Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

In search of Perruchet's dissociation between expectancy and control in the flanker task

Ellen Christy

ellen.christy@sydney.edu.au

The University of Sydney

Hilary Don

The University of Sydney

Max Hill

The University of Sydney

Evan Livesey

The University of Sydney

Article

Keywords: conflict adaptation, congruency sequence effect, expectancy

Posted Date: July 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10176854/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on October 28th, 2026.
See the published version at https://doi.org/10.1038/s41598-026-78183-0.

Abstract

Adaptation to cognitive conflict has been demonstrated in the form of reduced congruency effects following runs of high conflict (incongruent) trials, and increased effects following runs of low conflict (congruent) trials, known as the *progressive congruency sequence effect* (*PCSE*). This phenomenon has been attributed to an expectancy for conflict, yet direct evidence linking expectancy to performance is limited. Here we leveraged a dissociation paradigm known as the Perruchet effect to test the role of expectancy in producing conflict adaptation. Using variations of the flanker task, we ran three experiments investigating how expectancy influences conflict adaptation. Experiment 1 assessed the effect of expectancy for conflict specifically on the PCSE by manipulating the feature relevant to the expectancy rating (conflict expectancy vs. response expectancy). Experiment 2 aimed to increase motivation to produce accurate expectancies by directly incentivising the rating task. Experiment 3 addressed unexpected patterns in expectancy by examining participant's ability to track conflict history. Across Experiments 1 and 2, participants did not show reliable conflict expectancy trends. Experiment 3 revealed that this may be because some participants exhibit poor explicit recall of conflict history. Importantly, the PCSE remained robust across all experiments and conditions. This series of experiments suggests that the PCSE in the flanker task operates independently of conscious expectancy, challenging expectancy-based accounts of adaptive control.

Introduction

Conflict occurs when there is a mismatch between the information we must act upon to achieve our current goals and other salient information in the current environment. Resolving conflict is often assumed to require *cognitive control*, the ability to moderate our thoughts and actions in line with our goals [1,2]. Cognitive control has been widely studied in conflict tasks. For example, in the Stroop task, where participants respond to the color of a printed word, "RED" appearing in red will yield faster responses than the word "BLUE" appearing in red, because the meaning and color are congruent in the first instance and incongruent in the second [3]. Similarly, in the flanker task, responses to the central target feature in "<<<" will be faster than those to the central target feature in "><>", because the flanking arrows are congruent with the target arrow in the first instance and incongruent in the second [4].

A considerable body of evidence demonstrates that our ability to manage conflict is shaped by our recent experience [5], and most of this evidence is consistent with the hypothesis that we adjust our cognitive control in anticipation of conflict. For instance, recent studies [6] have shown that presenting an explicit cue that signals whether the upcoming trial is likely to be high-conflict or low-conflict can influence congruency effects; a high-conflict cue reduces the congruency effect and a low-conflict cue increases it. This suggests that an expectation for conflict can change how distracting information is processed proactively. Indeed, the idea that expectancy for conflict might be crucial for cognitive control dates back to a key finding by Gratton and colleagues [7] that the congruency effect is larger after experiencing a congruent trial than after an incongruent trial. They suggested that this phenomenon, now often referred to as the *congruency sequence effect* (CSE), is driven by the expectancy for the same

type of trial as the one that was just completed [7]. Although there have been explanations for the CSE accounting for mechanisms at different levels of information processing [1,5,8–10], the notion that prediction and expectancies play an important role in control remains influential [11,12].

Although researchers have proposed that expectancy for conflict drives adaptive control, relatively few studies have attempted to manipulate and measure explicit expectancy for conflict, and study how it influences performance. A series of experiments reported by Jiménez & Méndez [13,14] investigated how recent trial history affected performance and expectancy in the Stroop task. Their design took inspiration from studies by Perruchet [15,16], demonstrating a dissociation between an individual's behaviour in anticipation of an event and their conscious expectancy of that event. Jiménez and Méndez [13] pseudorandomised the sequence of congruent and incongruent trials such that there were runs of consecutive congruent and consecutive incongruent trials adhering to a binomial distribution of two equally likely events. They then analysed performance on the Stroop task as a function of the length and nature of the preceding trials, i.e. *conflict history*. Figure 1 illustrates how run lengths of congruency may manifest in Stroop and flanker tasks, as well as an idealised version of their results. Following longer runs of incongruent trials, participants respond faster to an incongruent trial and slower to a congruent trial, demonstrating a progressive reduction in the congruency effect. Conversely, following longer runs of congruent trials, participants were slower to respond to the subsequent incongruent trial, but faster to respond to the congruent trial, demonstrating an increased congruency effect. For clarity and following Jiménez and Méndez [13], we will refer to the resulting linear trend in congruency effect as the *progressive CSE* (PCSE).

The PCSE suggests *positive recency* in how an individual responds to conflict (i.e. experiencing high conflict trials leads to stronger preparedness to deal with further high conflict). However, in contrast, participants rated their expectancy of a congruent trial higher as the number of incongruent trials increased, and lower as the number of congruent trials increased. This pattern is consistent with the *gambler's fallacy*, an expectation that, in an equal distribution of binary outcomes, an alternation is more likely following a streak of repetitions, demonstrating *negative recency* [17]. Importantly, this demonstrates that conscious expectancies do not align with the pattern of performance. Thus, Jiménez and Méndez [13] suggested that the PCSE must be driven by other adaptive control mechanisms that do not require deliberate top-down control based on the individual's current beliefs.

This dissociation may seem at odds with the presumed role of conflict expectancy in guiding cognitive control [12,18,19]. However, Jiménez and Méndez [14] noted key limits to the dissociation. Firstly, the PCSE was only observed strongly when expectancy and performance on the Stroop task were measured in separate blocks rather than concurrently. They speculated that while the conditions may already be present for participants to show negative recency in expectancy, the act of asking them to report their expectancy may strengthen their conscious prediction, which would not have been present in the regular performance blocks. This would suggest that conscious expectancies, when made salient to the individual, can play an important role in preparing for conflict, even if the PCSE is not produced by those

expectancies. Indeed, such an effect has also been observed across choice response tasks employing Perruchet's paradigm [20–22].

In addition, [23] found distinctly different patterns of expectancy, compared to Jimenez and Mendez, when they asked participants to rate expectancy for upcoming trial conflict while they performed a Stroop task. Rather than finding negative recency, they found that participants on average indicated greater expectancy for congruent trials after a run of congruent trials than after a run of incongruent trials, showing positive recency consistent with a *hot-hand fallacy* [24]. Overall, Duthoo and colleagues' [23] results indicated consistency between participants' explicit expectancies and the level of Stroop interference observed, concluding that the CSE may be strongly determined by the expectancy for repetitions of the same trial type. Thus, existing research has produced inconsistent results regarding the role of top-down expectancy in driving adaptive control, and has been constrained specifically to the Stroop task. There are therefore questions remaining regarding the generality of these findings, and how a focus on explicitly rating expectancy could change the level of adaptive control when completing conflict tasks.

The current study

The present study therefore investigated how the formation of explicit expectancies influences adaptive control in the flanker task. We used variants of the PCSE design and analysis, seeking evidence for positive and negative recency effects in the flanker effect itself and in explicit expectancy for conflict. Critically, we also manipulated how we probed for expectancy, to motivate participants to generate expectancies while they performed the flanker task.

Jiménez and Méndez [14] found that the PCSE was reduced when probing expectancies for conflict concurrently to performing a conflict task, that is, during the inter-trial interval (ITI) between Stroop trials, but not when expectancies were probed in separate blocks. This might suggest that asking participants to report their conscious expectancies fundamentally changes the processes that generate adaptive control. However, comparing blocks of trials with and without ratings confounds several factors that could impact performance and the PCSE. For instance, switching between ratings and performing the conflict task generates additional task load, which may reduce reliability of subtle RT trends [20,21]. In the current study, we took a different approach, manipulating expectancy generation in different ways. We still retain a comparison between trials that follow an expectancy rating and those that do not. However, following a design used by Lee Cheong Lem and colleagues [21], these were randomly intermixed rather than blocked.

In Experiment 1, we compared two groups asked to generate explicit expectancies about different aspects of the flanker task. One group was asked to rate their expectancy for upcoming trial conflict (whether the next trial will be "easy" or "hard"), while the other group was asked to rate their expectancy for the response type of the upcoming trial (whether the next trial will require a left or a right response). If formation of conscious expectancies regarding conflict influences the processes responsible for

adaptive control, we should see a relative reduction in the strength of the PCSE in the conflict expectancy group. In Experiment 2 we investigated how the motivation to produce a valid expectancy rating may influence the likelihood of observing negative versus positive recency in expectancy ratings, and the impact of conscious expectancies on the PCSE. In Experiment 3, we tested a hypothesis that tracking past levels of conflict is inherently more difficult than tracking other, more distinct trial features, which may explain some of the inconsistency in expectancy ratings reported in prior literature and in our experiments. These experiments address the question of what role conscious expectancies play in adaptive control, as well as test how attention and motivation influence the processes underlying conflict adaptation.

Experiment 1

Experiment 1 investigated how conscious expectancies for conflict influence the PCSE in the flanker task. One group was asked to rate their expectancy for high- versus low-conflict trials (conflict expectancy group), and another was asked to rate their expectancy for an orthogonal task feature, the left versus right response (response expectancy group). If expectancy about conflict specifically influences our ability to adapt to conflict, then we should observe differences in the PCSE between groups. If the PCSE operates independently of expectancy about conflict specifically, then there should be no difference between expectancy groups.

Method

Participants

One-hundred and thirty-eight first- and second- year psychology undergraduates in the University of Sydney completed the experiment for partial course credit. Of these, nine of the participants who did not reach the accuracy criterion of 80% were removed from analysis. Analyses were conducted on the remaining 129 participants (100 female, 29 male; mean age = 19.57), with 65 participants in the conflict expectancy group and 64 in the response expectancy group. This experiment was not preregistered.

The procedures from all three experiments were carried out in accordance with the regulations as outlined under the National Statement on Ethical Conduct in Human Research.

All protocols have been approved by the University of Sydney Human Ethics Committee (ethics protocol: 2023/HE000715).

