participants in the former group were more 'frustrated' after an error than the participants in the latter group (even though the number of errors was the same).

General Discussion

Many studies have reported that people slow down after errors, but with a few notable exceptions. Furthermore, findings from the gambling literature indicate that slowing is not necessarily observed after other types of sub-optimal outcomes. Instead, people often speed up if they lose points or money. This raises the question of what task characteristics determine whether people speed up or slow down after sub-optimal outcomes. Here we further tested the idea that *objective* controllability over the outcome might play a critical role not only in gambling related contexts (Dyson et al., 2018) but also in tasks in which we usually observe post-error slowing. But before we could do so, we first had to rule out other basic task characteristics as explanations for slowing or speeding after sub-optimal outcomes.

In Experiments 1A-B, we added a self-pace element to a color discrimination task in order to see whether speeding vs. slowing depends on the self-pace element. After all, previous work indicates that the pace and timing of events could influence post-error slowing. Additionally, we added reward in Experiment 1B for correct responses and punishment for incorrect responses, as people might respond differently to motivationally salient reward/punishment (compared to less salient errors in cognitive psychology experiments). In both experiments, we observed post-error slowing. Making the task self-paced thus does not abolish (or reverse) post-error slowing effects. Furthermore, a between-experiment comparison revealed that adding re-

ward/punishment did not modulate post-error slowing (much) either.

In Experiments 2 and 3, we therefore focused on the main feature of interest, namely controllability over the outcome. In Experiment 2, we used a within-subject design. Participants performed a color discrimination task in two different conditions: in the 'easy' blocks, they could distinguish between the stimuli, whereas in the 'hard' blocks, they had to guess (as we always presented the same stimulus, with a predetermined outcome). As soon as we introduced the 'hard' condition, post-error slowing was abolished and we even observed a shift towards speeding. This effect was observed in both conditions, indicating that the effect of control over the outcome might be more global than we initially assumed. In line with the assumption that control over the outcome has an influence on post-error slowing, Regev and Meiran (2014) found post-error slowing in more complex situations that require more control (i.e., high control *demands*) compared to low control demands. Interestingly, this post-error slowing effect was not observed when these conditions were intermixed, also suggesting a global influence of the experimental context.

In order to further test the effect of control over the outcome, we ran another experiment with a between-subject design. Each participant was assigned to one of three conditions: 'easy' (as in Experiment 2), 'hard but doable' (as in Experiment 2), or 'impossible' (the same as the 'hard' condition, except that the participants were told from the beginning that it would be impossible to distinguish between the stimuli). We found that participants from the easy group – who had indeed control over the outcome – slowed down after errors. These participants could improve their performance on the task by accumulating more evidence. However, participants in the other two groups – who had an error rate of 50% and no control over the outcome as the outcome was predetermined – sped up after error trials. These participants might have realized that no matter how much evidence they accumulated, they would not be able to improve their performance on the task and therefore sped up. Thus, it seems that the controllability over outcome is indeed the crucial difference between tasks or situations in which we observe slowing after sub-optimal outcomes compared to tasks or situations in which we observe speeding after sub-optimal outcomes.

