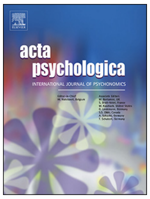
Acta Psychologica 198 (2019) 102869

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy



Multiple expectancies underlie the congruency sequence effect in confound-minimized tasks

Christopher D. Erb[a,*], Andrew J. Aschenbrenner[b]

[a] School of Psychology, University of Auckland, 23 Symonds Street, Building 302, Auckland 1010, New Zealand
[b] Department of Neurology, Washington University in St. Louis, 4488 Forest Park Avenue, St. Louis, MO 63108, USA

ARTICLE INFO

Keywords:
Conflict monitoring
Cognitive control
Congruency sequence effect
Repetition expectancy

ABSTRACT

The congruency sequence effect (CSE) occurs when the congruency effect observed in tasks such as the Eriksen flanker task is smaller on trials preceded by an incongruent trial relative to trials preceded by a congruent trial. The CSE has been attributed to a range of factors including repetition expectancy, conflict monitoring, feature integration, and contingency learning. To clarify the debate surrounding the CSE and the mechanisms underlying its occurrence, researchers have developed confound-minimized congruency tasks designed to control for feature-integration and contingency-learning effects. A CSE is often observed in confound-minimized tasks, indicating that the effect is driven by repetition expectancy, conflict monitoring, or a combination of the two. Here, we propose and test a variant of the repetition expectancy account that emphasizes how multiple expectations can be formed simultaneously based upon the congruency type (congruent vs. incongruent) and the congruency repetition type (congruency repetition vs. congruency alternation) of the most recent trial. Data from confound-minimized versions of the prime-probe task were found to support this novel account. Data from confound-minimized versions of the Eriksen flanker, Simon, and Stroop tasks indicate that feature-integration confounds often remain in these tasks, potentially undermining the conclusions of previous work. We discuss the implications of these findings for ongoing theoretical debates surrounding the CSE.

1. Introduction

Congruency tasks such as the Eriksen flanker task (Eriksen & Eriksen, 1974), the Simon (1969) task, and the Stroop (1935) task have played a central role in the development and assessment of fundamental psychological constructs such as interference control and response inhibition (e.g., MacLeod, 1991; Mullane, Corkum, Klein, & McLaughlin, 2009; Ridderinkhof, Van Den Wildenberg, Segalowitz, & Carter, 2004). Each of these tasks feature congruent trials in which different stimuli or stimulus features cue the same response, and incongruent trials in which different stimuli or stimulus features cue competing responses. For example, in the flanker task, participants are presented with a stimulus array and are instructed to respond according to the centermost stimulus in the array. On congruent trials, each stimulus in the array cues the same response (e.g., ←←←←←). On incongruent trials, the centermost stimulus cues a different response than the surrounding stimuli (e.g., →→←→→).

Performance on congruency tasks is generally assessed in terms of the *congruency effect*, which indexes the difference in performance between incongruent and congruent trials, with larger congruency effects interpreted to reflect less effective control. For example, the congruency effect is commonly used to measure developmental and individual differences in cognitive control (e.g., Aschenbrenner & Balota, 2017; Erb & Marcovitch, 2018a, 2018b; Friedman et al., 2008; Waszak, Li, & Hommel, 2010), and to evaluate the effects of disorder and disease (e.g., Gründler, Cavanagh, Figueroa, Frank, & Allen, 2009; Mullane et al., 2009; Wylie et al., 2009). In recent years, however, particular theoretical and empirical emphasis has been placed on identifying how the congruency effect is modulated by recent experience. For example, a *congruency sequence effect* (CSE) is commonly observed in which the congruency effect is smaller on trials preceded by an incongruent trial relative to those preceded by a congruent trial.

The CSE has been attributed to a wide range of factors, including repetition expectancy, conflict monitoring, feature integration, and contingency learning (for reviews, see Braem, Abrahamse, Duthoo, & Notebaert, 2014; Duthoo, Abrahamse, Braem, Boehler, & Notebaert, 2014; Egner, 2007, 2017; Schmidt, 2018; Schmidt & De Houwer, 2011). According to the *repetition-expectancy account*, the CSE occurs because participants expect that the current trial will match the congruency of the previous trial (Duthoo, Wühr, & Notebaert, 2013; Gratton, Coles, &

---

* Corresponding author.
E-mail address: christopher.erb@auckland.ac.nz (C.D. Erb).

https://doi.org/10.1016/j.actpsy.2019.102869
Received 29 November 2018; Received in revised form 6 June 2019; Accepted 7 June 2019

Donchin, 1992). This expectation results in better performance on congruency-repetition trials (*cC* and *iI trials*, where the lowercase denotes the congruency of the previous trial and the uppercase letter denotes the congruency of the current trial) than on congruency-alternation trials (*iC* and *cI trials*), which produces a smaller congruency effect on trials preceded by an incongruent trial relative to those preceded by a congruent trial.

The *conflict-monitoring account* proposes that the CSE is the product of conflict-driven modulations in top-down control (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Egner, Ely, & Grinband, 2010; Ullsperger, Bylsma, & Botvinick, 2005). On this view, performance is enhanced on iI trials relative to cI trials because control was recently recruited to resolve conflict on iI trials but not on cI trials. Performance is impaired on iC relative to cC trials because top-down control is not needed on congruent trials. Consequently, participants respond faster when they behave in a more automatic manner (cC trials) than when control was recently recruited (iC trials).

The *feature-integration account* highlights how the degree of overlap between the stimulus and response features the previous and current trial can influence an individual's ability to form the appropriate stimulus-response pair on the current trial (Hommel, 2004; Hommel, Proctor, & Vu, 2004). For example, performance can be facilitated on full-overlap trials in which the stimulus-response pair formed on the previous trial (e.g., →→←→→ = LEFT) must also be formed on the current trial. Similarly, performance can be impaired on partial-overlap trials in which only one member of the stimulus-response pair formed on the previous trial (e.g., →→←→→ = LEFT) is featured in the appropriate stimulus-response pair on the current trial (e.g., ←←←←← = LEFT). Importantly, full overlaps only occur on congruency-repetition trials, whereas partial overlaps are more likely to occur on congruency-alternation trials in standard two-alternative forced-choice (2AFC) congruency tasks.

Finally, the *contingency-learning account* proposes that the CSE can occur when certain stimulus and response features co-occur more than others over the course of task (Schmidt & De Houwer, 2011). For example, if a particular word in the Stroop task occurs in a congruent text color (e.g., the word “RED” in red text) more frequently than in other incongruent text colors (e.g., "RED" in blue or green text), participants might be sensitive to the higher contingency between word meaning and text color on congruent trials relative to incongruent trials. This could then lead participants to focus more on word meaning following a congruent trial, resulting in worse performance on cI relative to iI trials (e.g., Erb, Moher, Sobel, & Song, 2016).

1.1. Evaluating the repetition-expectancy account with confound-minimized tasks

In recent years, particular emphasis has been placed on developing confound-minimized tasks that limit the contributions of feature-integration and contingency-learning effects to performance (e.g., Aschenbrenner & Balota, 2017; Jiménez & Méndez, 2013, 2014; Kim & Cho, 2014; Mayr, Awh, & Laurey, 2003; Schmidt & Weissman, 2014). For example, to limit the degree of overlap between the stimulus and response features of the current and previous trial, researchers have developed four-response versions of congruency tasks in which the trials alternate between different stimulus sets that correspond to different response options. In the flanker task, for instance, participants can be presented with an array of upward and downward facing arrows on one trial (e.g., ↓↓↑↓↓) and an array of leftward and rightward facing arrows on the next trial (e.g., ←←→←←). This ensures that no aspect of the stimulus or response from the previous trial appears on the current trial. Further, to counteract contingency-learning effects, the tasks are designed so that none of the stimulus and response features co-occur more frequently than the others.

Research with confound-minimized tasks has revealed significant CSEs in the prime-probe, flanker, Simon, and Stroop task (e.g., Aschenbrenner & Balota, 2017; Kim & Cho, 2014; Schmidt & Weissman, 2014). These findings indicate that factors other than feature integration and contingency learning can generate the CSE. However, the findings do not conclusively differentiate between the repetition-expectancy account and the conflict-monitoring account given that both accounts predict that the CSE will remain after feature-integration and contingency-learning confounds are minimized. A number of studies have therefore sought to test the repetition-expectancy account by having participants explicitly predict the congruency of the upcoming trial or by manipulating how frequently congruency repetitions and alternations occur in different conditions (e.g., Duthoo et al., 2013; Jiménez & Méndez, 2013, 2014).

Jiménez and Méndez (2014), for example, presented participants with a confound-minimized version of the Stroop task that, in addition to blocks of standard trials, featured blocks of trials that required participants to predict whether upcoming trials would be congruent or incongruent. In the standard blocks, the researchers observed the CSE and found that the size of the CSE was modulated by the number of congruency repetitions or alternations that recently occurred. For example, the difference between cC and cI trials was particularly large when the two preceding trials were both congruent (ccC and ccI trials) and, similarly, the difference between iC and iI trials was particularly small when the two preceding trials were both incongruent (iiC and iiI trials).

Crucially, participants' explicit predictions in the prediction blocks did not match the pattern observed in response times in the standard blocks (Jiménez & Méndez, 2014). For example, although participants' response times in the standard blocks were particularly fast on trials that featured the same congruency as the two previous trials (ccC and iiI trials), participants were more likely to predict a congruency alternation when the two previous trials were the same congruency. For instance, following a cC trial, participants were more likely to predict that the upcoming trial would be incongruent (i.e., a ccI trial) as opposed to congruent (i.e., a ccC trial). Thus, participants' explicit predictions in the prediction blocks appeared inconsistent with their response times in the standard blocks and, consequently, inconsistent with the repetition-expectancy account.

Interestingly, when Jiménez and Méndez (2014) evaluated response times in the prediction blocks (as opposed to the standard blocks), the researchers did observe an association between participants' explicit predictions and response times. Participants responded significantly faster when they accurately predicted the congruency of the upcoming trial, consistent with the repetition-expectancy account, as well as previous research by Duthoo et al. (2013). In light of these and other findings, Duthoo et al. (2014, pg. 5) concluded in their review of the literature on the repetition-expectancy account that, “expectancy can exert an influence on control above and beyond conflict-induced adjustments, yet only when these expectancies are explicitly manipulated or registered”. Similarly, Jiménez and Méndez (2013, 2014) interpreted their results as evidence that the CSE observed in standard, confound-minimized congruency tasks reflect conflict monitoring rather than explicit repetition expectancies.

Although the studies reviewed above suggest that repetition expectancies play a limited role in driving the CSE in confound-minimized congruency tasks, it is important to note that these studies investigated participants' *explicit predictions* rather than their *implicit expectations*. Further, the results of these studies indicate that requiring participants to form explicit predictions fundamentally altered their performance on the task. It is therefore unclear whether implicit expectations regarding the qualities of an upcoming trial may be contributing to the CSE when participants are not instructed to generate explicit predictions.