Stimuli and Apparatus

All experiments were coded on MATLAB using the Psychtoolbox package [25]. The experiments were run on Dell computers running Windows 11, with wired keyboards to collect the participants' responses.

Task stimuli included a sequence of five left and/or right arrows that were presented in white print against a black screen (see Figure 2A for an example). For congruent trials, the distractor arrows (flankers) were the same as the target arrow on the centre, i.e. '<<<<<' for left response and '>>>>>' for right response. For incongruent trials, distractor arrows faced in the opposite direction to the target arrow, i.e. '>><>>' for left response & '<<><<' for right response.

Experimental Design

The test trials comprised 2 response types (left vs. right) × 2 conflict trial types (congruent vs. incongruent) × 2 rating trial types (no rating vs. rating). The sequence of congruent and incongruent trials was pseudorandomised according to a binomial distribution of run lengths (see Table 1), as typically conducted across Perruchet experiments, ensuring an equal number of congruent and incongruent trials overall and for each run length. To achieve this, runs of trials were concatenated, randomly choosing a run of congruent or incongruent to begin, then alternating regularly between runs of congruent and incongruent trials. Each time, the run was chosen without replacement from the pool of runs created according to the frequencies in Table 1. This yields a total of 128 runs, amounting to 254 trials. Two trials (one congruent and one incongruent; 256 trials in total) were then added so that the effect of the final run was fully measurable and the total randomised list was divisible into blocks of 64 trials. This randomisation was then repeated to yield a total of 512 trials. The sequence of left and right response trials was pseudorandomised in the same manner, allowing for an independent analysis based on runs of preceding left versus right trials. Runs of response were randomised with the constraint that these runs did not correlate with runs of congruence, based on a procedure first used by Barrett and Livesey [26]. The sequence of no-rating and rating trials was also randomised, with each trial type occurring a random 50% of the trials.

Table 1

Distribution of congruent and incongruent trials

	Incongruent							**Congruent**						**Total**
Run length	7	6	5	4	3	2	1	1	2	3	4	5	6	7	
Number of runs	1	1	2	4	8	16	32	32	16	8	4	2	1	1	128
Number of trials	7	6	10	16	24	32	32	32	32	24	16	10	6	7	254

Figure 5B outlines the sequence of events for each trial. In trials without expectancy ratings, the flanker task was preceded by a blank ITI of 3250 ms. In trials with expectancy ratings, the blank ITI lasted for 250 ms before the expectancy rating was shown for 3000 ms, during which the marker appeared at the

midpoint of the slider and the participant rated their expectancy by adjusting the slider with the ‘t’ and ‘h’ keys, where ‘t’ moved the marker to the left and ‘h’ moved it to the right.

Participants were randomly assigned to one of two conditions: those in the conflict expectancy group were asked to rate their expectancy of the level of conflict present in the upcoming trial, phrased in terms of easy (congruent) vs. hard (incongruent), while those in the response expectancy group were asked to predict the required response on the upcoming trial (left vs. right). Ratings for ‘easy’ were on the left end of the scale, and for ‘hard’ on the right end; ratings for ‘left’ response were on the left end, and for ‘right’ response on the right end. Following the expectancy rating or blank ITI, a fixation cross appeared for 800 ms, followed by the flanker display. Participants had a maximum of 1000 ms to complete the flanker trial and were given feedback informing them of their performance: ‘Correct’ for a correct answer within the time limit, ‘Wrong key!’ for an incorrect answer, ‘Responded too early!’ for a response before the 1000 ms window, and ‘Too slow!’ for a response beyond 1000 ms or a lack of response. The feedback lasted for 2000 ms.

Procedure

Prior to beginning the experiment, demographic data of age and gender were collected from participants, along with informed consent. They were then given brief verbal instructions, placing emphasis on the distribution of trials. That is, it was stressed that there would be an equal number of ‘easy’ and ‘hard’ trials, as well as left and right response trials, but that these were randomly determined. They were also explicitly informed that ‘easy’ referred to trials with all arrows pointing in the same direction (congruent), and ‘hard’ referred to trials with the central arrow pointing opposite to the flanking arrows (incongruent), ensuring that they were not misinterpreting the meanings of ‘easy’ and ‘hard’. These verbal instructions were reiterated in detailed written instructions that showed visual examples of the easy (congruent) and hard (incongruent) displays that the participants would need to complete. Participants then completed a practice block of 12 trials, before being given further instructions about the use of the expectancy rating scale. They then completed eight test blocks of 64 trials each, making up 512 trials.

At the end of each block, the text ‘END OF BLOCK’ with the block number appeared on screen, along with ‘press a key to continue’, where they were allowed to take a break until they were ready to continue. Once they pressed a key, ‘Get ready’ showed up on screen, and the next block began. When the experiment ended, ‘End of the experiment’ appeared on screen. Response time (RT) of the flanker task was used as the measure of performance for data analyses, alongside expectancy ratings, as a measure for conscious expectancy.

Data analyses

Responses made on the first trials of the block were not included in the analyses (1.56% of trials), as well as trials following an error (5.21% of trials). RTs that were beyond 3 SDs of the mean RT were flagged as

outliers and excluded from analysis (1.47% of trials). RT data from all remaining correct responses (87.4% of trials) were included in the analyses, which were run using RStudio [27].

We are most interested in observing the flanker effect, that is, the difference between mean incongruent RT and mean congruent RT, and expectancies as a function of conflict history (i.e. the preceding runs of congruent/incongruent trials). For our main analyses, we used repeated-measures analysis of variance (ANOVA), comparing the flanker effect and expectancy at each level of run length, and drawing linear trends from these means using linear contrasts.

Trials that were preceded by more than four consecutive trials (that is, had a run length of more than four) were pooled together, resulting in eight levels of run length. For analyses, in keeping with past studies on the Perruchet effect, these 8 levels were arranged with the longest run lengths at the extremes; from highest conflict history to lowest conflict history, they were ordered 4+ incongruent, 3 incongruent, 2 incongruent, 1 incongruent, 1 congruent, 2 congruent, 3 congruent, 4+ congruent (as shown on the x-axis in Figure 1). In doing so, linear contrasts provide a direct test for positive versus negative recency trends. For linear trend analyses, these were coded as runs values of [-4, -3, -2, -1, 1, 2, 3, 4]. A similar approach was used to analyse response history (runs of left and right responses) for the response expectancy condition, with 8 levels organised from 4+ left at one extreme to 4+ right at the other (see Supplementary Materials for further analyses based on runs of same versus different responses). Because the Perruchet dissociation is specifically designed to place emphasis on the presence of linear trends, for the within-subjects factor run length, we report only the outcome of linear trend contrasts (and interactions with these contrasts), however full ANOVA results including run length main effects can be found in Supplementary Materials.

To simplify reporting of results, we report analyses using the congruency effect—that is, the *difference* between RT on incongruent and congruent trials—as the dependent variable when reporting the PCSE. Full analyses separating congruent and incongruent trials as levels of an independent factor can be found in the see Supplementary Materials.

For all experiments, ratings were linearly transformed so that the maximum expectancy ratings for “easy” (i.e. congruent) was given score of +1 and maximum prediction of “hard” (i.e. incongruent) was given a -1. For left/right ratings, maximum expectancy ratings for “right” were given a score +1 and maximum expectancy ratings for “left” were given a score of -1. Hence, ratings less than 0 reflect greater expectancy of an incongruent trial (or a left trial), and ratings greater than 0 reflect greater expectancy for a congruent trial (or a right trial). These scores were used as the dependent variable for analyses on expectancies on all experiments.

Results

Flanker effect

Figure 3A depicts the flanker effect for each level of run length, for each trial type and group. We conducted a 2 × (2 × 8) mixed repeated-measures analysis of variance (ANOVA) on this RT difference, with rating group (conflict expectancy group vs. response expectancy group), rating (no rating in previous ITI vs. rating in previous ITI), and run length as factors. Figure 6A represents the average flanker effect as a function of run length. The main effect of rating group was not significant, $F(1, 127) = 0.63$, $p = .430$, $\eta_p^2 = .005$, indicating that the mean flanker effect did not differ between groups. The run length linear trend contrast was significant, $t(127) = 10.812$, $p < .001$, $\eta_p^2 = .48$, indicating the presence of a positive recency effect in which the flanker effect was smallest following longer runs of incongruent trials and largest following longer runs of congruent trials. This confirms the presence of the PCSE. A main effect of rating was also significant, $F(1, 127) = 19.62$, $p < .001$, $\eta_p^2 = .134$, indicating that making a rating reduced the flanker effect. However, the rating factor did not significantly interact with the linear trend across run length, $t(127) = -0.306$, $p = .760$, $\eta_p^2 < .001$, meaning there was no difference in the run length slopes (the PCSE) between no rating and rating conditions. Critically, there was also no difference in the run length linear trend (i.e. strength of the PCSE) as a function of rating groups, $t(127) = 1.027$, $p = .306$, $\eta_p^2 < .001$, and no three-way interaction effect between rating (rating vs no rating trials), rating group (conflict expectancy group vs response expectancy group) and the linear trend in run length, $t(127) = -0.851$, $p = .396$, $\eta_p^2 < .001$. The absence of interactions with run length linear trend suggests that the strength of the PCSE was not modulated by the qualities of the participant's explicit expectancies, nor whether they had just made such a rating.

Expectancy

Figure 3 depicts the average expectancy ratings as a function of conflict history for the conflict expectancy group (Figure 3B) and a function of response history for the response expectancy group (Figure 3C). Expectancy ratings for the response expectancy group exhibit a negative trend, which is consistent with past choice RT studies [20–22,26,28]; on the other hand, the trend is not as clear for the conflict expectancy group. Separate repeated measures ANOVAs for each ratings group showed that the conflict expectancy group did not show a significant linear trend over run length, , $t(64) = -1.084$, $p = .283$, $\eta_p^2 = .02$, but there was a significant downwards linear trend in the response expectancy group, corresponding to negative recency, $t(63) = -2.313$, $p = .024$, $\eta_p^2 = .08$, and consistent with what has been observed across Perruchet studies.

