Assessing the Temporal Learning Account of the List-Wide Proportion Congruence Effect

Emily R. Cohen-Shikora

Jihyun Suh

Julie M. Bugg

Washington University in St. Louis

Author Note

Emily R. Cohen-Shikora, Jihyun Suh, and Julie M. Bugg, Department of Psychological and Brain Sciences, Washington University in St. Louis.

Correspondence concerning this article should be sent to Emily R. Cohen-Shikora, Department of Psychological and Brain Sciences, Campus Box 1125, Washington University in St. Louis, St. Louis, MO 63130, ecohensh@wustl.edu.

We would like to thank Joshua Jackson, Michael Strube, the Cognitive Control and Aging Lab, and David Balota and the Cognitive Psychology Lab, for their helpful feedback on this project, and Jackson Colvett for his thoughtful comments on an earlier draft. We are also grateful to reviewers Keith Hutchison, Serje Robidoux, and James Schmidt for their insight.

Abstract

In this article, we assess an alternative account of a key experimental pattern thought to index top-down control. The list-wide proportion congruence effect is the well-documented pattern whereby the congruency effect (i.e., Stroop effect) is attenuated in lists containing mostly incongruent trials relative to lists containing mostly congruent trials. This pattern has typically been interpreted as a signature of a top-down control mechanism that modulates attention to the word dimension based on the global probability of encountering conflict between the word and color. However, Schmidt (2013a; 2013b) has proposed an alternative account that stresses relative temporal differences in responding between mostly incongruent and mostly congruent lists rather than relative differences in the control of attention. To assess this temporal learning account, we evaluate the evidence reported by Schmidt (2013a) and report new analyses of three previously published datasets in which a list-wide proportion congruence effect was observed after controlling for other potential confounds. These analyses targeted three key topics: effects of reaction time transformations, statistical support for temporal learning, and measurement of temporal rhythm. The evidence for the temporal learning account was neither strong nor consistent, and there was a highly significant list-wide proportion congruence effect that survived multiple attempts to control for temporal learning. Accordingly, we conclude that the temporal learning account is not currently a robust alternative to the global control account in explaining list-wide proportion congruence effects.

Keywords: cognitive control, attention; temporal learning, Stroop, proportion congruent, list-wide proportion congruence

Assessing the Temporal Learning Account of the List-Wide Proportion Congruence Effect

Goals change on a moment-by-moment basis, with these changes often corresponding to shifts in context. To maximize performance, one must shift attention in a context-sensitive fashion, an ability ascribed to *cognitive control*. Our phenomenological experience of heightening attention when a task demands it, and adapting attention in response to changes in goals or contexts, suggests the existence of a top-down cognitive control mechanism. Researchers have gathered empirical evidence for top-down cognitive control using various experimental paradigms, and in the current paper we focus on a key piece of evidence from the list-wide proportion congruence (list-wide PC) paradigm.

Since the seventies, the list-wide PC manipulation has been extensively employed to index the context-sensitive modulation of attention in tasks such as Stroop color-naming (e.g., Cheesman & Merikle, 1986; Glaser & Glaser, 1982; Kane & Engle, 2003; Lindsay & Jacoby, 1994; Logan, 1980; Logan & Zbrodoff, 1979; Lowe & Mitterer, 1982; Shor, 1975; West & Baylis, 1998). The list-wide PC manipulation varies the relative frequency of congruent (e.g., the word "blue" displayed in blue ink) to incongruent (e.g., the word "red" displayed in blue ink) trials across list contexts (e.g., mostly congruent vs. mostly incongruent lists; for reviews and a user's guide to the manipulation, see Bugg, 2012, 2017; Bugg & Crump, 2012). The *congruency (Stroop) effect* is the highly robust pattern whereby response times are slowed (and error rates are sometimes exacerbated) on incongruent relative to congruent trials. When list-wide PC is manipulated, the list-wide PC effect is observed— mostly incongruent (MI) lists produce smaller congruency effects relative to mostly congruent (MC) lists. That is, the magnitude of the congruency effect is negatively related to the probability of encountering conflict (Blais, Harris, Guerrero, & Bunge, 2012; for computational models, see Blais, Robidoux, Risko, & Besner, 2007; Botvinick, Braver, Barch, Carter, & Cohen, 2001). Accordingly, researchers traditionally have interpreted the list-wide PC effect as reflecting cognitive control such that attention to the word is minimized in MI (high conflict) lists relative to MC (low conflict) lists[1] (e.g., Kane & Engle, 2003; Lindsay & Jacoby, 1994; Melara & Algom, 2003; Logan & Zbrodoff, 1979; Lowe & Mitterer, 1982).