This question is particularly relevant in light of an intriguing finding reported by Jiménez and Méndez (2014); namely, that the CSE observed in the standard (i.e., non-predictive) Stroop blocks was driven entirely by trials that were preceded by a congruency-repetition trial in which the congruency of trial n – 1 repeated that of trial n – 2 (i.e., ccC,

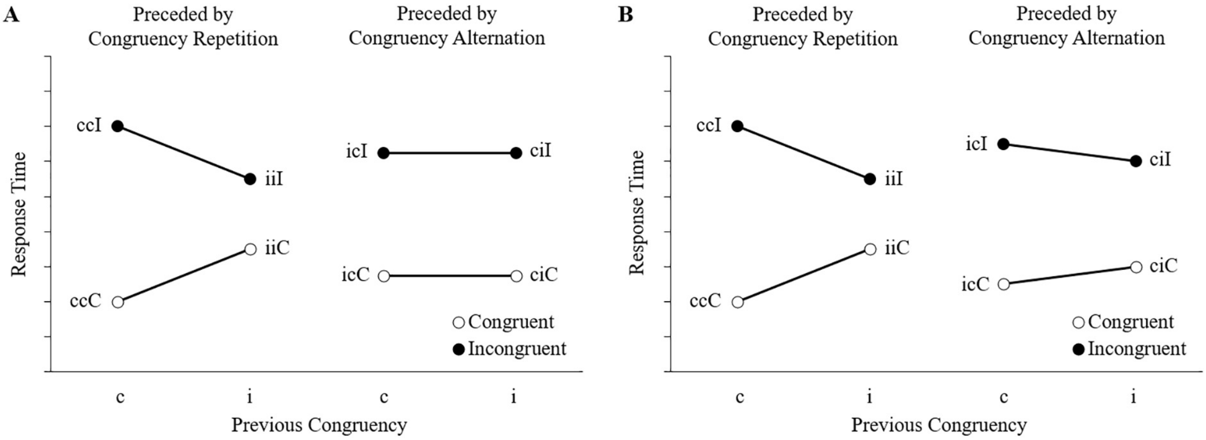


Fig. 1. (A) Illustration of the pattern of response time effects observed by Jiménez and Méndez (2014), with a significant CSE on trials preceded by a congruency-repetition trial (i.e., a cC or iI trial) but no evidence of a CSE on trials preceded by a congruency-alternation trial (i.e., an iC or cI trial). (B) Illustration of the pattern of effects that might be predicted by the repetition-expectancy account and the conflict-monitoring account, with a more robust CSE on trials preceded by a congruency-repetition trial relative to trials preceded by a congruency-alternation trial.

iiC, ccI, and iiI trials). When Jiménez and Méndez restricted their analyses to trials that were preceded by a congruency-alternation trial in which the congruency of trial n – 1 did not match that of trial n – 2 (i.e., icC, ciC, icI, and ciI trials), the CSE was absent entirely. This pattern of effects is illustrated in Fig. 1A.

The effect of the previous trial's repetition type (congruency repetition vs. congruency alternation) observed by Jiménez and Méndez (2014) appears to be incompatible with standard interpretations of both the repetition-expectancy account and the conflict-monitoring account. Both accounts were originally developed to explain how the congruency of the previous trial (as opposed to the congruency of multiple previous trials) could modulate performance on the current trial, yet the results of Jiménez and Méndez suggests that this type of modulation may not occur in confound-minimized congruency tasks. For example, the conflict-monitoring account, as it is generally presented, proposes that conflict-driven adjustments in top-down control should occur after both congruency-repetition and congruency-alternation trials, although the account certainly allows for the possibility that the CSE would be more pronounced following congruency-repetition trials (as illustrated Fig. 1B). In other words, the conflict-monitoring account predicts a difference in *degree* between trials preceded by a congruency-repetition trial and trials preceded by a congruency-alternation trial, not a difference in *kind*. Similarly, the standard version of the repetition-expectancy account allows for the possibility that the CSE would be more pronounced following congruency-repetition trials but it is unclear on this account why the CSE would be entirely absent following congruency-alternation trials.

Here, we propose and test a novel extension of the repetition-expectancy account in light of Jiménez and Méndez's (2014) findings. We call this account the *multiple-expectancies* account. In contrast to the original repetition-expectancy account, this account claims that (a) multiple expectations regarding the qualities of an upcoming trial can be formed simultaneously and (b) that these expectations can be consistent or inconsistent with one another.

1.2. The multiple-expectancies account

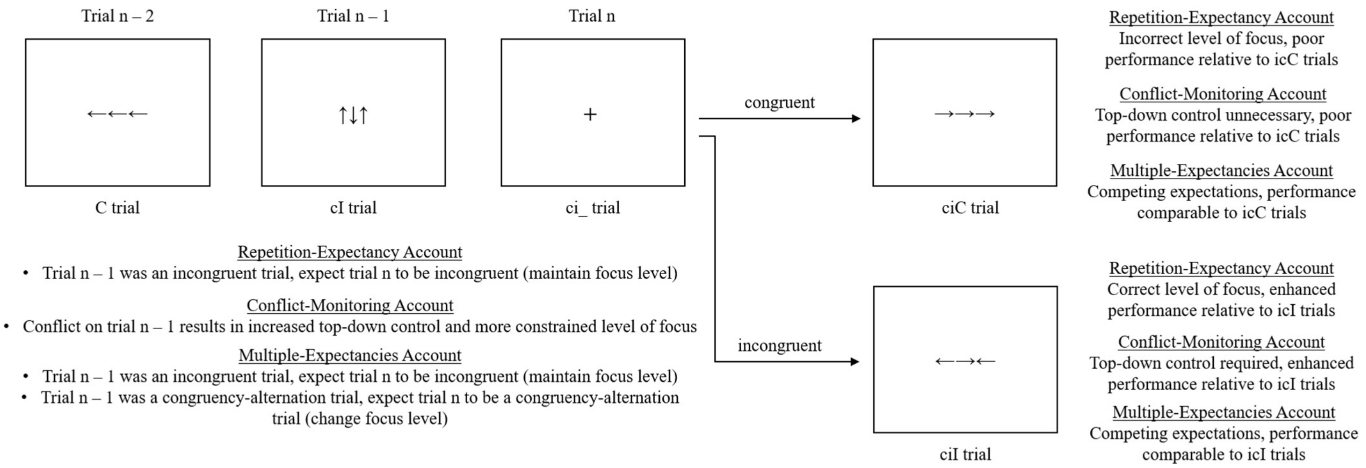
According to the standard version of the repetition-expectancy account, the CSE occurs because participants expect that the level of focus adopted on one trial will be appropriate for the next trial. In the flanker task, for example, incongruent trials require participants to constrain their focus on the target stimulus, whereas congruent trials do not require this level of constrained focus. On congruency-repetition trials, performance is facilitated because the level of focus adopted on the previous trial happens to be appropriate for the current trial. On congruency-alternation trials, performance is impaired because the level of focus adopted on the previous trial is not appropriate for the current trial.

In addition to the level of focus adopted on the previous trial, another salient factor that may influence participants' expectations regarding the upcoming trial concerns whether the most recent trial required participants to maintain or change their level of focus. For instance, transitioning from a congruent trial to an incongruent trial requires participants to change from a less constrained to a more constrained level of focus. Given that such a change of focus occurs on congruency-alternation trials but not on congruency-repetition trials, participants may expect to change their level of focus following congruency-alternation trials and to maintain the same level of focus following congruency-repetition trials. That is, in addition to expecting (1) that the congruency of the upcoming trial will match that of the previous trial (e.g., that a congruent trial will follow a congruent trial and that an incongruent trial will follow an incongruent trial), participants may also expect (2) that the upcoming trial will match the congruency repetition type (congruency alternation vs. congruency repetition) of the previous trial (e.g., that a congruency-alternation trial will follow a congruency alternation-trial and that a congruency-repetition trial will follow a congruency-repetition trial). We will refer to these types of expectations as *congruency-type* and *repetition-type* expectations because the former concerns whether the recently completed trial was congruent or incongruent, whereas the latter concerns whether the recently completed trial was a congruency-alternation trial or a congruency-repetition trial.

It is important to reiterate that within this framework participants' expectations can be consistent or inconsistent with one another. This point is illustrated in Table 1 with regard to congruency-type and repetition-type expectations. For example, both the congruency and the repetition type of a cC trial will generate expectations that the next trial will be congruent because cC trials (1) are congruent and (2) repeat the congruency of the preceding trial, which generates the expectation that the upcoming trial will also be a congruency-repetition trial. If the next trial happened to be congruent (i.e., a ccC trial), these expectations would be confirmed and performance would be facilitated. However, if the next trial happened to be incongruent (i.e., a ccI trial), the expectations formed on the previous trial would be disconfirmed and performance would be impaired. This is illustrated in Table 1 with a combined score of +2 for trial types featuring two accurate expectations (ccC and iiI trials) and a combined score of −2 for trial types featuring two inaccurate expectations (iiC and ccI trials).

Table 1
Illustration of congruency-type and repetition-type expectations as a function of a trial's current congruency type (C vs. I), the congruency type of the preceding trial (c vs. i), and the congruency repetition type of the preceding trial (repetition vs. alternation). Confirmed expectations are scored with a +1, whereas disconfirmed expectations are scored with a −1. Higher combined scores are proposed to correspond to better performance within a congruency type, whereas lower combined scores are proposed to reflect worse performance within a congruency type.

Trial type	Previous congruency type	Congruency-type expectation		Previous congruency repetition type	Repetition-type expectation		Combined score
		Expected	Confirmed?		Expected	Confirmed?	
ccC	c	C	+1	Repetition	Repeat (C)	+1	+2
iiC	i	I	−1	Repetition	Repeat (I)	−1	−2
ccI	c	C	−1	Repetition	Repeat (C)	−1	−2
iiI	i	I	+1	Repetition	Repeat (I)	+1	+2
icC	c	C	+1	Alternation	Alternate (I)	−1	0
ciC	i	I	−1	Alternation	Alternate (C)	+1	0
icI	c	C	−1	Alternation	Alternate (I)	+1	0
ciI	i	I	+1	Alternation	Alternate (C)	−1	0



Repetition-Expectancy Account
- Trial n – 1 was an incongruent trial, expect trial n to be incongruent (maintain focus level)

Conflict-Monitoring Account
- Conflict on trial n – 1 results in increased top-down control and more constrained level of focus

Multiple-Expectancies Account
- Trial n – 1 was an incongruent trial, expect trial n to be incongruent (maintain focus level)
- Trial n – 1 was a congruency-alternation trial, expect trial n to be a congruency-alternation trial (change focus level)

Repetition-Expectancy Account
Incorrect level of focus, poor performance relative to icC trials

Conflict-Monitoring Account
Top-down control unnecessary, poor performance relative to icC trials

Multiple-Expectancies Account
Competing expectations, performance comparable to icC trials

Repetition-Expectancy Account
Correct level of focus, enhanced performance relative to icI trials

Conflict-Monitoring Account
Top-down control required, enhanced performance relative to icI trials

Multiple-Expectancies Account
Competing expectations, performance comparable to icI trials

Fig. 2. Illustration of a ciC trial and a ciI trial. Explanations of the repetition-expectancy account, conflict-monitoring account, and multiple-expectancies account are presented at the bottom left of the figure, whereas the predictions regarding performance on each trial type are provided on the right of the figure.