Discussion

There was a strong effect of conflict history in the flanker task, demonstrating a clear PCSE. However, we found no evidence that the type of expectancy rating participants were asked to make had any influence on the PCSE. Instead, positive recency was evident across both conditions. Jiménez and Méndez [14] speculated that the act of being asked to provide an explicit expectancy rating may generate explicit

predictions that work in opposition to the processes responsible for the PSCE. We did not find a weaker PSCE in the conflict expectancy group. This suggests that prompting participants to predict the level of conflict in the upcoming trial does not significantly impact their ability to adapt to changing levels of conflict.

It is worth acknowledging that we did not observe a reliable negative recency trend in the expectancy ratings for conflict. In comparison, the response expectancy condition displayed Perruchet's dissociation reliably, with evidence of both negative recency in expectancy and positive recency in RT trends across response history (see Supplementary Materials). The discrepant trend we observed in our congruency expectancy condition, then, suggests that there is something different about the way people generate expectancies regarding conflict. Regardless, one aspect of the results is clear: directing the participant's explicit focus towards conflict history was not sufficient to overcome the processes responsible for the PCSE. Experiments 2 and 3 sought to investigate factors that may be responsible for this pattern of results.

Experiment 2

The lack of an expectancy trend for conflict in Experiment 1 is surprising, given past research [16]. Experiment 2 thus sought to investigate whether this was a function of participants' motivation to generate expectancy ratings. To that end, we introduced a points system to incentivise participants to pay more attention to the sequence of congruency. Crucially, one condition earned and lost points based on the accuracy of the expectancy ratings (performance + expectancy points group). When their expectancy for conflict matched the congruency of the next trial, they earned points, scaling with the strength of their rating. However, if their expectancy did not match the congruence, they lost points, again scaling with the strength of their rating. Importantly, the average points accumulated from expectancy ratings amounts to zero as there is a 50% chance of getting the expectancy rating correct at any given point. To avoid this impeding the motivational purpose of introducing the points, participants also earned points for accurate performance in the flanker task, scaling with the speed of their responses. The control condition only earned points for their performance, but not their expectancy ratings (performance points group).

Experiment 1 necessitated a two-response version of the flanker task so that we could compare expectancy for two orthogonal but binary dimensions of the task (high versus low conflict trials compared to left versus right response trials). However, one notable limitation of the standard 2-AFC flanker task used in Experiment 1, which uses only left and right arrow stimuli, is that there is inevitably a high degree of featural overlap between the current and previous trial; they are likely to possess the same target and response, or the same distractors, or both. Demonstrations of the standard CSE shown with this version of the task have been criticised because sequential effects can be largely attributed to the extent of this overlap, if one assumes that partial overlap leads to interference and thus slower responding, while full or zero overlap leads to fast responding because of a lack of this interference [29,30]. Failure to account for the feature integration effect means that the PCSE in Experiment 1 may be

partly driven by first-order featural overlap effects. To address this, following the procedure suggested by Schmidt and Weissman [31], in Experiment 2 we used a more complex version of the flanker task which alternated regularly between two response sets, left-right arrows (as with the flanker task used in Experiment 1), and up-down arrows. This prevents the problem of having overlapping target and distractor features on consecutive trials which may aid in enhanced responding between consecutive trials of the same kind.

Method

Participants

The experiment was preregistered and can be accessed at https://aspredicted.org/fd8du8.pdf. The required sample size was calculated using G*power [32]. As we were aiming to obtain a reliable gambler's fallacy trend, we used the effect size of the gambler's fallacy trend obtained in Experiment 1's response expectancy condition, which was $\eta_p^2 = .08$, to compute for the required sample size. Using a power of .8 and = .05 and running a one-tailed test of difference in means from constant, we obtained a result of n=74. As we had two experimental conditions, we used this size for each condition, entailing a minimum of 148 participants. A total of 156 psychology undergraduates from the University of Sydney participated in the experiment for partial course credit and were randomly allocated to either condition. Six participants did not meet the accuracy cutoff of 80% and were subsequently removed from analyses. The remaining 150 participants' data were used for analysis (123 female, 27 male; mean age = 19.27).

Stimuli and Apparatus

Experiment 2 had two sets of stimuli. The first set used left and right facing arrow heads identical to Experiment 1. Stimuli of the second set were rotated 90º such that they faced up and down, arranged vertically. The left/right and up/down sets were presented in strict alternation (each occurring every other trial), with the orientation on the first trial chosen randomly for each participant. Apparatus used was identical to the first experiment.

Experimental Design

The test trials were made up of 2 conflict trial types (congruent vs. incongruent) × 2 rating trial types (no expectancy rating vs. expectancy rating) × 4 response types (left, right, up and down arrows). The sequence of congruent and incongruent trials was pseudorandomised using the same procedure as Experiment 1, as well as the randomisation for the no-rating and rating trials. The trial presentation was very similar to Experiment 1, with the following minor time adjustments: on trials with a blank ITI (no-rating trials), a blank screen was shown for 2750 ms, whereas on rating trials, the blank ITI was presented for 250 ms before the expectancy rating screen appeared, lasting for 2500 ms. The fixation

cross was then presented for 1000 ms, followed by the flanker display. The participant was given a maximum of 1000 ms to respond to the flanker task. Feedback was then presented for 1500 ms.

The participant was given the same response feedback as in Experiment 1, but with the addition of information about the points they earned in that trial. Both groups were awarded points for making an accurate speeded response to the flanker task, which was calculated by taking the trial RT rounded to the nearest tenth of a second and subtracting it from 12; for example, a correct response with an RT of 685 would correspond to 5 points gained, while a correct response with an RT of 423 would earn them 8 points. This was determined by considering the average RT of 506 ms in the flanker task from Experiment 1, which meant that on average, participants will earn about 7 points from the flanker task.

To accommodate the points system in the performance + expectancy points condition, we also adjusted the scale of the expectancy ratings. Instead of a continuous scale with points between 0–100, Experiment 2 used a discrete scale with levels 1–5 on each side of the scale, with the marker's starting location (and the midpoint) being 0. That is, pressing the 't' key would shift the strength of their expectancy toward 'easy' by 1, to a maximum of 5, and pressing the 'h' key would shift their expectancy towards 'hard' by 1, also to a maximum of 5. In the performance + expectancy points condition, this directly corresponded to the points participants could earn or lose by making an expectancy rating: making an expectancy rating of '5' for easy before a congruent trial would award the participant an extra five points, on top of the points they would earn from the speeded response; on the other hand, making a rating of '2' for easy before an incongruent trial would deduct two points from their total amount of points. The maximum amount of points they could lose from making an incorrect expectancy rating was still lower than the average amount they would earn from accurately completing the flanker task. This ensured participants did not view the points from rating as detrimental, which may discourage them from engaging with the expectancy ratings. In principle, all participants should see their points tally gradually aggregating across trials.

Feedback was thus given accordingly: for the performance points group, participants received two lines of feedback on all trials. In trials where the participant responded correctly, they were shown "Correct" and "Points from task: " with the points they earned from the speeded response; for incorrect responses, they were shown the relevant feedback message, along with "Points from task: 0". This was also true for no-ratings trials in the performance + expectancy points condition. However, on rating trials, they were shown additionally the points they earned from making a rating, indicated by the string "Points from expectancy rating:". If they lost points instead, the number contained a negative sign. At the end of each block, participants were also informed of their current point total, with the string 'Current points total:'.

Procedure

The procedure was similar to that of Experiment 1, with the following exceptions: they were informed about the four response options and regarding the points system, where making faster responses in the flanker task would earn them more points. For the performance + expectancy points group, they were

additionally told that they would earn or lose points depending on the strength and accuracy of their expectancy ratings.

Data analyses

The RT data from correct responses and expectancy ratings were analysed in the same way as with Experiment 1. Responses made on the first trials of the block (1.56% of trials), trials following an error (6.65% of trials), and RTs that were beyond 3 SDs of the mean RT (1.24% of trials) were removed from analyses. The remaining trials with correct responses () were included in the analyses.

Results

Flanker effect

Figure 4A represents the magnitude of the flanker effect (the difference between incongruent trial RT and congruent trial RT) as a function of run length. A 2 × (2 × 8) ANOVA was conducted with the flanker effect as the dependent variable and group (performance points group vs. performance + expectancy points group), rating, and run length as factors. As in Experiment 1, there was a significant linear trend in run length, indicating the presence of the PCSE, $t(145) = 10.497$, $p < .001$, $\eta_p^2 = .43$. In addition, there was a significant main effect of group, where the average flanker effect was smaller for the performance points group than for the performance + expectancy points group, $F(1, 145) = 10.93$, $p = .001$, $\eta_p^2 = .070$. Although the flanker effect appeared to be marginally smaller on rating trials than no-rating trials, the main effect of rating did not reach statistical significance, $F(1, 145) = 3.82$, $p = .052$, $\eta_p^2 = .026$. However, there was an interaction effect between rating and group, where the difference in flanker effect between no-rating and rating trials was larger in the performance points group than the performance + expectancy points group, $F(1, 145) = 6.15$, $p = .014$, $\eta_p^2 = .041$. Unlike Experiment 1, we found a significant interaction between rating and linear trend in run length, indicating that the PCSE was stronger for no-rating trials compared to rating trials, $t(145) = -2.522$, $p = .013$, $\eta_p^2 = .04$ (though the linear trend was significantly positive for both rating and no-rating conditions, see Supplementary Materials). The PCSE was not significantly different between groups when averaged over rating conditions, $t(145) = .458$, $p = .648$, $\eta_p^2 < .001$, and the three-way interaction between rating, group, and run length linear trend was not significant, $t(145) = -.088$, $p = .930$, $\eta_p^2 < .001$.

Expectancy

Figure 4B demonstrates the average expectancy ratings for each run length. The graph displays a slightly upwards trend in the performance points group, which is an unexpected pattern of results. A 2 × (8) ANOVA was run with group and run length as factors. There was a significant main effect of group, $F(1,$

148) = 5.37, p = .022, η_p^2 = .035, with average expectancy ratings for the performance points group leaning closer to 'hard' (i.e. expectancy of high conflict), and that of the performance + expectancy points group leaning closer to 'easy' (i.e. expectancy of low conflict). Averaging across groups, there was no significant linear trend across run length, t(148) = 0.869, p = .386, η_p^2 = 0.03, and even though the two groups seem to yield different patterns in expectancy, these trends were also not significantly different, t(145) = 1.604, p = .111, η_p^2 = 0.02. Examining each group individually, there was no linear trend in expectancy ratings for the performance + expectancy points group, t(74) = 0.625, p = .534, η_p^2 = 0.005. Although the seemingly *positive* trend (i.e. positive recency consistent with the hot-hand fallacy) for the performance points group is surprising, this trend did not reach statistical significance, t(74) = -1.942, p = .056, η_p^2 = 0.05.