At first sight, the findings of Experiments 1A and 1B also appear consistent with the orienting account. That is, in these experiments, we observed post-error slowing in start and choice RTs, combined with an increased error rate for post-error trials. Note that the results on the robust measure of post-error slowing (see supplementary material on OSF) indicate that this overall patterns is not due to prolonged lapses in attention or global fluctuations in task performance. Furthermore, the slowing in the start RT (i.e., for the first response after the error) seemed to be more pronounced than in the choice RT (i.e., the second response after the error). This is also in line with the idea that post-error effects are due to the (re)orienting of attention (Notebaert et al., 2009; Wessel, 2018), at least when there is insufficient time for control adjustments. In this respect, it is interesting to note that our task was self-paced, so participants could have had enough time to reorient the attention back to the main task and make control adjustments if they wished to do so. Here we can only speculate why they did not do this. To our knowledge this is the first study investigating post-error adjustments in an entirely self-paced task. One possibility is that in a self-paced task context, there is a trade-off between slowing the task pace (and make control adjustments) and speeding up in order to increase the possible reward rate (see e.g. Bogacz et al., 2006), which is similar to the suggestion by Damaso et al. (2020). Importantly though, the main results of Experiments 2 and 3 appear inconsistent with the orienting account. After all, the orienting account would predict no slowing but also no speeding in the hard and impossible conditions, as correct and incorrect trials occurred with equal probability. Of course, one could argue that the difference between the 'easy' and the two 'hard' conditions in Experiment 3 could still be explained by the frequency of errors: in the 'easy' conditions, errors are rare and might lead to an orienting response, which would explain the slowing (and decreased accuracy). However, errors were equally 'rare' in the easy condition of Experiment 2 and yet participants numerically sped up after introducing the 'hard' condition. So overall, it appears that the orienting account alone cannot explain the shift from slowing to speeding. Instead, we propose that control over the outcome is the main feature that determines how people respond to errors (instead of frequency of error).

In fact, one could even argue that control over the outcome might have also played a role in the original findings by Notebaert et al. (2009), which led to the proposal of the orienting account. In their study, they used a procedure in which they made sure that participants made a great amount of errors to investigate the influence of error frequency. They found that participants in the condition with more errors (65% errors or 35% correct, only slightly above the chance rate of 25% correct) did not show post-error slowing. Based on our findings, one could speculate that these results were partially influenced by the control over the outcome a participant could exert. After all, in order to achieve a higher error rate, the stimulus evidence quality became very low and participants' task was close to gambling, in the sense they only had little control over the outcome.

Thus, our findings are in line with the results of Damaso et al. (2020) who proposed that speeding after errors might be because participants realized that more response caution (i.e., slowing) would not lead to increased accuracy. Due to the frequent 'disappointments', participants might have gotten impatient, showing less response caution ('recklessness' as Williams et al., 2016, called it). This post-error 'recklessness' seems to be akin to the 'frustration effect' that has been described before in the animal learning literature (i.e., when a blocked reward – at least initially –leads to invigorated behavior). We also observed this in the 'hard but doable' group where participants showed a bigger post-error speeding effect in the choice RT than the participants in the 'impossible' group: in the 'hard but

doable' group participants might still assume they have control over the outcome (at least initially), thus, a sub-optimal outcome is even more 'frustrating' and leads to even more 'recklessness'. Such speeding may seem maladaptive but it could be an adaptive process after all. For example, in line with the reward-rate optimization idea (Bogacz et al., 2006), participants might be able to increase the opportunities of success by increasing the amount of trials that can be completed in a certain amount of time.

Extending these assumptions to an even broader context, this also suggests that frustration as described here can be adaptive as well (Amsel, 1992). As already mentioned, frustration can help closing the gap between the current state and the goal or simply help to escape a negative affective state through invigorated behavior (Carver & Scheier, 1990; Frijda, 2010). Importantly, the adaptivity seems to be dependent on the context in which frustration is displayed. If the outcome or environment are controllable and predictable, rash action is likely to be maladaptive. However, when the environment is unpredictable or uncontrollable, sometimes acting rashly (i.e., more 'impulsive') can be beneficial (for similar examples in developmental and animal research see e.g., Dingemanse et al., 2004; Kidd et al., 2013).

Surprisingly, Experiment 3 showed that the effects in all three groups dissipated over time. Although the decreasing post-error slowing in the easy condition can be explained in terms of decreased orienting to accumulating errors, the pattern in the two difficult condition is more surprising. One tentative explanation is that post-error speeding dissipates over time due to learned helplessness following frustration. For instance, Mikulincer (1988) showed that participants who are repeatedly exposed to unsolvable problems initially tried harder but eventually give up. We observed a similar pattern in an earlier gambling study (Eben et al., 2020): in a not very engaging gambling task, we found general speeding on gambling trials (compared with non-gambling trials); furthermore, participants initially showed more speeding after (gambled) losses compared to (gambled) wins, but this effect dissipated over time. It seems that we observe a similar pattern here. As the task in Experiment 3 was not very engaging, participants might have stopped caring about the outcomes of their actions as the experiment progressed, resulting in reduced difference between post-error and post-correct trials. Another indication that participants eventually gave up on doing the task is a general speeding towards the end of the experiment (see Figures 2 and 3 in the supplementary materials). Whether this motivational account (or more specifically, reduced motivation as the experiment progresses) could also explains the reduced post-error slowing in the easy condition needs further testing though.