On certain trial types, congruency-type and repetition-type expectations can diverge. For example, the congruency type of an cI trial would lead participants to expect that the upcoming trial would be incongruent (i.e., participants would expect that the next trial would be an ciI trial), whereas the repetition type of an cI trial would lead participants to expect a congruency alternation (i.e., participants would expect that the next trial would be an ciC trial). Assuming that the strength of these expectations is roughly equal given that the expectations have an equivalent chance of being confirmed or disconfirmed, performance on trials featuring divergent expectations should be middling relative to trials featuring two correct or two incorrect expectations. This is illustrated in Table 1 with a combined score of 0 for all trial types featuring one accurate and one inaccurate expectation (icC, ciC, icI and ciI trials). Fig. 2 illustrates the predictions of the repetition-expectancy account, conflict-monitoring account, and multiple-expectancies account regarding performance on ciC and ciI trials.

This version of the multiple-expectancies account therefore predicts that a robust CSE will occur on trials preceded by a congruency-repetition trial in confound-minimized congruency tasks, with enhanced performance on ccC and iiI trials and impaired performance on iiC and ccI trials. Crucially, the account also predicts that the CSE will be reduced or absent on trials preceded by a congruency-alternation trial, depending on the relative strength of the congruency-type and repetition-type expectations. In other words, the account presents an explanation of the general pattern of effects observed by Jiménez and Méndez (2014). However, at present it is unclear the extent to which the results of Jiménez and Méndez generalize across a range of confound-minimized congruency tasks given that a systematic investigation of the effect of prior congruency repetition type has yet to be conducted.

The current study aims to provide a comprehensive test of the multiple-expectancies account by reanalyzing data from two published studies comprising a total of eight congruency tasks: (1) a study by Schmidt and Weissman (2014) consisting of two experiments that featured confound-minimized versions of the prime-probe task and (2) a study by Aschenbrenner and Balota (2017) consisting of two experiments in which participants completed confound-minimized versions of the flanker, Simon, and Stroop tasks. If the multiple-expectancies account is correct, then robust CSEs will be observed on trials following a congruency-repetition trial, whereas CSEs will be reduced or absent from performance on trials following a congruency-alternation trial (as illustrated in Fig. 1A). If either the repetition-expectation account or the conflict-monitoring account is correct, then evidence of the CSE should be observed in trials following both congruency-repetition and congruency-alternation trials, although a more robust CSE can be expected to occur on trials following a congruency-repetition trial (as illustrated in Fig. 1B).

2. Methods

2.1. Participants

Schmidt and Weissman (2014) obtained data from 16 participants (mean age = 21.2, range = 18–27) in Experiment 1 and 16 (mean age = 21.0, range = 18–30) in Experiment 2. Aschenbrenner and Balota (2017) collected data from 40 young adults (Mean age = 20.3, SD = 2.2; Mean years of education = 13.4, SD = 3.2) in Experiment 1 and 52 young adults (Mean age = 20, SD = 1.4; Mean years of education = 14.1, SD = 1.5) in Experiment 2. A sample of older adult participants were also collected for each experiment, however to maintain consistency with the larger literature on the CSE, which mainly utilized younger adults, the older participants from Aschenbrenner and Balota (2017) are not considered here.

2.2. Tasks

Each experiment discussed below implemented a confound-minimized design such that different stimuli were presented on alternating trials (e.g., left vs. right arrows on odd-numbered trials and up vs. down facing arrows on even-numbered trials). This design ensures that no overlap occurs between the stimulus and response features of trial n and trial n – 1. However, given that the tasks alternate between different stimulus sets, the stimulus and response features of trial n could overlap with those of trial n – 2. We return to this consideration in Sections 4 and 5.

Schmidt and Weissman (2014) used a prime-probe design in which a distractor stimulus was displayed at fixation followed by a target. Experiment 1 utilized arrow stimuli in which the distractor consisted of an array of identical arrows (e.g., →→→→→) followed by either a matching (→) or mismatching (←) target. Experiment 2 utilized word stimuli rather than arrows (e.g., the distracter would be a series of spatial words such as “LEFT” stacked vertically, followed by a matching probe “LEFT” or mismatching probe “RIGHT”). Each experiment consisted of 8 blocks of 96 trials preceded by 24 practice trials. In both experiments, the distractor appeared 2000 milliseconds after the onset of the distractor array from the previous trial.[1]

Aschenbrenner and Balota (2017) administered three congruency tasks in each experiment. A flanker task in which participants identified the center letter of an array (e.g., HHKHH), a Simon task in which participants identified location words and ignored spatial location (e.g., ABOVE printed below a fixation point) and a standard Stroop paradigm in which they identified colors while ignoring the word (e.g., "BLUE" printed in red text). Experiments 1 and 2 of their study were identical with the following two exceptions: Experiment 2 included a response deadline procedure and Experiment 2 consisted of fewer total trials (384 trials per task in Experiment 1, 192 trials per task in Experiment 2). The tasks used in the two experiments were designed so that each trial would only initiate after participants pressed a button. Consequently, the tasks did not feature set inter-stimulus intervals. The average inter-stimulus interval is estimated to have been between 2400 and 3000 milliseconds. Full details for each study can be found in the original manuscripts.

2.3. Statistical analyses

In order to evaluate the extent to which the CSE is modulated by the congruency type and the congruency repetition type of the previous trial, we conducted a series of linear mixed effects (LME) models that included random intercepts across participants using the lme4 (Bates et al., 2015) package in R. The congruency type of the current trial (C vs. I), the congruency type of the previous trial (C vs. I), and the congruency repetition type of the previous trial (repetition vs. alternation) were entered into the model as dummy-coded factors and all two-way and three-way interactions were included. The significance of the main effects and interactions were evaluated using chi-square tests with Type III sums of squares implemented in the ANOVA function from the car (Fox & Weisberg, 2011) package in R. We then conducted planned follow-up contrasts of the CSE within levels of the previous trial's congruency repetition type. Due to computation restrictions, denominator degrees of freedom are not calculated in the analysis of error rates.

Given our focus on the CSE, only model terms that involve the interaction between the congruency type of the current and previous trial will be discussed (i.e., the CSE); however, a complete account of the results can be found is available in Section 1 of the Supplementary material. In order to avoid any confounds due to post-error slowing, only accurate trials preceded by two accurate trials were included in our response time analyses. Error rate analyses included all trials, regardless of whether an error had been committed on either of the preceding two trials. We did not want to obscure any trends in the data that arise from specific trial sequences and thus chose not to implement any trimming of reaction times prior to analysis given that particular trial types may be disproportionally affected by trimming procedures.

The data files from the Schmidt and Weissman (2014) study are available online at https://doi.org/10.1371/journal.pone.0102337. The data and analysis files used in the current study's reanalysis of the Schmidt and Weissman (2014) data are available upon request. The data and analysis files for the Aschenbrenner and Balota (2017) reanalysis are available online at https://osf.io/6nh4g/?view_only=ba2a0bfc8c8543abafa7b7d8e79ab984.

3. Results

3.1. Schmidt and Weissman, Experiments 1 and 2

Response times revealed a significant CSE by previous congruency repetition type interaction in Experiment 1, (χ^2 = 46.83, p < .001), with follow-up contrasts confirming the CSE was significant following congruency-repetition trials, F(1,10708) = 80.73, p < .001, but was not significant following congruency-alternation trials, F(1,10708) = 0.57, p = .45 (see Fig. 3A). Similarly, error rates in Experiment 1 revealed a significant CSE by previous congruency repetition type interaction, (χ^2 = 15.16, p < .001), with follow-up contrasts confirming the CSE was significant following congruency-repetition trials, F(1,inf) = 14.73, p < .001, but was not significant following congruency-alternation trials, F(1,inf) = 3.41, p = .06.

Response times in Experiment 2 also revealed a significant CSE by previous congruency repetition type interaction, (χ^2 = 33.20, p < .001). Follow-up contrasts confirmed a significant CSE following congruency-repetition trials, F(1,10647) = 61.20, p < .001, but a non-significant effect following congruency-alternation trials, F(1,10647) = 0.13, p = .72 (see Fig. 3B). Error rates in Experiment 2 did not reveal a reliable CSE by previous congruency repetition type interaction, (χ^2 = 1.22, p = .27). Follow-up contrasts confirmed there was no CSE following congruency-repetition trials, F(1,inf) = 1.70, p = .19, nor following congruency-alternation trials, F(1,inf) = 0.07, p = .79.

3.2. Aschenbrenner and Balota, Experiment 1

3.2.1. Flanker task

Response times revealed a significant interaction between the CSE and previous congruency repetition type, (χ^2 = 13.82, p < .001). As shown in Fig. 4A, the CSE was significant following congruency-repetition trials, F(1,13633) = 17.07, p < .001, but was not significant following congruency-alternation trials, F(1,13633) = 1.24, p = .27. Error rates did not reveal a significant interaction between CSE and previous congruency repetition type, (χ^2 = 0.34, p = .56). Follow-up contrasts confirmed no CSE was present following congruency-repetition trials, F(1,inf) = 0.04, p = .85, nor following congruency-alternation trials, F(1,inf) = 0.40, p = .53.

---

[1] It is important to consider inter-stimulus intervals and response-to-stimulus intervals when evaluating CSEs, as recent research indicates that the CSE can dissipate as these intervals increase (Duthoo, Abrahamse, Braem, & Notebaert, 2014; Egner et al., 2010). For instance, Egner et al. (2010) found that the CSE dissipate over inter-stimulus intervals bins ranging from 500 to 7000 milliseconds. Additionally, Duthoo et al. (2013) note that very brief response-to-stimulus intervals may not allow participants to prepare for expected events.