Discussion

In Experiment 2, we again demonstrated a robust PCSE, but failed to observe clear expectancy rating trends in either group. It is possible that the target response (left-right response expectancy in Experiment 1) is simply more salient than conflict, which is not directly task-relevant. For example, in conflict tasks, participants are not typically asked to overtly indicate whether a trial was congruent or incongruent. Thus, one possibility is that people do not explicitly remember conflict history as well as they do response history. Experiment 3 was therefore designed to test the question of whether the ability to explicitly recall conflict history is a modulating factor in the expectancy trends people generate regarding congruency.

Experiment 3

Experiment 3 sought to address whether participants simply fail to properly retain information about the level of conflict experienced on recent trials in a way that can be used to make explicit predictions. Chen and Wyble [33] noted participants' difficulty recalling basic task-relevant attributes in attention-demanding speeded tasks, referring to the apparent deficit as *attribute amnesia*. We hypothesised that a failure to encode conflict (analogous to attribute amnesia) might explain why expectancies for conflict are not generated in a way that is sensitive to recent trial history. It is possible that expectancies for other outcomes (such as left and right responses) are more sensitive to recent trial history because they are more salient and goal-relevant. Experiment 3 thus tested for an analogue of attribute amnesia by occasionally probing participants about the level of conflict they had just encountered.

This addition serves two purposes: to observe whether participants remember what they previously experienced, and if so, to see if their ability to recall past conflict level influences their pattern of expectancy ratings as a function of conflict history. We hypothesised that those who have higher recall of conflict on recent trials would exhibit a stronger gambler's fallacy trend in their expectancy ratings than those with poorer recall of conflict. Furthermore, we were interested in whether a stronger ability to

report recently encountered conflict was also related to the magnitude of the PCSE in flanker task performance.

Method

Participants

The preregistration for this experiment can be found at https://aspredicted.org/we8w9v.pdf. We aimed for the sample size used in the performance + expectancy points condition of Experiment 2 to yield comparable results. A total of 79 psychology undergraduates from the University of Sydney participated in the experiment for partial course credit. Two did not meet the accuracy cutoff of 80% and were removed from analyses. One participant was excluded for performing well below chance in the attention check trials (accuracy = 5%) suggesting they had not understood the task. The remaining 76 participants' data were used for analysis (60 female, 16 male; mean age = 19.05).

Stimuli, Apparatus, and Experimental Design

All stimuli, apparatus and design elements were identical to the performance + expectancy points condition in Experiment 2, apart from the additional attention check questions introduced. A randomly selected 64 trials ($1/8^{th}$ of all trials) were designated as attention check trials, on which the participant reported the level of conflict on the previous trial in response to a probe question presented during the ITI. The rest of the trials were divided into rating and no-rating trials.

The trial presentation for no-rating trials and rating trials were identical to the performance + expectancy points condition in Experiment 2. On attention check trials, the question "What type of trial was the previous trial?" was presented, followed by two further lines of text prompting "Press 'w' for Easy" and "Press 's' for Hard". Responding correctly resulted in the chosen option turning green, while responding incorrectly resulted in the option to turning red. Regardless of whether they made a response, the attention test prompt disappeared after 3000 ms and was replaced by the fixation cross, followed by the flanker display. Unlike the expectancy and flanker responses, the attention test itself was not incentivised with points.

To accommodate the addition of the attention check questions, we modified the key assignments for Experiment 3. The expectancy ratings were now adjusted using the 'a' and 'd' keys, where pressing 'a' moved the marker to the left of the scale, incrementing expectancy for 'easy' (congruent) by 1, and pressing 'd' moved the marker to the right, increasing expectancy for 'hard' (incongruent) by 1. These key mappings directly reflect the orientation in which they observed the prompts on screen for each of their relevant tasks (a horizontal expectancy rating scale, and vertically arranged options for the attention test), and prevent the need for hand movements between trials, as well as ensure that there are no response priming influences in responding to different tasks.

Procedure

Procedures used were identical to the performance + expectancy condition of Experiment 2, except for the new key assignments and the addition of the attention check questions, as described above.

Data analyses

The RT data and expectancy ratings were analysed in the same way as with the previous experiments, with the exception that there was only one group (and thus no *a priori* between-subjects factor) in Experiment 3.

Accuracy in the attention test was calculated as proportion of correct responses. While we had preregistered an analysis looking into the subset of people whose attention accuracy was above 80%, post-hoc examination of the data revealed two clear clusters of participants, with one cluster performing at about chance and the other, above chance (> 64% accuracy). Using either threshold leads to the same pattern of results, so we report analyses using these clusters as a factor and report the preregistered analyses in Supplementary Materials. Additionally, in order to investigate the relationship between attention check accuracy and expectancy ratings, we calculated participants' average expectancy ratings slopes across the eight run length levels. This was done by running a linear regression using run length as a factor against expectancy ratings and taking the slope of the equation for each participant.

Results

Flanker effect and expectancy

Figure 5A outlines the average flanker effect for each run length. A (2 × 8) ANOVA was conducted with the flanker effect as the dependent variable and rating and run length as factors. We again found a significant linear trend in flanker effect as a function of run length, indicating the PCSE, $t(74) = 7.133$, $p < .001$, $\eta_p^2 = .41$. There was no main effect of rating, $F(1, 74) = 0.46$, $p = .499$, $\eta_p^2 = .006$, and although the linear trend appears to be slightly stronger in the no-rating condition, there was no significant interaction between rating condition and the linear trend in run length, $t(74) = -1.874$, $p = .065$, $\eta_p^2 = .05$ (and both conditions yielded a significant linear trend when analysed separately, see Supplementary Materials).

Figure 5B depicts the average trend of expectancies across run length. We ran a repeated measures ANOVA on expectancy with run length as a factor. Consistent with Experiment 2, the linear trend was non-significant, $t(74) = 0.967$, $p = .337$, $\eta_p^2 = .01$. While the pattern of expectancies (i.e. lack of reliable linear trend) was similar to the performance + expectancy points condition in Experiment 2, the mean expectancy in this experiment leaned slightly in favour of expectancies for higher conflict (mean ratings below zero).

Attention check

Mean accuracy on the attention check questions was 0.777 (SD = 0.159 , range = 0.371–0.984). The distribution of these scores can be seen in Figure 6A. A Pearson's correlation test showed a significant negative correlation between attention accuracy scores and expectancy rating slopes, such that those who performed more accurately in the attention tests had more negative expectancy slopes, as plotted in Figure 6A, r = -.321, p < .001, indicating. The scatterplot also reveals two general clusters: those with attention check accuracies in the 40–60% range, and those with accuracies above >70%. According to a simple binomial test, an accuracy greater than 64% would be considered significant for an individual participant, meaning that these clusters can be seen as those performing at chance and above chance respectively. As such, we decided to run a post-hoc exploratory analysis where we split participants into those two attention groups (N(chance group) = 15; N(above chance group) = 61) and reran the prior statistical analyses with attention group as a factor, to see whether different levels of explicit conflict recall will yield different trends in expectancy ratings.

Figure 6B displays mean expectancy ratings by run length, divided by attention groups. We reran the ANOVA on expectancy ratings with attention group (chance group vs. above chance group) as an additional between-subjects factor. Attention group was not a significant factor on mean expectancy, $F(1, 74) = 1.81$, $p = .182$, $\eta_p^2 = .024$, but it interacted with the linear trend across run length, $t(74) = -2.95$, $p = .004$, $\eta_p^2 = .11$. Contrast analyses on each attention group's trends revealed that the chance group exhibited a significant positive recency trend, $t(14) = 2.28$, $p = .039$, $\eta_p^2 = .27$; by contrast, the above chance group displayed a significant negative recency trend, $t(60) = -2.21$, $p = .031$, $\eta_p^2 = .08$.

Given the clear differences in expectancy trends between attention groups, we tested whether there were any differences in the flanker effect slopes between these groups. Figure 6C outlines the flanker effect by run length, split by attention groups. We reran the ANOVA on the flanker effect with attention as an additional between-subjects factor. Attention group did not significantly affect the flanker effect on average, $F(1, 73) = 1.82$, $p = .182$, $\eta_p^2 = .024$, nor did it interact with rating, $F(1, 73) = 0.35$, $p = .558$, $\eta_p^2 = .005$. The linear trend in run length did not interact with group, $t(73) = -.437$, $p = .663$, $\eta_p^2 < .001$, suggesting no differences between the groups in the magnitude of the PCSE. Similarly, there was no difference in PCSE between rating and no rating trials, $t(73) = -1.004$, $p = .319$, $\eta_p^2 = .01$, nor was there a three-way interaction effect between attention group, rating, and linear trend run length, $t(73) = -.721$, $p = .473$, $\eta_p^2 < .001$. Analyses ran on each attention group showed that the PCSE was significantly linear for both groups; above chance group: $t(60) = 5.972$, $p < .001$, $\eta_p^2 = .37$; chance group: $t(13) = 4.370$, $p < .001$, $\eta_p^2 = .60$.

Discussion

Consistent with previous experiments, Experiment 3 found a significant PCSE in the flanker effect. Expectancy ratings again yielded a null trend overall, highlighting that, on average, people do not generate clear patterns of expectancies with respect to conflict history. Our attention check manipulation, however, reveals an important qualification to this finding. When participants were divided into those who were able to reliably report the level of conflict experienced on the previous trial and those who were not, differences in the pattern of expectancies became evident. Those performing above chance demonstrated clear negative recency, consistent with the pattern of expectancy typically observed in the Perruchet effect. Interestingly, those who could not reliably report the conflict they had just experienced produced a *positive* recency trend (i.e. consistent with the hot hand fallacy) rather than no trend at all. Our at-chance group consists of only 15 participants and thus our ability to draw confident conclusions about the behaviour of this group is limited. However, the pattern suggests that the null results obtained in the previous experiments may well also be an average of attentive participants demonstrating one trend and inattentive participants demonstrating another.