In summary, we propose that both slowing and speeding in response to sub-optimal outcomes can be considered as goal-directed behaviors. Our main assumption is that outcomes of actions are constantly compared to the goal (which can be responding correctly or obtaining a reward). In the case of a sub-optimal outcome, there is a discrepancy between the outcome and the goal, which triggers adjustments of behavior. This can either lead to a decrease of "effort", if the individual is doing well in reaching their goal, or it can lead to "trying harder" if the individual is not doing well in reaching their goal (Carver, 2006). But what constitutes "trying harder" and more specifically, which action tendencies will have the highest utility (Moors et al., 2017), will strongly depend on the overall context. In a context where participants have control over the outcome, trying harder means slowing down and changing task settings (see also Dyson et al., 2018). By contrast, in contexts in which participants have no control, there are not many options to "try harder" and slowing down would be time consuming and even ineffective (Damaso et al., 2020); thus, in such situations, it might seem more advantageous to speed up to have more possible successes in the same amount of time or to simply to minimize the time on the task (given that successes are random; Damaso et al., 2020). In sum, we propose that both slowing (response caution, Wessel, 2018) and speeding ('recklessness' or 'impulsive action', Frijda, 2010; Williams et al., 2016) might be goal-directed and adaptive (or maladaptive) depending on the context (Damaso et al., 2020; Wessel, 2018).

Author Contributions

CE and LV initiated the first study. CE, LV and FV designed the first study and all authors contributed to the design of the following studies. CE and LV programmed the experiments. CE, LV and FV performed the analyses. CE and FV wrote the first draft and LV, ZC, WN, and II provided critical revisions on the manuscripts. All authors approved the final version. Note that LV and ZC acted as 'co-pilots' for the whole project.

Funding Information

This work was supported by an ERC Consolidator grant awarded to FV (European Union's Horizon 2020 research and innovation programme, grant agreement No 769595).

Competing Interests

The authors have no competing interests.

Data Accessibility Statement

All raw and processed data, code, and materials of all experiments can be found on OSF (https://osf.io/azjey/). The preregistrations for Experiments 1B-3 are also on OSF (Experiment 1B, Experiment 2, Experiment 3).

Submitted: April 22, 2022 PDT, Accepted: February 21, 2023 PDT

References

Amsel, A. (1958). The role of frustrative nonreward in noncontinuous reward situations. *Psychological Bulletin*, *55*(2), 102–119. https://doi.org/10.1037/h0043125

Amsel, A. (1992). *Frustration Theory: An Analysis of Dispositional Learning and Memory*. Cambridge University Press. https://doi.org/10.1017/cbo9780511665561

Barcelo, F., Escera, C., Corral, M. J., & Periáñez, J. A. (2006). Task Switching and Novelty Processing Activate a Common Neural Network for Cognitive Control. *Journal of Cognitive Neuroscience*, *18*(10), 1734–1748. https://doi.org/10.1162/jocn.2006.18.10.1734

Beatty, P. J., Buzzell, G. A., Roberts, D. M., & McDonald, C. G. (2018). Speeded response errors and the error-related negativity modulate early sensory processing. *NeuroImage*, *183*, 112–120. https://doi.org/10.1016/j.neuroimage.2018.08.009

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700–765. https://doi.org/10.1037/0033-295x.113.4.700

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652. https://doi.org/10.1037/0033-295x.108.3.624

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). https://doi.org/10.5334/joc.72

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*(1), 395–411. https://doi.org/10.32614/rj-2018-017