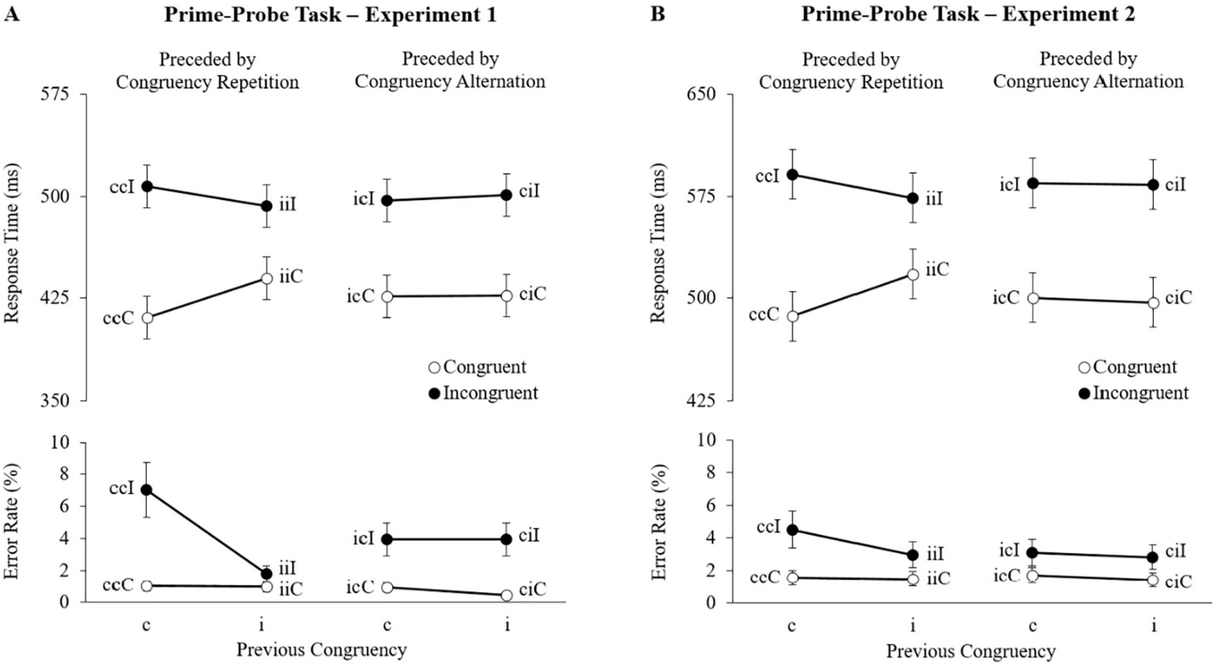


Fig. 3. Results from the reanalysis of Schmidt and Weissman (2014). Response time and error rate performance in (A) Experiment 1 and (B) Experiment 2 as a function of current congruency (C vs. I), previous congruency (c vs. i), and the previous trial's congruency repetition type (repetition vs. alternation). Error bars display standard errors.

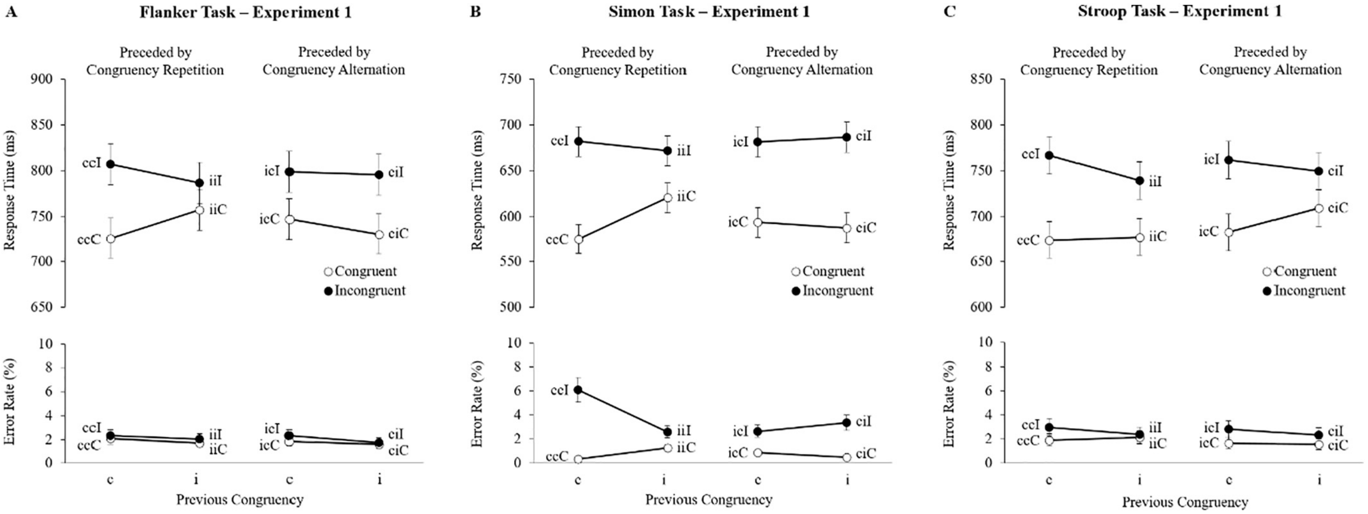


Fig. 4. Results from the reanalysis of Aschenbrenner and Balota (2017) Experiment 1. Response time and error rate performance in the (A) Eriksen flanker task, (B) Simon task, and (C) Stroop task as a function of current congruency (C vs. I), previous congruency (c vs. i), and the previous trial's congruency repetition type (repetition vs. alternation). Error bars display standard errors.

3.2.2. Simon task

Response times in the Simon task revealed a significant CSE by previous congruency repetition type interaction task, ($\chi^2 = 26.62$, $p < .001$), indicating a reliable CSE following congruency-repetition trials, F(1,13627) = 37.26, $p < .001$, but a non-significant effect following congruency-alternation trials, F(1,13627) = 1.46, $p = .23$ (see Fig. 4B). Error rates also revealed a significant CSE by previous congruency repetition type interaction, ($\chi^2 = 31.42, p < .001$). Follow-up comparisons revealed the CSE was significant following congruency-repetition trials, F(1,inf) = 31.74, $p < .001$, and was significant following congruency-alternation trials, F(1,inf) = 4.95, $p = .03$. Notably, the pattern of effects observed on trials following congruency-alternation trials was the inverse of the standard CSE, with a smaller congruency effect on trials preceded by a congruent trial.

3.2.3. Stroop task

The CSE by previous congruency repetition type interaction was not significant in response times in the Stroop task, ($\chi^2 = 0.15$, $p = .70$). Follow-up comparisons confirmed the CSE was significant following both congruency-repetition trials, F(1,6774) = 3.91, $p = .05$, as well as following congruency-alternation trials, F(1,6774) = 6.34, $p = .01$ (see Fig. 4C). Error rates revealed no evidence for a CSE by previous congruency repetition type interaction, ($\chi^2 = 0.23$, $p = .63$), indicating no CSE following congruency-repetition trials, F(1,inf) = 1.04, $p = .31$, nor following congruency-alternation trials, F(1,inf) = 0.10, $p = .76$.

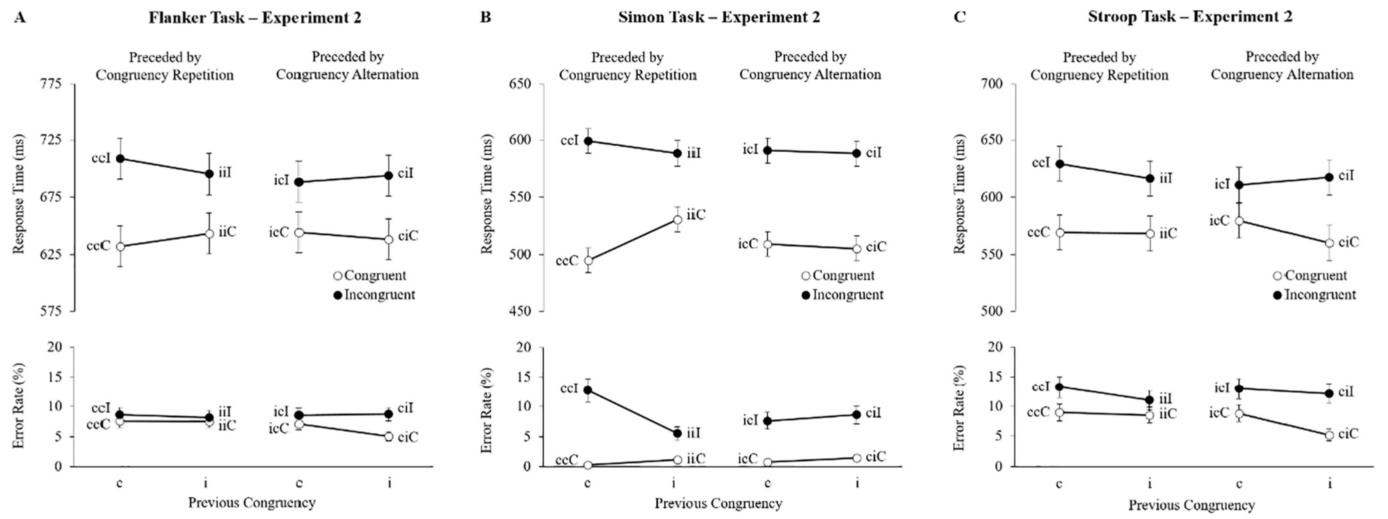


Fig. 5. Results from the reanalysis of Aschenbrenner and Balota (2017) Experiment 2. Response time and error rate performance in the (A) Eriksen flanker task, (B) Simon task, and (C) Stroop task as a function of current congruency (C vs. I), previous congruency (c vs. i), and the previous trial's congruency repetition type (repetition vs. alternation). Error bars display standard errors.

3.3. *Aschenbrenner and Balota, Experiment 2*

3.3.1. *Flanker task*

Response times revealed a significant interaction between the CSE and previous congruency repetition type in the flanker task, ($\chi^2 = 4.15$, $p = .04$), and follow-up contrasts confirmed that the CSE was significant following congruency-repetition trials, F(1,6931) = 3.91, $p = .05$, but was not significant following congruency-alternation trials, F(1,6930) = 0.81, $p = .37$ (see Fig. 5A). The interaction between the CSE and previous congruency repetition type was not significant in error rates, ($\chi^2 = 2.12$, $p = .15$), and follow-up contrasts confirmed the CSE was not significant following congruency-repetition trials, F (1,inf) = 0.07, $p = .80$, nor following congruency-alternation trials, F (1,inf) = 3.12, $p = .08$.

3.3.2. *Simon task*

Response times in the Simon task revealed a significant CSE by previous congruency repetition type interaction, ($\chi^2 = 9.83$, $p = .002$), and the follow-up contrasts confirmed the pattern was in the expected direction. Specifically, the CSE was significant following congruency-repetition trials, F(1,3616) = 18.85, $p < .001$, but was not significant following congruency-alternation trials, F(1,3615) = 0.02, $p = .90$ (see Fig. 5B). The CSE by previous congruency repetition type interaction was marginal in error rates, ($\chi^2 = 3.53$, $p = .06$), however the follow-up contrasts indicated the CSE was reliable following congruency-repetition trials, F(1,inf) = 10.11, $p = .002$, but was not significant following congruency-alternation trials, F(1,inf) = 1.17, $p = .28$.