Despite the evidence that something akin to attribute amnesia significantly affects conflict level expectancy trends, the same cannot be said for the PCSE. Attention check accuracy did not have any reliable influence on the magnitude of the PCSE in the flanker task itself. In fact, among the group of participants who passed the attention check, we finally found more convincing evidence of Perruchet's dissociation in full. This adds to the evidence across all three experiments that probing for conscious expectancies about upcoming conflict does not reduce the effect of conflict history on cognitive control processes. It suggests that the PCSE, at least in the context of the flanker task, is largely independent of conscious expectancy. This has important implications for theories of adaptive control, particularly with respect to the (potentially limited) influence of conscious deliberate processes imposed intentionally by the individual.

General Discussion

This study set out to investigate the influence of conscious expectancy of conflict on adaptive control in the flanker task, making use of a dissociation logic that can separate trends in expectancy from trends in behaviour [16]. Here, we summarise the results of our three experiments in relation to the aims of group manipulations of conscious expectancies and its effect on adaptive control in the flanker task, and the production of expectancy trends that consistently show evidence of negative or positive recency.

In Experiment 1, we found no evidence that motivating the formation of conscious expectancies about conflict specifically had any impact on the magnitude of the PCSE. In Experiment 2, we found no evidence that motivating participants to produce valid and accurate expectancies affected the magnitude of the PCSE. Experiment 3 provided even further motivation to monitor the sequence of conflict experienced, and demonstrated evidence that participants vary in their ability to do so. Yet still, participants reliably produce the PCSE on the flanker task regardless of whether they could or could not reliably report the previous trial's conflict, despite the two groups showing significant differences in

pattern of expectancies. Overall, these experiments present consistent evidence that the PCSE is not influenced by the manner in which participants form conscious expectancies about impending conflict.

In measuring expectancy for conflict in the Stroop task [14,23], inconsistent patterns of expectancy have been found. Our results seem to suggest consistent expectancy trends in the flanker task are also not easy to find. While mean expectancy in each experiment exhibited weak trends towards negative recency, the result was not significant in Experiment 1 and 2, and only statistically robust in Experiment 3 after removing individuals who did not perform above chance in attention checks that probed memory for conflict on the previous trial. This is despite numerous observations of a negative recency pattern consistent with the gambler's fallacy in other choice response paradigms [16]. We hypothesised that one of the reasons why expectancy trends for conflict are inconsistent across individuals, studies and procedures is that consciously monitoring levels of conflict is inherently more difficult than monitoring other salient and discrete trial features like the required response. Consistent with this hypothesis, in Experiment 3, when probed about the conflict that was just observed, some participants did not perform above chance, and accuracy on these attention check questions predicted the presence of a negative recency trend; it appears only participants who remember conflict can use it to form explicit expectancies about future events that are in line with the gambler's fallacy.

Unlike the expectancy trends, however, whatever mechanism is responsible for adaptive control in the flanker task is not limited by participants' ability to remember and use prior conflict to generate explicit predictions. The PCSE was present in all conditions and largely unaffected by manipulations of participants' explicit expectancies. There are several types of mechanism that may be responsible for such an effect. One possibility is that adaptive control is driven by predictive mechanisms that operate independently of what the participant explicitly reports. For instance, associations between predictive cues in the task (e.g., the experimental context, the fixation cross) may incrementally strengthen and weaken as a consequence of completing several congruent or incongruent trials in a row, in effect automatically producing adaptive control adjustments that prepare for more of the same [34]. Such an explanation would be consistent with the theoretical motivation for Perruchet's original design [15], which dissociated expectancy from conditioned responding in an eyeblink conditioning paradigm. Just as Perruchet concluded that conditioning strength was determined by automatic learning and prediction mechanisms and not the individual's conscious expectation of the unconditioned stimulus, here it may be the case that learning guided control occurs automatically. Although it is usually aligned with conscious expectancies, there will be rare instances in which the two processes produce different predictions, as is the case in Perruchet's dissociation.

There are, however, other theories of adaptive control that do not conceive of sequential effects as a prediction-driven phenomenon at all [30,35,36]. While the simplest explanation in terms of S-R feature binding is ruled out by the use of alternating response sets in Experiments 2 and 3, it still may be the case that the PCSE is driven by other types of adaptive mechanisms. One possibility is that the sequential effects are a form of simple recency-based adjustment, with control setting adjusted upwards or downwards depending on the level of conflict encountered, as posited by the conflict-monitoring

theory [37]. Perruchet's dissociation has prompted novel research into the underlying mechanisms responsible for trends in behaviour in several conditioning and motor priming paradigms, with evidence provided for both associative and non-associative processes [22,26,38–41]. Future research should explore these issues with a specific focus on adaptive control and the PCSE in conflict tasks.

Conclusions

In summary, we obtained reliable and robust PCSEs on the flanker effect across our three experiments. They remained robust regardless of the feature for which they were generating expectancies (Experiment 1), regardless of the whether accurate predictions were incentivised (Experiment 2), and regardless of whether they made an expectancy rating immediately before the flanker trial. Significant recency trends for conflict expectancy were surprisingly hard to find, and this issue was not simply resolved by incentivising the expectancy rating task, suggesting that people do not develop expectancies for conflict the same way they do with stimulus/response outcomes. Only by directly investigating people's ability to recall conflict history did we finally obtain a negative recency trend, which was specific to those who were attentive of conflict history. Despite this major difference, expectancy trends did not have any influence on the PCSE of the flanker task. These series of experiments demonstrate that PCSEs are not driven by conscious expectancy for conflict. They lay a groundwork for further research involving the PCSE that manipulates different aspects of adaptive control, and potentially discriminating between alternative theories of cognitive control.

Declarations

Competing interests statement

The authors declare no competing interests.

Author contributions statement

E.C. and E.L. came up with the earlier versions of the study, and M.H. and H.D. contributed in conceiving the later versions. The programming of the experiment and the analyses scripts was done by E.C. and E.L. Data collection was done by E.C. and M.H., and all authors contributed in data analyses. E.C. wrote the first draft of the manuscript. All authors contributed in the editing of the draft, and approved the final draft before submission.

Funding statement

No funding was received for this research.

Author Contribution

E.C. and E.L. came up with the earlier versions of the study, and M.H. and H.D. contributed in conceiving the later versions. The programming of the experiment and the analyses scripts was done by E.C. and E.L. Data collection was done by E.C. and M.H., and all authors contributed in data analyses. E.C. wrote the first draft of the manuscript. All authors contributed in the editing of the draft, and approved the final draft before submission.

Acknowledgement

We would like to acknowledge Dr. Daniel Pearson, who offered valuable contributions into the interpretation and design of the experiments.

Data Availability

All data obtained and analysed in the experiment, along with R scripts for the data analyses, can be found in the OSF repository at the following link: https://osf.io/uwxgp/overview?view_only=50a108fded194236bcc6c4a800002e7a

References

1. Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S. & Cohen, J. D. Conflict monitoring and cognitive control. *Psychol. Rev.* **108**, 624–652 (2001).
2. Miller, E. K. & Cohen, J. D. An Integrative Theory of Prefrontal Cortex Function. *Annu. Rev. Neurosci.* **24**, 167–202 (2001).
3. Stroop, J. R. Studies of interference in serial verbal reactions. *J. Exp. Psychol.* **18**, 643–662 (1935).
4. Eriksen, B. A. & Eriksen, C. W. Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* **16**, 143–149 (1974).
5. Egner, T. Creatures of habit (and control): a multi-level learning perspective on the modulation of congruency effects. *Front Psychol* **5**, (2014).
6. Bugg, J. M. & Smallwood, A. The next trial will be conflicting! Effects of explicit congruency pre-cues on cognitive control. *Psychol. Res.* **80**, 16–33 (2016).
7. Gratton, G., Coles, M. G. H. & Donchin, E. Optimizing the use of information: Strategic control of activation of responses. *J. Exp. Psychol. Gen.* **121**, 480–506 (1992).
8. Hommel, B. & Colzato, L. Visual attention and the temporal dynamics of feature integration. *Visual Cognition*. **11**, 483–521 (2004).
9. Mordkoff, J. T. & Observation Three reasons to avoid having half of the trials be congruent in a four-alternative forced-choice experiment on sequential modulation. *Psychon Bull. Rev.* **19**, 750–757

(2012).

10. Schmidt, J. R. & De Houwer, J. Now you see it, now you don't: Controlling for contingencies and stimulus repetitions eliminates the Gratton effect. *Acta. Psychol.* **138**, 176–186 (2011).
11. Braver, T. S. The variable nature of cognitive control: a dual mechanisms framework. *Trends Cogn. Sci.* **16**, 106–113 (2012).
12. Shenhav, A., Botvinick, M. M. & Cohen, J. D. The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function. *Neuron* **79**, 217–240 (2013).
13. Jiménez, L. & Méndez, A. It is not what you expect: Dissociating conflict adaptation from expectancies in a Stroop task. *J. Exp. Psychol. Hum. Percept. Perform.* **39**, 271–284 (2013).
14. Jiménez, L. & Méndez, A. Even with time, conflict adaptation is not made of expectancies. *Front Psychol* **5**, (2014).
15. Perruchet, P. A pitfall for the expectancy theory of human eyelid conditioning. *Pav J. Biol. Sci.* **20**, 163–170 (1985).
16. Perruchet, P. Dissociating conscious expectancies from automatic link formation in associative learning: A review on the so-called Perruchet effect. *J. Experimental Psychology: Anim. Learn. Cognition*. **41**, 105–127 (2015).
17. Cowan, J. L. The Gambler's Fallacy. *Philos. Phenomenol. Res.* **30**, 238 (1969).
18. Gratton, G., Coles, M. G. H. & Donchin, E. Optimizing the use of information: Strategic control of activation of responses. *J. Exp. Psychol. Gen.* **121**, 480–506 (1992).
19. Braver, T. S. The variable nature of cognitive control: a dual mechanisms framework. *Trends Cogn. Sci.* **16**, 106–113 (2012).
20. Livesey, E. J. & Costa, D. S. J. Automaticity and Conscious Control in Single and Choice Response Time Versions of the Perruchet Effect. *Q. J. Experimental Psychol.* **67**, 646–664 (2014).
21. Lee, C., Lem, V. A., Harris, J. A. & Livesey, E. J. Testing the limits of the Perruchet effect in choice response time tasks. *J. Experimental Psychology: Anim. Learn. Cognition*. **41**, 385–394 (2015).
22. Destrebecqz, A., Vande Velde, M., San Anton, E., Cleeremans, A. & Bertels, J. Saving the Perruchet effect: A role for the strength of the association in associative learning. *Q. J. Experimental Psychol.* **72**, 1379–1386 (2019).
23. Duthoo, W., Wühr, P. & Notebaert, W. The hot-hand fallacy in cognitive control: Repetition expectancy modulates the congruency sequence effect. *Psychon Bull. Rev.* **20**, 798–805 (2013).
24. Gilovich, T., Vallone, R. & Tversky, A. The hot hand in basketball: On the misperception of random sequences. *Cogn. Psychol.* **17**, 295–314 (1985).
25. Kleiner, M., Brainard, D. & Pelli, D. What's new in Psychtoolbox-3? in Perception - ECVP Abstract Supplement. European Conference on Visual Perception (ECVP-2007), August 27–31, Arezzo, Italy (2007). 10.1177/03010066070360S101
26. Barrett, L. C. & Livesey, E. J. Dissociations between expectancy and performance in simple and two-choice reaction-time tasks: A test of associative and nonassociative explanations. *J. Experimental*

Psychology: Learn. Memory Cognition. **36**, 864–877 (2010).