Carver, C. S. (2006). Approach, Avoidance, and the Self-Regulation of Affect and Action. *Motivation and Emotion*, *30*(2), 105–110. https://doi.org/10.1007/s11031-006-9044-7

Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, *97*(1), 19–35. https://doi.org/10.1037/0033-295x.97.1.19

Chen, Z., Doekemeijer, R. A., Noël, X., & Verbruggen, F. (2022). Winning and losing in online gambling: Effects on within-session chasing. *PLOS ONE*, *17*(8), e0273359. https://doi.org/10.1371/journal.pone.0273359

Corr, P. J., & Thompson, S. J. (2014). Pause for Thought: Response Perseveration and Personality in Gambling. *Journal of Gambling Studies*, *30*(4), 889–900. https://doi.org/10.1007/s10899-013-9395-4

Cyders, M. A., Littlefield, A. K., Coffey, S., & Karyadi, K. A. (2014). Examination of a short English version of the UPPS-P Impulsive Behavior Scale. *Addictive Behaviors*, *39*(9), 1372–1376. https://doi.org/10.1016/j.addbeh.2014.02.013

Damaso, K., Williams, P., & Heathcote, A. (2020). Evidence for different types of errors being associated with different types of post-error changes. *Psychonomic Bulletin & Review*, *27*(3), 435–440. https://doi.org/10.3758/s13423-019-01675-w

de Bruijn, E. R. A., Mars, R. B., & Hulstijn, W. (2004). "It wasn't me… or was it?" How false feedback effects performance. In M. Ullsperger & M. Falkenstein (Eds.), *Errors, conflicts, and the brain. Current opinions on performance monitoring* (pp. 118–124). MPI cognitive neuroscience. http://hdl.handle.net/2066/64750

de Heide, R., & Grünwald, P. D. (2021). Why optional stopping can be a problem for Bayesians. *Psychonomic Bulletin & Review*, *28*(3), 795–812. https://doi.org/10.3758/s13423-020-01803-x

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Dingemanse, N. J., Both, C., Drent, P. J., & Tinbergen, J. M. (2004). Fitness consequences of avian personalities in a fluctuating environment. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *271*(1541), 847–852. https://doi.org/10.1098/rspb.2004.2680

Dixon, M. J., MacLaren, V., Jarick, M., Fugelsang, J. A., & Harrigan, K. A. (2013). The Frustrating Effects of Just Missing the Jackpot: Slot Machine Near-Misses Trigger Large Skin Conductance Responses, But No Post-reinforcement Pauses. *Journal of Gambling Studies*, *29*(4), 661–674. https://doi.org/10.1007/s10899-012-9333-x

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. https://doi.org/10.1016/j.jmp.2012.04.001

Dutilh, G., Vandekerckhove, J., Forstmann, B. U., Keuleers, E., Brysbaert, M., & Wagenmakers, E.-J. (2012). Testing theories of post-error slowing. *Attention, Perception, & Psychophysics*, *74*(2), 454–465. https://doi.org/10.3758/s13414-011-0243-2

Dyson, B. J., Sundvall, J., Forder, L., & Douglas, S. (2018). Failure generates impulsivity only when outcomes cannot be controlled. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(10), 1483–1487. https://doi.org/10.1037/xhp0000557

Eben, C., Chen, Z., Billieux, J., & Verbruggen, F. (2022). Outcome sequences and illusion of control – part II: the effect on post-loss speeding. *International Gambling Studies*, *0*(0), 1–20. https://doi.org/10.1080/14459795.2022.2135227

Eben, C., Chen, Z., Vermeylen, L., Billieux, J., & Verbruggen, F. (2020). A direct and conceptual replication of post-loss speeding when gambling. *Royal Society Open Science*, *7*(5), 200090. https://doi.org/10.1098/rsos.200090

Fievez, F., Derosiere, G., Verbruggen, F., & Duque, J. (2021). *Post-error slowing reflects the joint impact of adaptive and maladaptive processes during decision making* [Tech. rep.]. Cold Spring Harbor Laboratory. https://doi.org/10.1101/2021.12.22.473805