3.3.3. *Stroop task*

Response times revealed a marginal CSE by previous congruency repetition type interaction in the Stroop task, ($\chi^2 = 3.44$, $p = .06$), however the follow-up contrasts indicated the CSE was not significant following either congruency-repetition trials F(1,3331) = 0.66, $p = .42$, nor following alternation trials F(1, 3331) = 3.26, $p = .07$ (see Fig. 5C). The CSE by previous congruency repetition type interaction was not significant in error rates, ($\chi^2 = 2.81$, $p = .09$). Follow-up contrasts indicated the CSE following congruency-repetition trials was not significant, F(1,inf) = 0.38, $p = .54$, nor following congruency-alternation trials, F(1,inf) = 2.96, $p = .09$.

4. Discussion

The results presented above demonstrate that the CSE observed in the majority of confound-minimized congruency tasks is fundamentally modulated by the congruency repetition type of the preceding trial. Of the seven datasets to reveal a significant CSE in response times or error rates, all seven revealed a CSE on trials preceded by a congruency-repetition trial, whereas only one revealed a CSE on trials preceded by a congruency-alternation trial (see Table 2). These findings are consistent with the version of the multiple-expectancies account introduced above, which proposed (a) that participants form expectations regarding the upcoming trial based upon the congruency type and the congruency repetition type of the most recent trial, and (b) that these expectations can be consistent or inconsistent with one another. On this view, the CSE is particularly pronounced on trials preceded by a congruency-repetition trial because both the congruency type and the

Table 2
Summary of the results across the eight different datasets with regard to whether the CSE was present (+) or absent (−) in response times and error rates. Note: * indicates that a significant interaction between previous and current congruency was observed but not in the direction associated with the CSE.

		Preceded by congruency repetition		Preceded by congruency alternation	
Study	Task	RT	Error rate	RT	Error rate
Schmidt and Weissman (2014)	Prime-Probe (Exp. 1)	+	+	–	–
	Prime-Probe (Exp. 2)	+	–	–	–
Aschenbrenner and Balota (2017)	Flanker (Exp. 1)	+	–	–	–
	Simon (Exp. 1)	+	+	–	*
	Stroop (Exp. 1)	+	–	+	–
	Flanker (Exp. 2)	+	–	–	–
	Simon (Exp. 2)	+	+	–	–
	Stroop (Exp. 2)	–	–	–	–

congruency repetition type of the preceding trial lead participants to adopt the appropriate level of focus on ccC and iiI trials and the inappropriate level of focus on iiC and ccI trials. The CSE is less likely to be observed on trials preceded by a congruency-alternation trial because the congruency type and congruency repetition type of the preceding trial encourage participants to adopt different levels of focus, resulting in similar levels of performance on icC relative to ciC trials and on icI relative to ciI trials.

The interpretation of the results offered by the multiple-expectancies account is predicated on the notion that the confound-minimized tasks used by Schmidt and Weissman (2014) and Aschenbrenner and Balota (2017) successfully minimized feature-integration confounds. If such confounds remained, however, an alternative interpretation of the results could be provided.[2] For instance, given the alternating structure of these tasks, it is conceivable that performance on trial n is influenced by the degree to which the stimulus and response features of trial n overlap with those of trial n – 2. That is, even though confound-minimized tasks ensure that trial n does not feature any overlap with trial n – 1, overlap between trial n and trial n – 2 may nevertheless impact performance. Indeed, data from Mayr et al. (2003, Experiment 2) indicates that the stimulus and response features of trial n – 2 can impact performance on trial n in confound-minimized flanker tasks.

If substantial feature-integration confounds do remain in the confound-minimized tasks currently employed in the literature, performance that is typically assumed to reflect conflict monitoring or repetition expectancy could actually arise from a combination of these effects with residual feature-integration confounds. For instance, the feature-integration account predicts that performance should be facilitated on full-overlap trials in which the stimulus and response features of trial n are identical to those of the previous trial (note that the previous trial of interest in this case is trial n – 2 rather than the immediately preceding trial). If the stimulus and response features of trial n overlap fully with those of trial n – 2 and the congruency of trial n matches that of trial n – 1, then performance could benefit from feature-integration effects as well as repetition expectancy, conflict monitoring, or both.

On this view, which we will refer to as the *confound-laden account*, the CSE would be particularly pronounced in trials following a congruency-repetition trial (i.e., ccC, iiC, ccI, and iiI trials) because the feature-integration, repetition-expectancy, and conflict-monitoring accounts predict effects in the same direction. For instance, performance on ccC and iiI trials would be more likely to benefit from feature-integration effects as well as repetition expectancy, conflict monitoring, or both. Performance on iiC and ccI trials, however, would be more likely to be impaired by feature-integration effects as well as repetition expectancy, conflict monitoring, or both.

The CSE would be reduced or absent in trials following a congruency-alternation trial (i.e., icC, ciC, icI, and ciI trials), on this view, because the feature-integration account predicts effects that are opposed to the effects predicted by the repetition-expectancy and conflict-monitoring accounts. For instance, the feature-integration account predicts impaired performance on icC trials because the stimulus and response features of trial n – 2 partially overlap with those of trial n. Alternatively, the repetition-expectancy account predicts facilitated performance on icC trials because the congruency of trial n – 1 matches that of trial n. Thus, the confound-laden account presents an alternative explanation of the pattern of results illustrated in Fig. 1A and observed in the majority of datasets evaluated in the current study.

Although the tasks featured in the current study do not allow for a direct comparison of the multiple-expectancies account and the confound-laden account, one can evaluate whether repetition-priming confounds (a particular subset of feature-integration confounds) remained in the tasks by comparing instances in which the response provided on trial n matched that of trial n – 2 (i.e., *response-repetition trials*) with instances in which the response provided on trial n did not match that of trial n – 2 (i.e., *response-alternation trials*). If repetition-priming confounds remained in the tasks, then the CSEs observed on trials preceded by a congruency repetition should be larger on response-repetition trials than response-alternation trials. This is because only response-repetition trials can be full-overlap trials (i.e., trials in which the stimulus and response features of trial n match those of trial n – 2), whereas response-alternation trials necessarily feature either no overlap or partial overlap with trial n – 2. In order to evaluate the confound-laden account, we therefore performed exploratory analyses investigating the effects of response repetition type on the CSEs observed in trials preceded by a congruency-repetition trial.[3]

5. Exploratory analyses

We analyzed response times and error rates as a function of current trial congruency, previous trial congruency, and response repetition type (i.e., whether the response required on trial n repeated or alternated relative to that of trial n – 2). We performed these analyses only on trials preceded by a congruency-repetition trial (e.g., ccI, ccC, iiC and iiI trials) and only for tasks that revealed a significant CSE in the primary analyses. Factors were dummy-coded and all two and three way interactions were included. Given the hypothesis being tested, we conducted follow-up comparisons of the previous by current trial congruency (CSE) interaction within levels of response repetition type, regardless of the significance of the three-way interaction.

5.1. *Schmidt and Weissman, Experiments 1 and 2*

In Experiment 1, the three-way interaction among current, previous trial congruency, and response repetition type was not significant in response times, ($\chi^2 = 0.60$, $p = .44$), indicating the CSE was not moderated by response repetition type. Follow-up comparisons confirmed the presence of a significant CSE in response-repetition trials, F(1,5425) = 31.06, $p < .001$, and in response-alternation trials, F(1,5425) = 50.03, $p < .001$ (see Fig. 6A). Similarly, in the analysis of error rates in Experiment 1, the three-way interaction was not significant, ($\chi^2 = 0.25$, $p = .62$). Follow-up comparisons revealed a marginal CSE in response-repetition trials, F(1,inf) = 3.26, $p = .07$, and a significant CSE in response-alternation trials, F(1,inf) = 11.01, $p = .001$ (see Fig. 6B).

In the analysis of response times in Experiment 2, the three-way interaction was not significant, ($\chi^2 = 0.10$, $p = .76$), with follow-up comparisons confirming the presence of a significant CSE in response-repetition trials, F(1,5353) = 30.17, $p < .001$, as well as in response-alternation trials, F(1,5353) = 30.24, $p < .001$ (see Fig. 6C).

5.2. *Aschenbrenner and Balota, Experiment 1*

5.2.1. *Flanker task*

The three-way interaction was significant in response times, ($\chi^2 = 3.78$, $p = .05$), and follow-up comparisons indicated there was a significant CSE in response-repetition trials, F(1,6735) = 17.50, $p < .001$. However, the CSE was not significant in response-alternation trials, F(1,6735) = 2.39, $p = .12$ (see Fig. 7A).

---

[2] We would like to thank D. H. Weissman for bringing this alternative interpretation to our attention.

[3] Note that evaluating the effect of response type on the CSEs observed in trials preceded by a congruency-alternation trial is complicated by the fact that putative feature-integration effects would be operating in the opposite direction of putative repetition-expectancy or conflict-monitoring effects (e.g., ciI trials featuring a response repeat).

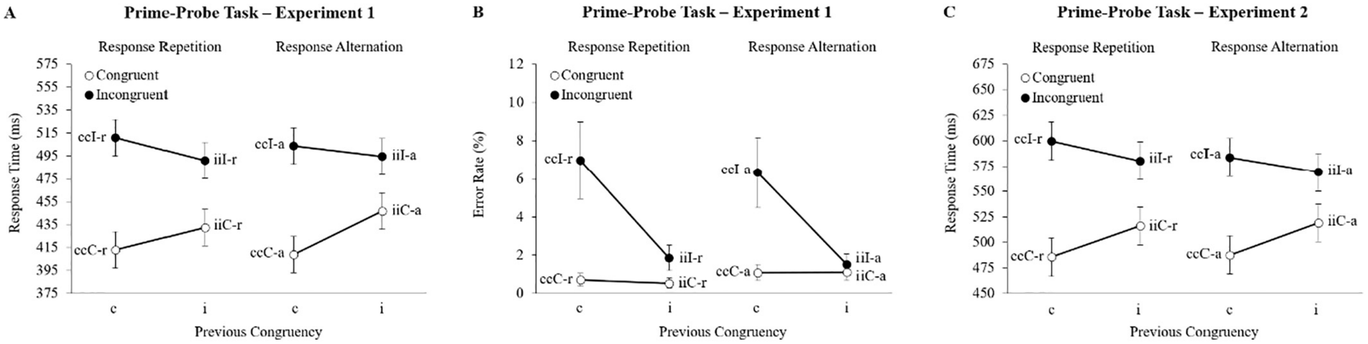


Fig. 6. Results from the exploratory reanalysis of Schmidt and Weissman (2014) evaluating the effects of response repetition type (repetition vs. alternation) on the subset of trials exhibiting a significant CSE in the previous analyses. (A) Response time and (B) error rate performance from Experiment 1 and (C) response time performance from Experiment 2. Error bars display standard errors.