27. RStudio Team. *RStudio: Integrated Development for R* (RStudio, PBC, 2020).
28. Destrebecqz, A. et al. The influence of temporal factors on automatic priming and conscious expectancy in a simple reaction time task. *Q. J. Experimental Psychol.* **63**, 291–309 (2010).
29. Mayr, U., Awh, E. & Laurey, P. Conflict adaptation effects in the absence of executive control. *Nat. Neurosci.* **6**, 450–452 (2003).
30. Hommel, B., Proctor, R. W. & Vu, K. P. L. A feature-integration account of sequential effects in the Simon task. *Psychol. Res.* **68**, 1–17 (2004).
31. Schmidt, J. R. & Weissman, D. H. Congruency Sequence Effects without Feature Integration or Contingency Learning Confounds. *PLoS ONE*. **9**, e102337 (2014).
32. Faul, F., Erdfelder, E., Lang, A. G. & Buchner, A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods*. **39**, 175–191 (2007).
33. Chen, H. & Wyble, B. Amnesia for Object Attributes: Failure to Report Attended Information That Had Just Reached Conscious Awareness. *Psychol. Sci.* **26**, 203–210 (2015).
34. Abrahamse, E., Braem, S., Notebaert, W. & Verguts, T. Grounding cognitive control in associative learning. *Psychol. Bull.* **142**, 693–728 (2016).
35. Schmidt, J. R. & De Houwer, J. Now you see it, now you don't: Controlling for contingencies and stimulus repetitions eliminates the Gratton effect. *Acta. Psychol.* **138**, 176–186 (2011).
36. Egner, T. Creatures of habit (and control): a multi-level learning perspective on the modulation of congruency effects. *Front Psychol* **5**, (2014).
37. Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S. & Cohen, J. D. Conflict monitoring and cognitive control. *Psychol. Rev.* **108**, 624–652 (2001).
38. Lee, C., Lem, V. A., Moul, C., Harris, J. A. & Livesey, E. J. Associations or repetitions? Testing the basis of the Perruchet effect in voluntary response speed. *J. Experimental Psychology: Learn. Memory Cognition*. **44**, 1971–1985 (2018).
39. Tran, D. M. D., Harris, J. A., Harris, I. M. & Livesey, E. J. Motor Conflict: Revealing Involuntary Conditioned Motor Preparation Using Transcranial Magnetic Stimulation. *Cereb. Cortex*. **30**, 2478–2488 (2020).
40. Weidemann, G. & Lovibond, P. F. The role of US recency in the Perruchet effect in eyeblink conditioning. *Biol. Psychol.* **119**, 1–10 (2016).
41. Weidemann, G., McAndrew, A., Livesey, E. J. & McLaren, I. P. L. Evidence for multiple processes contributing to the Perruchet effect: Response priming and associative learning. *J. Experimental Psychology: Anim. Learn. Cognition*. **42**, 366–379 (2016).

Figures

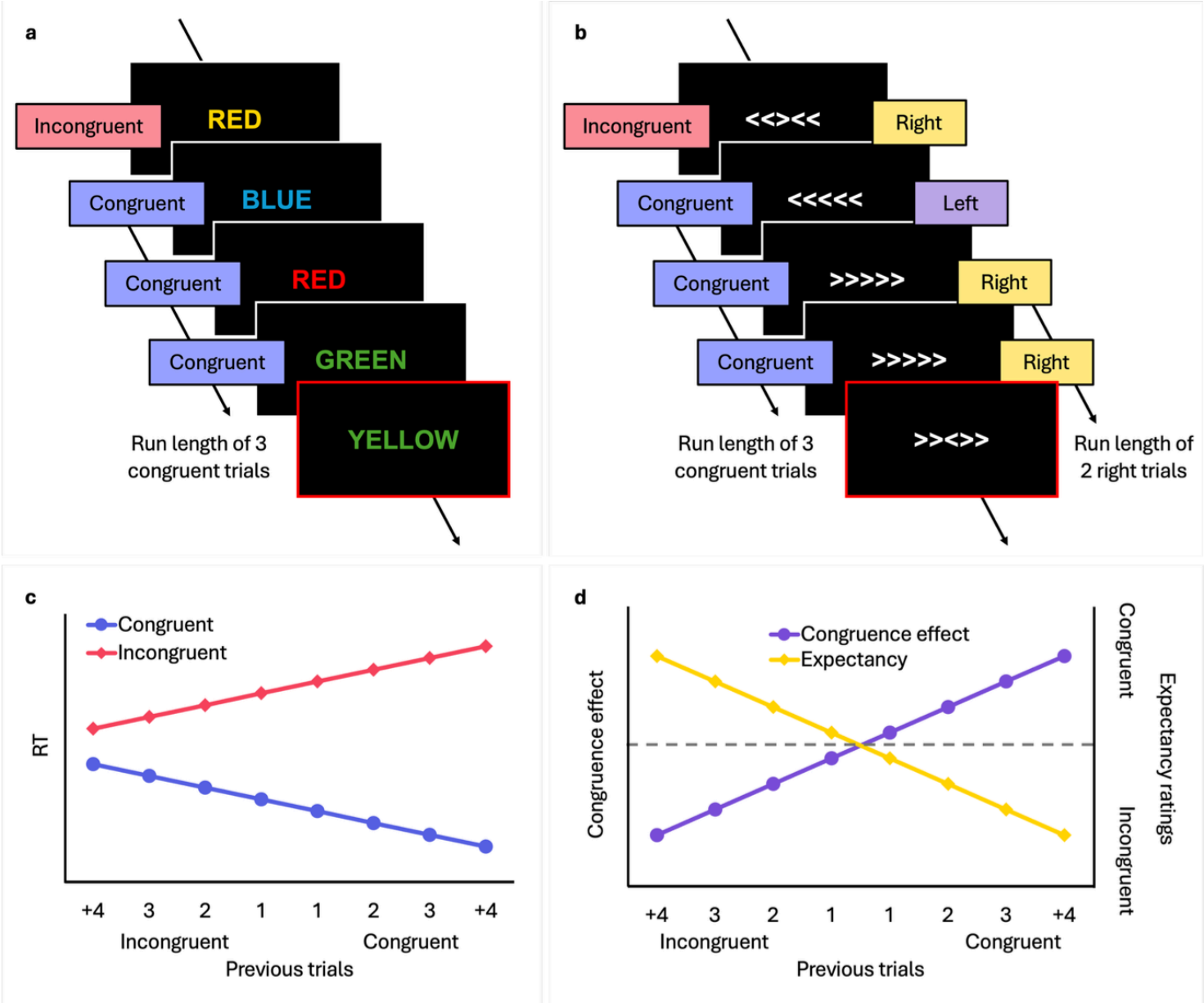


Figure 1

Schematics of the progressive congruency sequence effect (PCSE) and its dissociation against expectancy. **a)** An example sequence of trials in the Stroop task. The *n* trial (denoted by the red border) is characterised as having a run length of 3 congruent trials, as it is preceded by three congruent trials. **b)** An example sequence of trials in the flanker task used in Experiment 1. The *n* trial (denoted by the red border) is characterised as having a run length of 3 congruent trials, as it is preceded by three congruent trials, and having a run length of 2 right trials, as it is preceded by two trials requiring a right response. **c)** An idealised graph of response time as a function of the length and type (high- vs low-conflict) of preceding run of trials, displaying the *progressive congruency sequence effect* (PCSE). **d)** An idealised graph of the predicted dissociation between the congruence effect, denoted by the difference in RT between incongruent and congruent trials of the same run length, and expectancy for conflict across the same run lengths.