Forster, S. E., & Cho, R. Y. (2014). Context Specificity of Post-Error and Post-Conflict Cognitive Control Adjustments. *PLoS ONE*, *9*(3), e90281. https://doi.org/10.1371/journal.pone.0090281

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Frijda, N. H. (2010). Impulsive action and motivation. *Biological Psychology*, *84*(3), 570–579. https://doi.org/10.1016/j.biopsycho.2010.01.005

Garavan, H., & Stout, J. C. (2005). Neurocognitive insights into substance abuse. *Trends in Cognitive Sciences*, *9*(4), 195–201. https://doi.org/10.1016/j.tics.2005.02.008

Hajcak, G., & Simons, R. F. (2008). Oops!.. I did it again: An ERP and behavioral study of double-errors. *Brain and Cognition*, *68*(1), 15–21. https://doi.org/10.1016/j.bandc.2008.02.118

Jentzsch, I., & Dudschig, C. (2009). Short Article: Why do we slow down after an error? Mechanisms underlying the effects of posterror slowing. *Quarterly Journal of Experimental Psychology*, *62*(2), 209–218. https://doi.org/10.1080/17470210802240655

Kidd, C., Palmeri, H., & Aslin, R. N. (2013). Rational snacking: Young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability. *Cognition*, *126*(1), 109–114. https://doi.org/10.1016/j.cognition.2012.08.004

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Langer, E. J., & Roth, J. (1975). Heads I win, tails it's chance: The illusion of control as a function of the sequence of outcomes in a purely chance task. *Journal of Personality and Social Psychology*, *32*(6), 951–955. https://doi.org/10.1037/0022-3514.32.6.951

Lenth, R. V. (2021). *Emmeans: Estimated marginal means, aka least-squares means*. [R package version 1.5.4]. https://github.com/rvlenth/emmeans

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review*, *108*(2), 393–434. https://doi.org/10.1037/0033-295x.108.2.393

Lüdecke, D. (2020). *Ggeffects: Create tidy data frames of marginal effects for ggplot from model outputs*. [R package version 1.0.1]. https://strengejacke.github.io/ggeffects/

Lüdecke, D. (2021). *Sjplot: Data visualization for statistics in social science*. [R package version 2.8.9]. https://strengejacke.github.io/sjPlot/

Mikulincer, M. (1988). Reactance and helplessness following exposure to unsolvable problems: The effects of attributional style. *Journal of Personality and Social Psychology*, *54*(4), 679–686. https://doi.org/10.1037/0022-3514.54.4.679

Moors, A., Boddez, Y., & De Houwer, J. (2017). The Power of Goal-Directed Processes in the Causation of Emotional and Other Actions. *Emotion Review*, *9*(4), 310–318. https://doi.org/10.1177/1754073916669595

Moors, A., Ellsworth, P. C., Scherer, K. R., & Frijda, N. H. (2013). Appraisal Theories of Emotion: State of the Art and Future Development. *Emotion Review*, *5*(2), 119–124. https://doi.org/10.1177/1754073912468165

Morey, R. D., & Rouder, J. N. (2018). *Bayesfactor: Computation of bayes factors for common designs*. [R package version 0.9.12-4.2]. https://richarddmorey.github.io/BayesFactor/

Nash, J. C. (2014). On best practice optimization methods in R. *Journal of Statistical Software*, *60*(2), 1–14. https://doi.org/10.18637/jss.v060.i02

Notebaert, W., Houtman, F., Opstal, F. V., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: An orienting account. *Cognition*, *111*(2), 275–279. https://doi.org/10.1016/j.cognition.2009.02.002

R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. http://www.R-project.org/

Rabbitt, P. M. A., & Phillips, S. (1967). Error-detection and correction latencies as a function of s-r compatibility. *Quarterly Journal of Experimental Psychology*, *19*(1), 37–42. https://doi.org/10.1080/14640746708400065