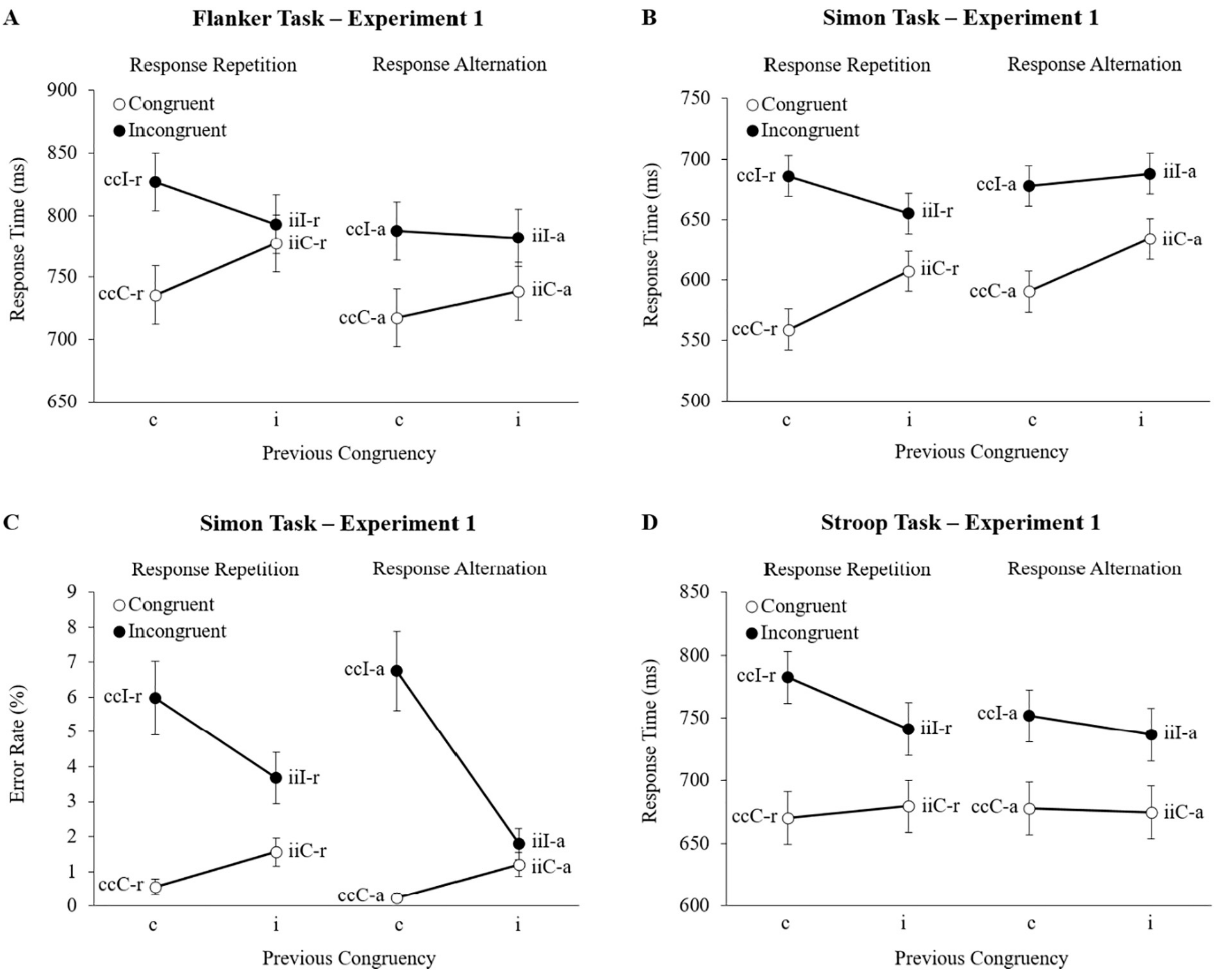


Fig. 7. Results evaluating the effects of response repetition type (alternation vs. repetition) on the CSEs observed in trials preceded by a congruency-repetition trial in Experiment 1 of Aschenbrenner and Balota (2017). (A) Response time performance in the flanker task, (B) response time and (C) error rate performance in the Simon task, and (D) response time performance in the Stroop task. Error bars display standard errors.

5.2.2. *Simon task*

The three-way interaction in response times was significant, ($\chi^2 = 6.57$, $p = .01$). Follow-up comparisons revealed a significant CSE in response-repetition trials, F(1,6830) = 39.92, $p < .001$. Although a significant interaction between current congruency and previous congruency was observed in response-alternation trials, F(1,6830) = 7.53, $p = .006$, the size of the interaction effect was smaller relative to the effect observed in response-repetition trials. Additionally, the pattern of effects observed in response-alternation trials did not match the pattern typically associated with the CSE (see Fig. 7B). A similar pattern was observed in the error rates where the three-way interaction was marginal, ($\chi^2 = 3.33$, $p = .07$), however a significant CSE emerged in both response-repetition trials, F(1,inf) = 10.73, $p = .001$, as well as in response-alternation trials, F(1,inf) = 21.01, $p < .001$ (see Fig. 7C).

5.2.3. *Stroop task*

The three-way interaction was not significant in response times, ($\chi^2 = 1.92$, $p = .16$), however follow-up comparisons indicated the CSE was only significant in response-repetition trials, F(1,3358) = 6.46, $p = .01$, but not in response-alternation trials, F(1,3359) = 0.35,

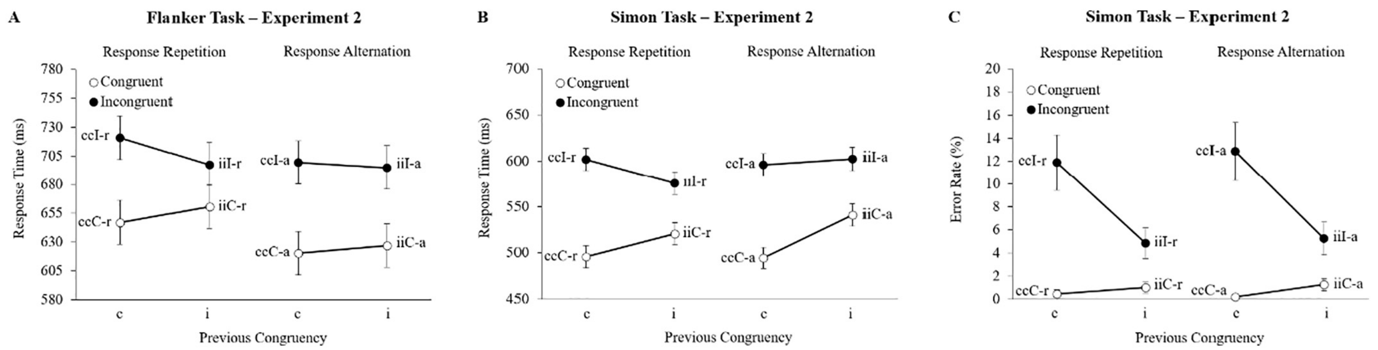


Fig. 8. Results from the exploratory reanalysis of Experiment 2 from Aschenbrenner and Balota (2017) evaluating the effects of response repetition type (alternation vs. repetition) on the subset of trials exhibiting a significant CSE in the previous analyses. (A) Response time performance in the flanker task, and (B) response time and (C) error rate performance in the Simon task. Error bars display standard errors.

p = .55.

5.3. Aschenbrenner and Balota, Experiment 2

5.3.1. Flanker task

In the response time analysis, the three-way interaction was not reliable, (χ^2 = 1.07, p = .30). However, the follow-up comparisons revealed that the CSE was significant in response-repetition trials, F (1,3412) = 4.43, p = .04, but not in response-alternation trials, F (1,3412) = 0.43, p = .51 (see Fig. 8A).

5.3.2. Simon task

The three-way interaction was not significant in response times, (χ^2 = 0.28, p = .60), and follow-up comparisons revealed a significant CSE in both response-repetition trials, F(1,1819) = 12.42, p < .001, and in response-alternation trials, F(1,1819) = 7.84, p = .005. However, the pattern of effects observed in response-alternation trials did not match the pattern typically associated with the CSE (see Fig. 8B). Results were similar in the analysis of error rates as the three-way interaction was not significant, (χ^2 = 0.66, p = .42), and the CSE was significant in both response-repetition trials, F(1,inf) = 3.94, p = .05, and response-alternation trials, F(1,inf) = 6.90, p = .009 (see Fig. 8C).

6. Discussion

In order to test the confound-laden account of the results presented in Section 3, we performed exploratory analyses evaluating whether the CSEs observed on trials preceded by a congruency-repetition trial were more pronounced when the response provided on trial n matched that of trial n – 2. If this were the case, it would suggest that the confound-minimized tasks commonly employed in contemporary research contain lingering repetition-priming confounds. This would, in turn, potentially undermine the conclusions presented in previous studies featuring the tasks.

The prime-probe tasks used by Schmidt and Weissman (2014) did not reveal evidence of lingering repetition-priming confounds. That is, the CSEs observed in response times and error rates in the tasks were not reduced in response-alternation trials relative to response repetition trials. These results therefore bolster the multiple expectancies account of performance in confound-minimized prime-probe tasks. However, it must be noted that the results of these exploratory analyses cannot rule out the possibility that other feature-integration confounds stemming from partial-overlap trials may have remained in the task (e.g., iiC-r trials featured the same target as trial n – 2 but different distractors, whereas iiC-a trials featured the same distractors as trial n – 2 but different targets).

In contrast to the prime-probe tasks, many of the congruency tasks used by Aschenbrenner and Balota (2017) did reveal evidence of lingering repetition-priming confounds. For instance, response times from the flanker task used in Experiment 1 of their study revealed a marginally significant three-way interaction (p = .05) among current congruency type, previous congruency type, and response repetition type, with follow-up tests revealing a significant CSE in response repetition trials but not response-alternation trials. Similarly, response times from the Simon task used in Experiment 1 of their study revealed a significant three-way interaction (p = .01). Although follow-up tests revealed significant interactions between current and previous congruency on both response-alternation trials and response-repetition trials, the CSE observe in response-alternation trials was smaller than the CSE observed in response-repetition trials. Further, in contrast to the pattern of effects most commonly associated with the CSE, response times were descriptively *slower* on iiI-a than ccI-a trials (where "-a" indicates a response alternation).