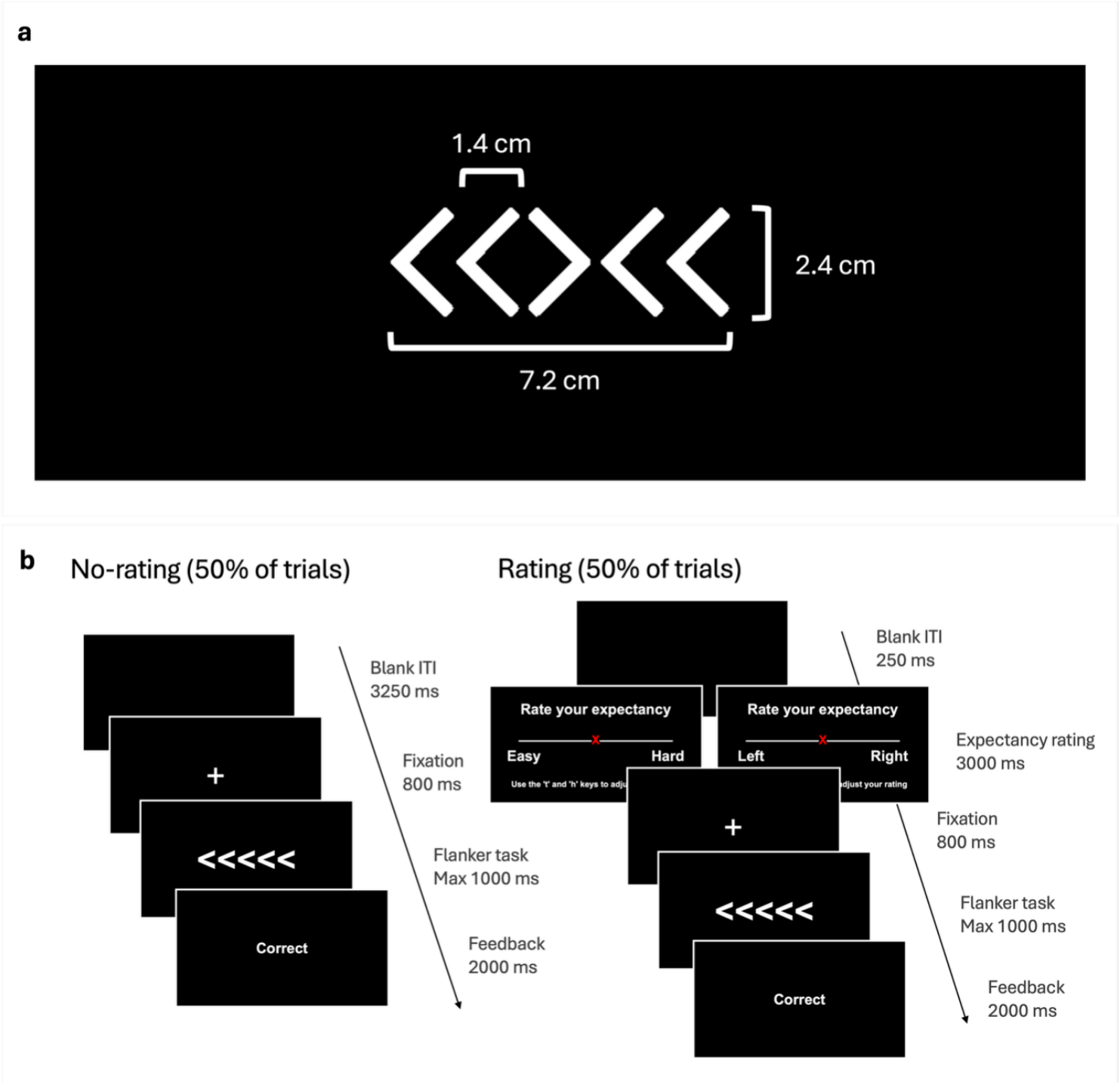


Figure 2

Example stimulus, and progression of a no-rating and rating trial. **a)** Figure depicts one stimulus (incongruent, right response trial) with the size specifications outlined, as presented on the monitors used in the lab. **b)**Figure depicts the sequence of screen presentations for the two trial types: no-rating and rating trials. In the rating trials, participants only saw the expectancy rating screen regarding the feature that they had been randomly assigned to, which was either upcoming trial conflict or upcoming trial response.

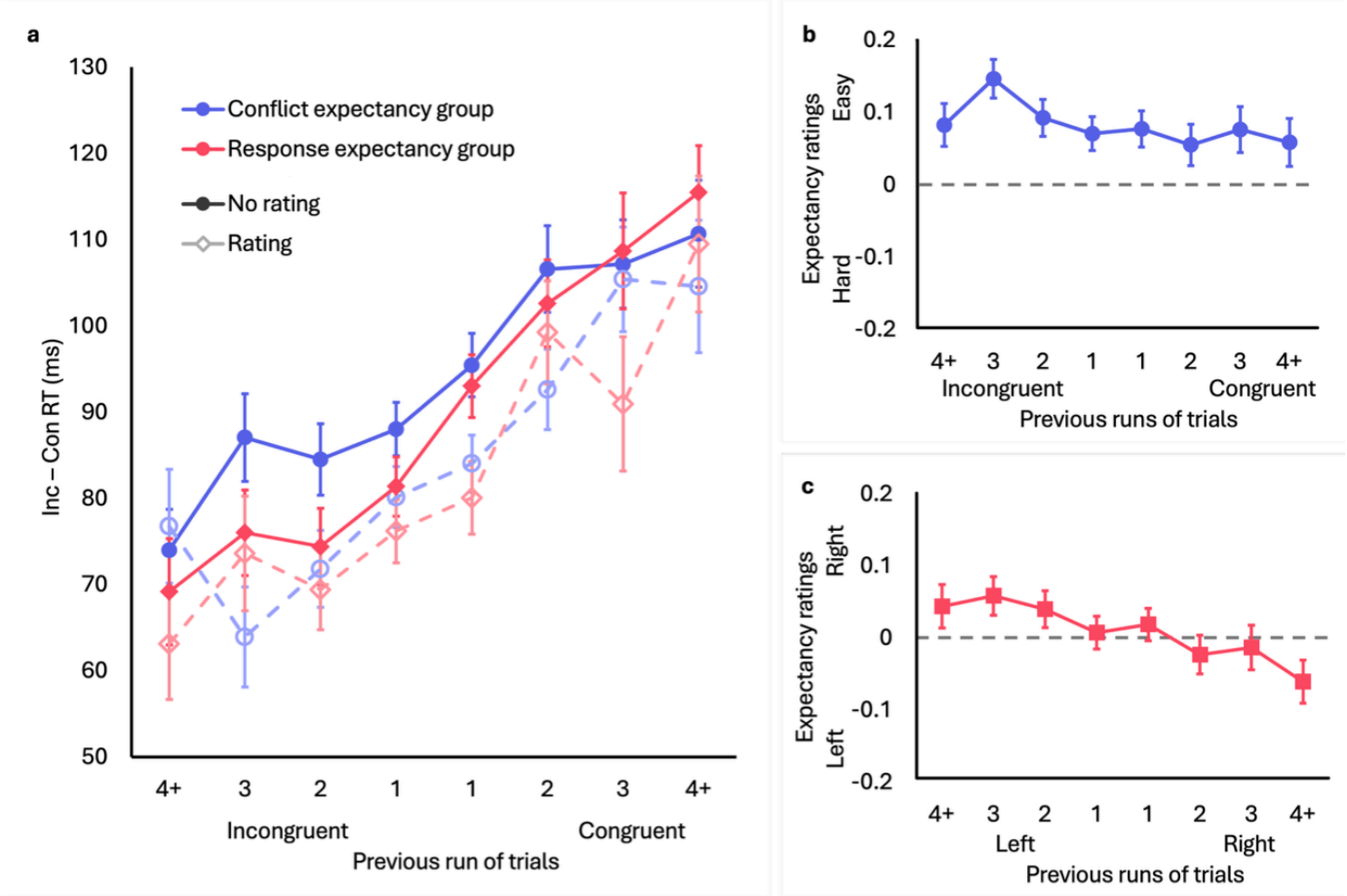


Figure 3

Flanker Effect RT and Expectancy Ratings by Previous Trial Runs. **A)** The flanker effect is denoted by the difference between incongruent trial RT and congruent trial RT. Dark solid lines denote trials without an expectancy rating, while light dashed lines are trials with expectancy ratings. The blue lines represent the conflict expectancy group, while the red lines represent the response expectancy group. Error bars correspond to one standard error of the mean (SEM). **B)** The blue line graph shows the expectancy ratings for the conflict expectancy group. The red line graph depicts the ratings for the response expectancy group. Standard error bars represent one SEM.