Rabbitt, P. M. A., & Rodgers, B. (1977). What does a Man do after he Makes an Error? An Analysis of Response Programming. *Quarterly Journal of Experimental Psychology*, *29*(4), 727–743. https://doi.org/10.1080/14640747708400645

Regev, S., & Meiran, N. (2014). Post-error slowing is influenced by cognitive control demand. *Acta Psychologica*, *152*, 10–18. https://doi.org/10.1016/j.actpsy.2014.07.006

Riesel, A., Weinberg, A., Endrass, T., Kathmann, N., & Hajcak, G. (2012). Punishment has a lasting impact on error-related brain activity: Punishment modulates error monitoring. *Psychophysiology*, *49*(2), 239–247. https://doi.org/10.1111/j.1469-8986.2011.01298.x

Robinson, D., Hayes, A., & Couch, S. (2021). *Broom: Convert statistical objects into tidy tibbles*. [R package version 0.7.10]. https://CRAN.R-project.org/package=broom

Rouder, J. N. (2014). Optional stopping: No problem for bayesians. *Psychonomic Bulletin & Review*, *21*(2), 301–308. https://doi.org/10.3758/s13423-014-0595-4

Schönbrodt, F. D., & Wagenmakers, E.-J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, *25*(1), 128–142. https://doi.org/10.3758/s13423-017-1230-y

Schönbrodt, F. D., Wagenmakers, E.-J., Zehetleitner, M., & Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods*, *22*(2), 322–339. https://doi.org/10.1037/met0000061

Singmann, H., Brown, S., Gretton, M., & Heathcote, A. (2020). *Rtdists: Response time distributions*. [R package version 0.11-2]. https://github.com/rtdists/rtdists/

Steinhauser, M., & Kiesel, A. (2011). Performance monitoring and the causal attribution of errors. *Cognitive, Affective, & Behavioral Neuroscience*, *11*(3), 309–320. https://doi.org/10.3758/s13415-011-0033-2

Stürmer, B. (2011). Reward and Punishment Effects on Error Processing and Conflict Control. *Frontiers in Psychology*, *2*, 335. https://doi.org/10.3389/fpsyg.2011.00335

Sullivan, R. M., Perlman, G., & Moeller, S. J. (2019). Meta-analysis of aberrant post-error slowing in substance use disorder: Implications for behavioral adaptation and self-control. *European Journal of Neuroscience*, *50*(3), 2467–2476. https://doi.org/10.1111/ejn.14229

Verbruggen, F., Chambers, C. D., Lawrence, N. S., & McLaren, I. P. L. (2017). Winning and losing: Effects on impulsive action. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(1), 147–168. https://doi.org/10.1037/xhp0000284

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q. F., Dropmann, D., Boutin, B., Meerhoff, F., Knight, P., Raj, A., van Kesteren, E.-J., van Doorn, J., Šmíra, M., Epskamp, S., Etz, A., Matzke, D., … Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, *25*(1), 58–76. https://doi.org/10.3758/s13423-017-1323-7

Wessel, J. R. (2018). An adaptive orienting theory of error processing. *Psychophysiology*, *55*(3), e13041. https://doi.org/10.1111/psyp.13041

Wickham, H. (2021). *Tidyverse: Easily install and load the tidyverse*. [R package version 1.3.1]. https://CRAN.R-project.org/package=tidyverse

Wilke, C. O. (2020). *Cowplot: Streamlined plot theme and plot annotations for ggplot2*. [R package version 1.1.1]. https://wilkelab.org/cowplot/

Williams, P., Heathcote, A., Nesbitt, K., & Eidels, A. (2016). Post-error recklessness and the hot hand. *Judgment and Decision Making*, *11*(2), 174–184. https://doi.org/10.1017/s1930297500007282

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: Confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1310–1321. https://doi.org/10.1098/rstb.2011.0416

Supplementary Materials

Peer Review History

Download: https://collabra.scholasticahq.com/article/73052-when-response-selection-becomes-gambling-post-error-slowing-and-speeding-in-self-paced-colour-discrimination-tasks/attachment/151900.docx?auth_token=eFmT_5YVE62Seyt6Kjo2