Given that repetition-priming confounds appear to have remained in the tasks used by Aschenbrenner and Balota (2017), it is unclear whether the results presented in Sections 3.2 and 3.3 were influenced by expectations regarding the congruency repetition type of the preceding trial. A clean test of the multiple-expectancies account requires tasks that are truly confound-minimized. Future research should therefore aim to develop versions of the flanker, Simon, and Stroop tasks that more effectively minimize feature-integration confounds. This could be accomplished, for example, by further expanding the tasks into eight-alternative forced-choice (8AFC) tasks in which participants would have four possible responses for each hand. Previous research with 8AFC congruency tasks indicates that such an approach could be effective (e.g., Hazeltine, Lightman, Schwarb, & Schumacher, 2011, Experiment 3), although steps would need to be taken (a) to ensure that confounds were eliminated rather obscured and (b) to minimize unwanted memory or task-switching demands.

Despite the limitations of the tasks used by Aschenbrenner and Balota (2017), our reanalysis of the tasks present two important conclusions. First, performance on these tasks is often modulated by the congruency repetition type of the preceding trial. Although future research is needed to identify the extent to which this modulation reflects factors beyond feature integration (such as multiple expectancies), our findings highlight the importance of evaluating the effect of previous congruency repetition type in tasks designed to minimize feature-integration confounds. Second, our findings indicate that many of the tasks purported to minimize feature-integration confounds may not be minimizing confounds as effectively as researchers may presume. Hence, caution should be used when analyzing and interpreting data from tasks using standard confound-minimized designs.

Why were the prime-probe tasks used by Schmidt and Weissman (2014) less susceptible to repetition-priming confounds than the congruency tasks used by Aschenbrenner and Balota (2017)? One salient difference between the tasks is that the tasks used by Aschenbrenner and Balota required participants to respond with one hand, whereas the tasks used by Schmidt and Weissman required participants to switch hands between each trial. Multiple studies have demonstrated that response mode differences (i.e., uni-manual vs. bi-manual responding) in confound-minimized congruency tasks can impact the occurrence of the CSE (e.g., Kim & Cho, 2014; Lim & Cho, 2018). Thus, it is possible that differences in response mode may have driven the differences in performance observed in the current study. Future research should therefore explore this possibility by directly comparing performance on uni-manual and bi-manual versions of the same four-alternative forced-choice congruency tasks.

7. General discussion

In order to identify the cognitive and neural mechanisms underlying the CSE, researchers have sought to develop confound-minimized versions of congruency tasks that limit the contribution of feature-integration effects and contingency-learning effects to performance. Research with these tasks has revealed substantial CSEs, indicating that the effect is the product of repetition-expectancy effects, conflict-monitoring effects, or some combination of the two. In light of previous research by Jiménez and Méndez (2014) indicating that the CSE observed in confound-minimized tasks was restricted to trials preceded by a congruency-repetition trial, we developed an extension of the repetition-expectancy account called the multiple-expectancies account. The multiple-expectancies account proposes (a) that multiple expectations regarding the qualities of an upcoming trial can be formed simultaneously, and (b) that these expectations can be consistent or inconsistent with one another.

To test the multiple-expectancies account, we reanalyzed data from two experiments by Schmidt and Weissman (2014) featuring confound-minimized versions of the prime-probe task and two experiments by Aschenbrenner and Balota (2017) featuring confound-minimized versions of the flanker, Simon, and Stroop tasks. Consistent with the results of Jiménez and Méndez (2014), we found that the majority of the datasets revealed a significant CSE in trials preceded by a congruency-repetition trial (i.e., ccC, iiC, ccI, iiI trials) but not in trials preceded by a congruency-alternation trial (i.e., icC, ciC, icI, ciI trials). Although these results initially appeared to support the multiple-expectancies account, it was brought to our attention that the results may have reflected lingering feature-integration confounds in which the stimulus and response features from trial n – 2 overlapped with the stimulus and response features of trial n.

To evaluate whether the confound-minimized tasks used in the current study did exhibit evidence of lingering repetition-priming confounds (a particular subset of the feature-integration confounds potentially occurring in the tasks), we analyzed the effect of response repetition type (i.e., whether the response of trial n repeated that of trial n – 2) on the CSEs observed in trials preceded by a congruency-repetition trial. The results of these exploratory analyses did not reveal evidence of remaining repetition-priming confounds in the prime-probe tasks, lending credence to the multiple-expectancies interpretation of performance. It is important to reiterate, however, that feature-integration confounds stemming from partial-overlap trials may have remained in the prime-probe tasks, potentially undermining the multiple-expectancies interpretation. In contrast to the prime-probe tasks, evidence of remaining repetition-priming confounds was observed in many of the tasks used by Aschenbrenner and Balota (2017).

Given that the predictions of the multiple-expectancies account are predicated on the assumption that feature-integration confounds are sufficiently minimized in the tasks, it is unclear whether the multiple-expectancies account applies to the CSEs observed in Aschenbrenner and Balota's (2017) data. Although further research is needed to directly address this question, our reanalysis of their data demonstrates that additional caution is required when analyzing data from tasks purported to minimize feature-integration confounds, as the confounds may be more widespread than is commonly acknowledged. Further, our results indicate that lingering confounds can be particularly difficult to identify given that feature-integration effects can be obscured by expectancy-based effects, conflict-monitoring effects, or a combination of the two.

The results of the prime-probe tasks appear to be incompatible with standard interpretations of the repetition-expectancy account and the conflict-monitoring account. Both accounts were originally developed to explain how performance on one trial can be modulated by the congruency of the preceding trial. Although these accounts certainly allow for the possibility that the CSE would be more robust on trials preceded by a congruency-repetition trial as opposed to a congruency-alternation trial, neither account predicts that the CSE would be absent entirely following a congruency-alternation trial, assuming that feature-integration confounds have been sufficiently minimized. Yet, both prime-probe datasets failed to reveal any evidence of a CSE on trials preceded by a congruency-alternation trial. It therefore appears that these accounts would require additional modification to accommodate the results of the current study. Thus, the results of the current study contribute to a growing body of findings from confound-minimized prime-probe tasks that call into question the conflict-monitoring account of the CSE (e.g., Weissman, Colter, Grant, & Bissett, 2017; Weissman, Egner, Hawks, & Link, 2015; for a discussion, see Schmidt, 2018).

It could be argued that although the CSE observed in confound-minimized tasks is not generally driven by single congruency repetitions, the mechanisms featured in the repetition-expectancy account or the conflict-monitoring account are nevertheless underlying the CSE albeit on a slightly longer timescale. For example, one could argue that participants only generate expectations for a congruency repetition following two or more trials of the same congruency. Similarly, one could argue that participants only recruit additional top-down resources following two or more incongruent trials. Although we view these possibilities as plausible alternatives to the multiple-expectancies account, it is currently unclear what mechanisms would underlie these more selective versions of the repetition-expectancy account and conflict-monitoring account. We would also like to emphasize that these versions would reflect an important departure from the original formulations and standard interpretations of the accounts.

It should be noted that alternative versions of the multiple-expectancies account can be formulated to account for the results of the current study. For instance, the observed prime-probe results could be explained by appealing to two separate congruency-repetition expectations in which participants form expectations about the congruency of trial n on the basis of trial n – 1's congruency and trial n – 2's congruency (as opposed to trial n – 1's congruency repetition type). Assuming that the strength of each of these congruency-type expectations is roughly equal, a CSE should be observed on trials preceded by a congruency-repetition trial but not on trials preceded by a congruency-alternation trial. Although one might assume that expectations based on the congruency of trial n – 1 would be stronger than those based on the congruency of trial n – 2, certain aspects of the task may serve to equate the strength of the expectations. For instance, trial n and trial n – 2 in the prime-probe tasks were similar in terms of which spatial dimension was relevant, even if the trials were separated by a longer period of time than trial n and trial n – 1.

7.1. Relation to 2AFC tasks

An important question raised by the current study concerns the extent to which performance on standard 2AFC congruency tasks is structured by the congruency repetition type of the preceding trial. Unfortunately, the presence of potential feature-integration confounds complicates the interpretation of performance on such tasks. This is because feature-integration effects can be expected to occur on trials preceded by a congruency-repetition trial as well as those preceded by a congruency-alternation trial. Consequently, if a CSE were to be observed on trials preceded by a congruency-alternation trial, it would be unclear if the observed effect stemmed from repetition expectancy, conflict monitoring, or feature integration.

To investigate whether feature-integration confounds do in fact complicate the interpretation of 2AFC congruency tasks in this manner, we reanalyzed behavioral data from a flanker task collected by Gründler et al. (2009). As anticipated, significant CSEs were observed regardless of the repetition type of the previous trial, with the size of the CSE on trials preceded by a congruency-repetition trial comparable to the CSE on trials preceded by a congruency-alternation trial. The results of this analysis are presented in Section 2 of the Supplementary materials, along with further description of the sample and task.

7.2. *Cumulative effect of multiple congruency repetitions?*

The congruency tasks used in the current study did not constrain the number of congruency repetitions that could occur. Consequently, trials preceded by a congruency-repetition trial may have been preceded by multiple other congruency-repetition trials. This raises the question of whether the effect of the previous trial's congruency repetition type observed in many of the datasets evaluated in the current study was built up over the course of many congruency-repetition trials. To evaluate this possibility, we performed additional analyses in which trials preceded by more than two repetitions of the same congruency were excluded.

As we discuss in further detail in Section 3 of the Supplementary materials, the results of these analyses were highly consistent with the results presented in Sections 3.1–3.3. The most notable difference between the two sets of analyses was that the CSE observed in trials following congruency-repetition trials in the Stroop task from Experiment 1 of Aschenbrenner and Balota (2017) was no longer significant after the number of previous congruency-repetition trials was constrained. Thus, the effect of the previous congruency repetition type observed in the majority of datasets evaluated in the current study does not appear to have been driven by long strings of congruency-repetition trials.

7.3. *Expectations or preparedness?*

Given that the theoretical account on offer in the current study developed out of the repetition-expectancy account, we adopted the usage of terms such as “expectancy” and “expectation”. In some ways, these terms can be misleading as they may cause readers to interpret the account as proposing that participants are generating explicit predictions or proactively forming expectations. The repetition-expectancy account (and derivatives thereof) need not be interpreted in this manner, however.

For instance, one could describe the account in terms of a participant's degree of preparedness for an upcoming trial type: if the congruency of the upcoming trial happens to match the congruency of the recently completed trial, participants will be better prepared because the level of focus adopted on the recent trial will be appropriate for the upcoming trial. Similarly, from the view of the multiple-expectancies account, if the recently completed trial featured a congruency repetition, participants will be better prepared for the upcoming trial to require the same level of focus. Thus, although we retained the expectancy framing in the current study, we do not endorse the position that participants are generating expectations in a controlled or explicit manner. Future research could attempt to compare the proactive expectations view and the preparedness view by, for example, manipulating the frequency with which congruency repetitions occur (e.g., as in Jiménez & Méndez, 2013).