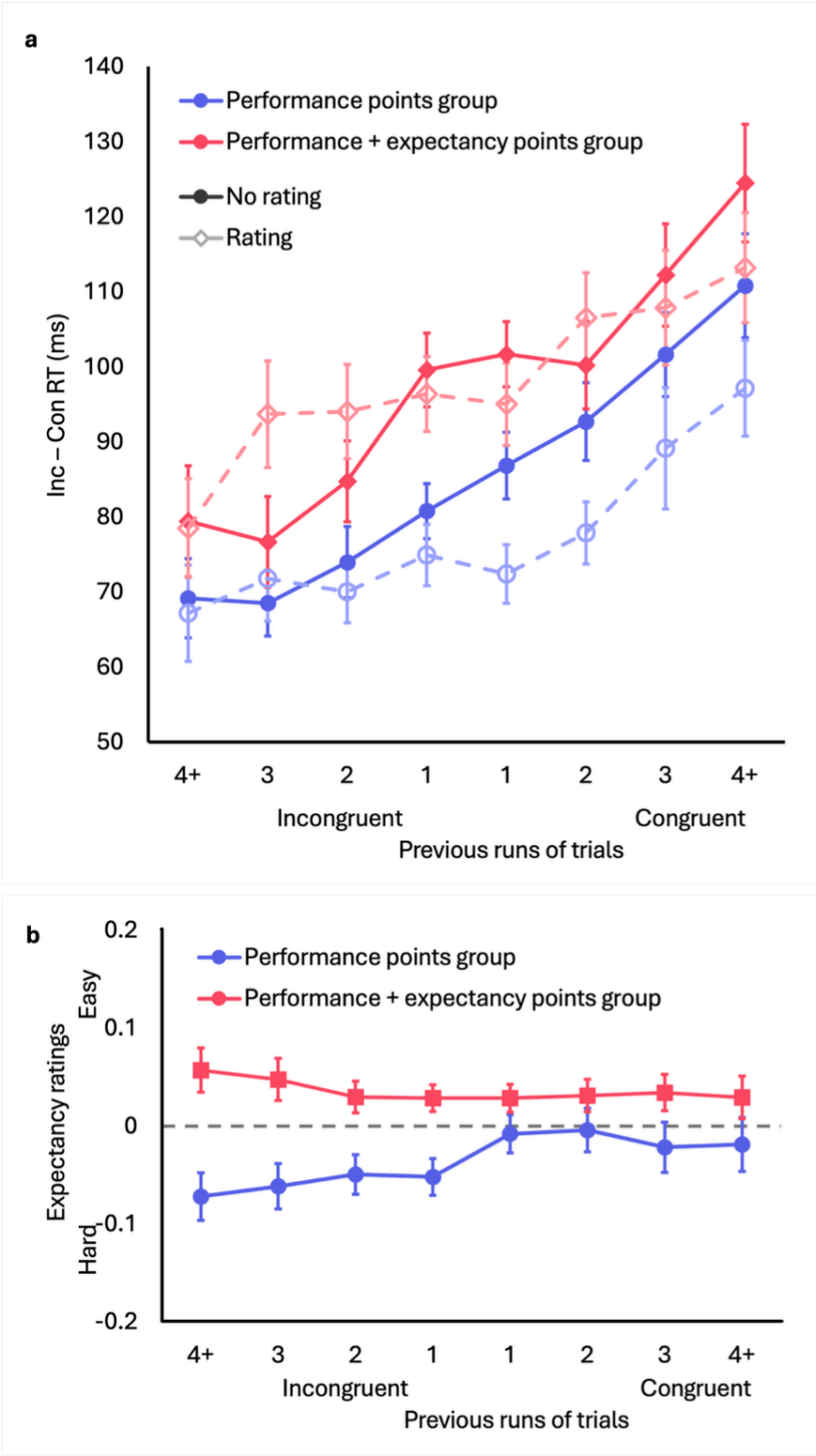


Figure 4

Flanker Effect RT and Expectancy Ratings by Previous Trial Runs. **a)** The flanker effect graph shows the difference between incongruent and congruent trial RT. Dark solid lines represent no rating trials, and light dashed lines correspond to trials with expectancy ratings. Blue lines refer to the performance points group, while red lines depict the performance + expectancy points group. Standard error bars represent one SEM. **b)**Average expectancy ratings for each run length, split by group conditions. The blue line

represents the pattern of expectancy ratings made by the performance points group, while the red line represents that of the performance + expectancy points group. Standard error bars represent one SEM.

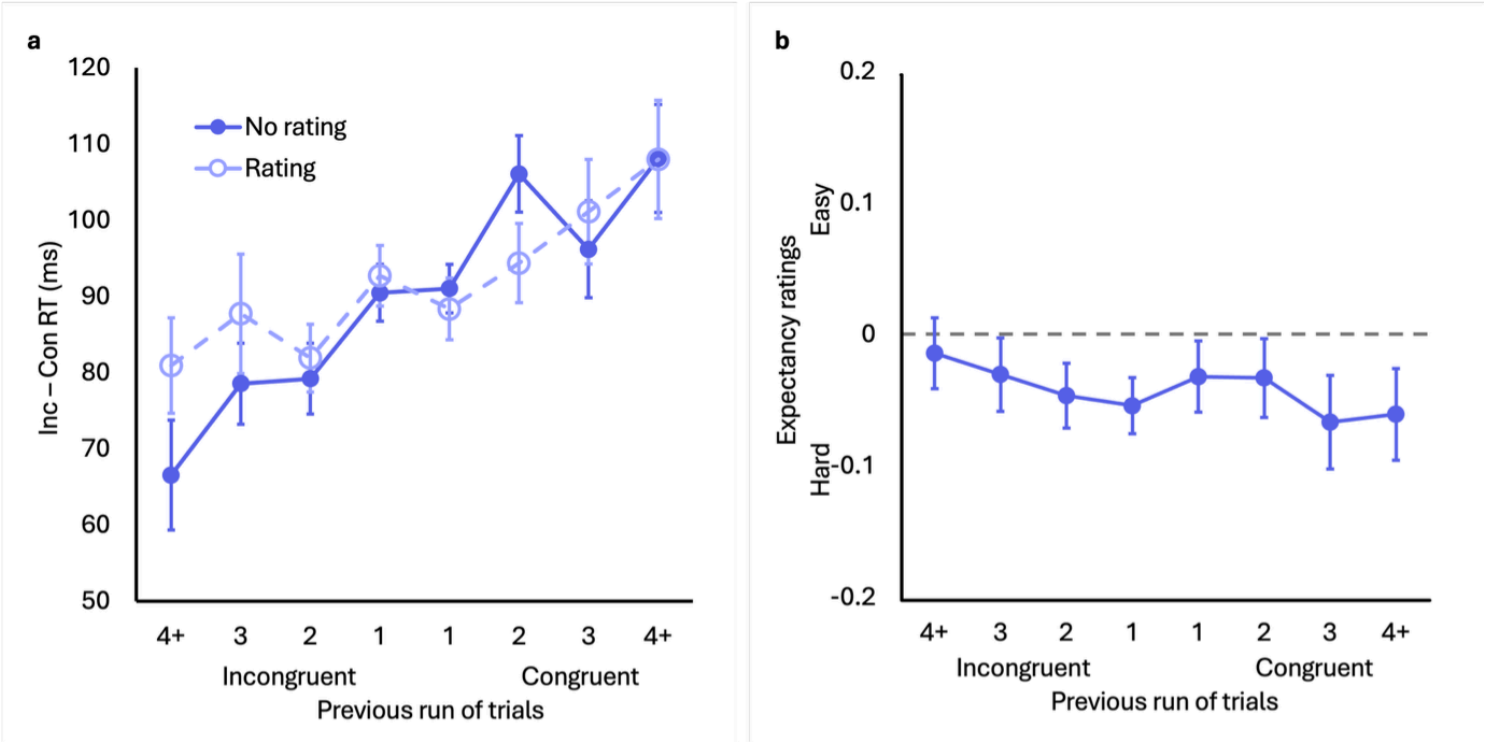


Figure 5

Flanker Effect and Expectancy by Previous Trial Runs. **a)** The flanker effect graph shows the difference between incongruent and congruent trial RT. Dark solid lines represent no rating trials, and light dashed lines correspond to trials with expectancy ratings. Standard error bars represent one SEM. **b)** Average expectancy ratings for each run length. Standard error bars represent one SEM.

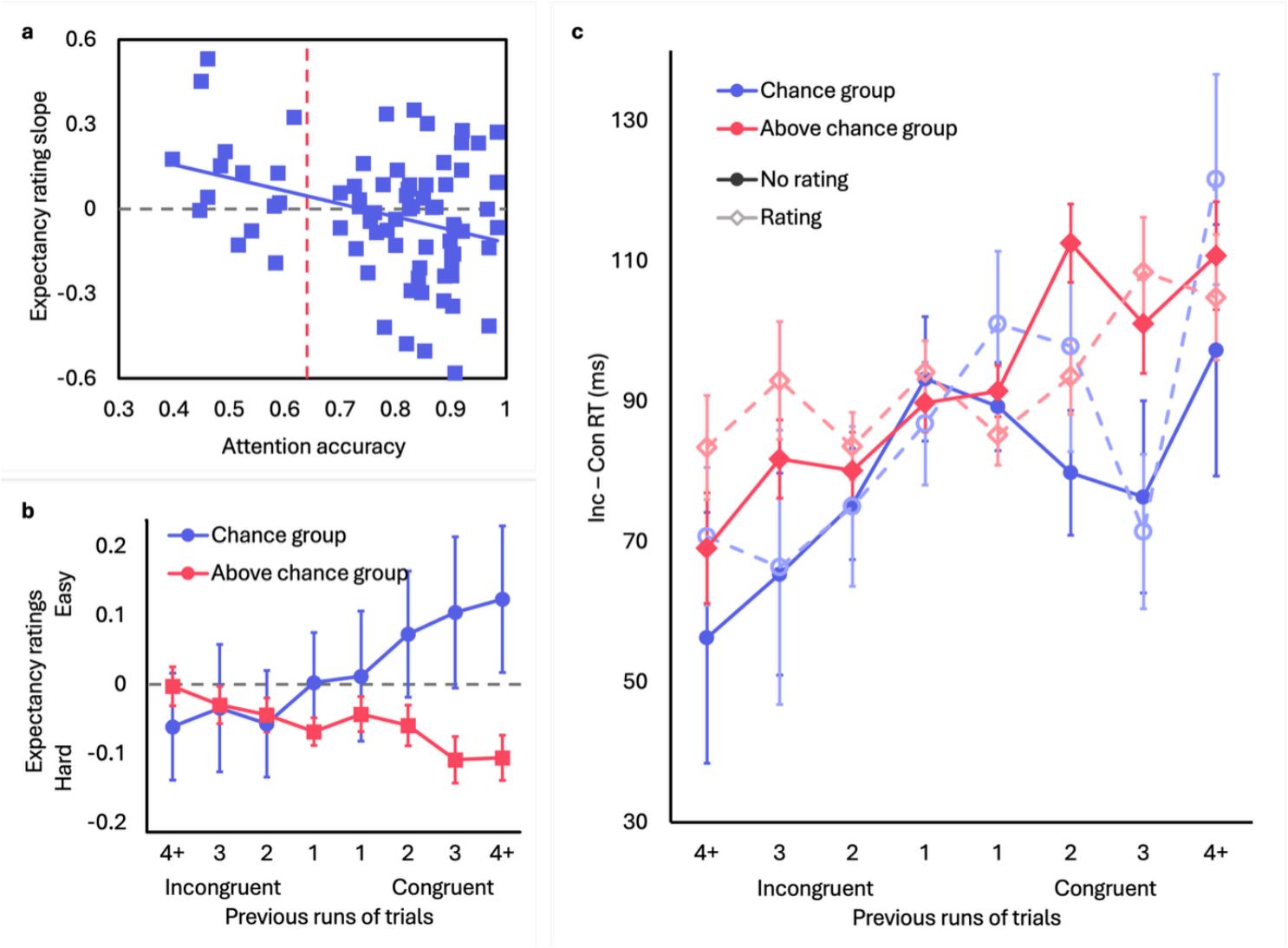


Figure 6

a) Scatterplot of attention accuracy scores by expectancy rating slopes, where each point on the graph represents where each participant lies. The blue solid line represents a linear regression of the scatterplot. The red dotted line indicates the threshold required for participants to be performing above chance in a binomial test (64% of attention check questions). **b)** Expectancy ratings for each attention group. The blue line represents the pattern of expectancy ratings made by the chance group, while the red line represents that of the above chance group. Standard error bars represent one SEM. **c)** Flanker effect RT for each attention group, split by trial type. Dark solid lines represent no rating trials, and light dashed lines correspond to trials with expectancy ratings. Blue lines correspond to the chance group, while red lines refer to the above chance group. Standard error bars represent one SEM.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterials.docx