The notion that participants may be prepared for particular events is reminiscent of accounts proposing that the CSE results from attentional settings being carried over passively from one trial to the next (e.g., Hubbard, Kuhns, Schäfer, & Mayr, 2017). One possible point of separation between the preparedness interpretation of the multiple-expectancies account and the passive-carryover account concerns performance on trials preceded by a congruency-alternation trial. According to the preparedness view, participants are primed to change their level of focus if the previous trial also required a change in focus level. Insofar as this preparedness for a change in focus level can be construed to reflect a passive carryover of attentional settings, the two interpretations can be considered consistent.

8. Conclusion

The results of the current study present strong evidence that the congruency repetition type of the previous trial fundamentally modulates the presence or absence of the CSE across a range of congruency tasks designed to minimize feature-integration and contingency-learning confounds. Results from the flanker, Simon, and Stroop tasks indicate that the effect of the previous trial's congruency repetition type was at least partially driven by lingering repetition-priming confounds, suggesting that feature-integration confounds may have remained in previous research using similar versions of the tasks. Crucially, the two versions of the prime-probe task evaluated in the current study did not reveal evidence of lingering repetition-priming confounds. To the extent that these versions of the task were free from other potential feature-integration effects, the results of the current reanalysis pose an important challenge to traditional interpretations of the repetition-expectancy account and the conflict-monitoring account. In place of these accounts, we have presented the multiple-expectancies account, which proposes that participants form multiple – and potentially inconsistent – expectations regarding the qualities of an upcoming trial. Although we anticipate and welcome debate concerning the relative merits of this account, the results of the current study present a clear and novel constraint on the theoretical landscape that has emerged around the CSE.

Author contributions

C. D. Erb wrote the majority of the introduction and discussion sections with critical feedback and revisions from A. J. Aschenbrenner. A. J. Aschenbrenner wrote the majority of the methods and results sections and performed the statistical analyses.

Acknowledgements

We would like to thank Dr. James Schmidt and Dr. Daniel Weissman for making the data from their 2014 study publicly accessible. We would also like to thank Dr. James Cavanagh for providing the behavioral data for our reanalysis of the study by Gründler et al. (2009).

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.actpsy.2019.102869.

References

Aschenbrenner, A. J., & Balota, D. A. (2017). Dynamic adjustments of attentional control in healthy aging. *Psychology and Aging, 32*(1), 1–15. https://doi.org/10.1037/pag0000148.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., … Rcpp, L. (2015). Package ‘lme4’. *Convergence, 12*(1).

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.

Braem, S., Abrahamse, E. L., Duthoo, W., & Notebaert, W. (2014). What determines the specificity of conflict adaptation? A review, critical analysis, and proposed synthesis. *Frontiers in Psychology, 5*(1134), https://doi.org/10.3389/fpsyg.2014.01134.

Duthoo, W., Abrahamse, E. L., Braem, S., Boehler, C. N., & Notebaert, W. (2014). The heterogeneous world of congruency sequence effects: An update. *Frontiers in Psychology, 5*(1001), https://doi.org/10.3389/fpsyg.2014.01001.

Duthoo, W., Abrahamse, E. L., Braem, S., & Notebaert, W. (2014). Going, going, gone? Proactive control prevents the congruency sequence effect from rapid decay. *Psychological Research, 78*(4), 483–493. https://doi.org/10.1007/s00426-013-0498-4.

Duthoo, W., Wühr, P., & Notebaert, W. (2013). The hot-hand fallacy in cognitive control: Repetition expectancy modulates the congruency sequence effect. *Psychonomic Bulletin & Review, 20*(4), 798–805. https://doi.org/10.3758/s13423-013-0390-7.

Egner, T. (2007). Congruency sequence effects and cognitive control. *Cognitive, Affective, & Behavioral Neuroscience, 7*(4), 380–390. https://doi.org/10.3758/CABN.7.4.380.

Egner, T. (2017). Conflict adaptation: Past, present, and future of the congruency sequence effect as an index of cognitive control. In T. Egner (Ed.). *The Wiley handbook of cognitive control* (pp. 64–78). Oxford: Wiley-Blackwell.

Egner, T., Ely, S., & Grinband, J. (2010). Going, going, gone: Characterizing the time-course of congruency sequence effects. *Frontiers in Psychology, 1*(154), https://doi.org/10.3389/fpsyg.2010.00154.

Erb, C. D., & Marcovitch, S. (2018a). Deconstructing the Gratton effect: Targeting dissociable trial sequence effects in children, pre-adolescents, and adults. *Cognition, 179*, 150–162. https://doi.org/10.1016/j.cognition.2018.06.007.

Erb, C. D., & Marcovitch, S. (2018b). Tracking the within-trial, cross-trial, and developmental dynamics of cognitive control: Evidence from the Simon task. *Child Development*. https://doi.org/10.1111/cdev.13111.

Erb, C. D., Moher, J., Sobel, D. M., & Song, J.-H. (2016). Reach tracking reveals dissociable processes underlying cognitive control. *Cognition, 152*, 114–126. https://doi.org/10.1016/j.cognition.2016.03.015.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*(1), 143–149. https://doi.org/10.3758/BF03203267.

Fox, J., & Weisberg, S. (2011). *Multivariate linear models in R. An R companion to applied regression.* Los Angeles: Thousand Oaks.

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General, 137*(2), 201–225. https://doi.org/10.1037/0096-3445.137.2.201.

Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General, 121*(4), 480–506. https://doi.org/10.1037/0096-3445.121.4.480.

Gründler, T. O., Cavanagh, J. F., Figueroa, C. M., Frank, M. J., & Allen, J. J. (2009). Task-related dissociation in ERN amplitude as a function of obsessive–compulsive symptoms. *Neuropsychologia, 47*(8–9), 1978–1987. https://doi.org/10.1016/j.neuropsychologia.2009.03.010.

Hazeltine, E., Lightman, E., Schwarb, H., & Schumacher, E. H. (2011). The boundaries of sequential modulations: Evidence for set-level control. *Journal of Experimental Psychology: Human Perception and Performance, 37*(6), 1898–1914. https://doi.org/10.1037/a0024662.

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences, 8*(11), 494–500. https://doi.org/10.1016/j.tics.2004.08.007.

Hommel, B., Proctor, R. W., & Vu, K. P. L. (2004). A feature-integration account of sequential effects in the Simon task. *Psychological Research, 68*(1), 1–17. https://doi.org/10.1007/s00426-003-0132-y.

Hubbard, J., Kuhns, D., Schäfer, T. A., & Mayr, U. (2017). Is conflict adaptation due to active regulation or passive carry-over? Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(3), 385–393. https://doi.org/10.1037/xlm0000306.

Jiménez, L., & Méndez, A. (2013). It is not what you expect: Dissociating conflict adaptation from expectancies in a Stroop task. *Journal of Experimental Psychology: Human Perception and Performance, 39*(1), 271–284. https://doi.org/10.1037/a0027734.

Jiménez, L., & Méndez, A. (2014). Even with time, conflict adaptation is not made of expectancies. *Frontiers in Psychology, 5*(1042), https://doi.org/10.3389/fpsyg.2014.01042.

Kim, S., & Cho, Y. S. (2014). Congruency sequence effect without feature integration and contingency learning. *Acta Psychologica, 149*, 60–68. https://doi.org/10.1016/j.actpsy.2014.03.004.

Lim, C. E., & Cho, Y. S. (2018). Determining the scope of control underlying the congruency sequence effect: Roles of stimulus-response mapping and response mode. *Acta Psychologica, 190*, 267–276. https://doi.org/10.1016/j.actpsy.2018.08.012.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*(2), 163–203. https://doi.org/10.1037/0033-2909.109.2.163.

Mayr, U., Awh, E., & Laurey, P. (2003). Conflict adaptation effects in the absence of executive control. *Nature Neuroscience, 6*(5), 450–452. https://doi.org/10.1038/nn1051.

Mullane, J. C., Corkum, P. V., Klein, R. M., & McLaughlin, E. (2009). Interference control in children with and without ADHD: A systematic review of Flanker and Simon task performance. *Child Neuropsychology, 15*(4), 321–342. https://doi.org/10.1080/09297040802348028.

Ridderinkhof, K. R., Van Den Wildenberg, W. P., Segalowitz, S. J., & Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and Cognition, 56*(2), 129–140. https://doi.org/10.1016/j.bandc.2004.09.016.

Schmidt, J. R. (2018). Evidence against conflict monitoring and adaptation: An updated review. *Psychonomic Bulletin & Review,* 1–19. https://doi.org/10.3758/s13423-018-1520-z.

Schmidt, J. R., & De Houwer, J. (2011). Now you see it, now you don't: Controlling for contingencies and stimulus repetitions eliminates the Gratton effect. *Acta Psychologica, 138*(1), 176–186. https://doi.org/10.1016/j.actpsy.2011.06.002.

Schmidt, J. R., & Weissman, D. H. (2014). Congruency sequence effects without feature integration or contingency learning confounds. *PLoS One, 9*(7), e102337. https://doi.org/10.1371/journal.pone.0102337.

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology, 81*(1), 174. https://doi.org/10.1037/h0027448.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643–662. https://doi.org/10.1037/h0054651.

Ullsperger, M., Bylsma, L. M., & Botvinick, M. M. (2005). The conflict adaptation effect: It's not just priming. *Cognitive, Affective, & Behavioral Neuroscience, 5*(4), 467–472. https://doi.org/10.3758/CABN.5.4.467.

Waszak, F., Li, S. C., & Hommel, B. (2010). The development of attentional networks: Cross-sectional findings from a life span sample. *Developmental Psychology, 46*(2), 337–349. https://doi.org/10.1037/a0018541.

Weissman, D. H., Colter, K. M., Grant, L. D., & Bissett, P. G. (2017). Identifying stimuli that cue multiple responses triggers the congruency sequence effect independent of response conflict. *Journal of Experimental Psychology: Human Perception and Performance, 43*(4), 677. https://doi.org/10.1037/xhp0000350.

Weissman, D. H., Egner, T., Hawks, Z., & Link, J. (2015). The congruency sequence effect emerges when the distracter precedes the target. *Acta Psychologica, 156*, 8–21. https://doi.org/10.1016/j.actpsy.2015.01.003.

Wylie, S. A., van den Wildenberg, W. P. M., Ridderinkhof, K. R., Bashore, T. R., Powell, V. D., Manning, C. A., & Wooten, G. F. (2009). The effect of Parkinson's disease on interference control during action selection. *Neuropsychologia, 47*(1), 145–157. https://doi.org/10.1016/j.neuropsychologia.2008.08.001.