Over the years alternative interpretations of the list-wide PC effect have been proposed. For instance, as will be described below, there was evidence that a contingency learning mechanism (Schmidt & Besner, 2008) could explain the list-wide PC effect (Bugg, Jacoby, & Toth, 2008; Blais & Bunge, 2010). However, interpretations such as this have since been refuted by demonstrations of list-wide PC effects for items equated in presentation frequency and PC (and thereby contingency; see Bugg, 2014; Bugg & Chanani, 2011; Bugg, McDaniel, Scullin, & Braver, 2011; Gonthier, Braver, & Bugg, 2016; Hutchison, 2011, for evidence within a single task; see Funes, Lupiáñez, & Humphreys, 2010; Torres-Quesada, Funes, & Lupiáñez, 2013; Wühr, Duthoo, & Notebaert, 2015, for evidence across tasks). But another alternative account has emerged—Schmidt (2013a; 2013b) recently proposed the temporal learning account, which has the potential to explain list-wide PC effects for such items. The temporal learning account posits that the list-wide PC effect results from temporal differences in responding across lists and not from modulations of attention. The overarching aim of this research article is to evaluate the evidence for this account.

We begin by briefly reviewing the evidence in favor of the global control account of the list-wide PC effect, including prior research that has documented conditions under which extant,

[1] The interpretation that cognitive control acts to minimize attention to the word dimension in an MI context relative to an MC context is supported by findings from the process-dissociation procedure (Lindsay & Jacoby, 1994); however, heightening of color information in the MI context may also occur (see Egner & Hirsch, 2005).

alternative explanations were not viable. Doing so serves to situate the present question in a broader theoretical and historical context and introduces the reader to the three previously published datasets that we re-analyzed for present purposes of evaluating the temporal learning account (Bugg, 2014; Gonthier et al., 2016; Hutchison, 2011). Then we turn to assessing the temporal learning account using a two-pronged approach. The first goal was to replicate the analyses from Schmidt (2013a), based on the Hutchison (2011) dataset, and apply those same analyses to two other datasets from a different lab (Bugg, 2014; Gonthier et al., 2016). The second goal was to evaluate and expand the analyses conducted by Schmidt (2013a) to provide a comprehensive and rigorous test of the temporal learning account.

The Control-Based Account

The list-wide PC effect has played a central role in theories and models because it demonstrates the context sensitivity of cognitive control (e.g., Botvinick et al., 2001; Blais et al., 2007; for reviews, see Bugg, 2012; 2017; Bugg & Crump, 2012). Although the instruction is to name the color of the stimulus regardless of list, the list-wide PC effect demonstrates that attention to the word dimension is attenuated (MI list) or exaggerated (MC list) depending on the context (i.e., Dishon-Berkovits & Algom, 2000; Melara & Algom, 2003). Initial evidence for a control account was provided by early studies (Logan, 1980; Logan & Zbrodoff, 1979; Logan, Zbrodoff, & Williamson, 1984; Lowe & Mitterer, 1982), which tended to stress the strategic nature of the attentional adjustments. For instance, Lowe and Mitterer (1982) stated that the differential deployment of attention across lists demonstrated that “attentional strategies may be actively chosen to suit prevailing conditions” (p. 684). However, a recent study that gauged awareness of the PC manipulation at the end of each list found that the magnitude of the list-wide PC effect was largely unrelated to participants’ awareness of the number of congruent trials

(Blais et al., 2012). These data suggest it is unlikely that a strategic mechanism like that described by Lowe and Mitterer underlies the list-wide PC effect. Rather, the mechanism may be non-strategic and based on implicit knowledge of the global probability of conflict (Blais et al., 2012; see also Bugg & Diede, 2017).

In recent years, theoretical debate has centered on whether the mechanism underlying the list-wide PC effect operates globally (i.e., the global control account) or locally. According to the global control account, attention is modulated at the list level based on the global probability of encountering conflict (e.g., proactively; Braver, Gray, & Burgess, 2007; DePisapia & Braver, 2006; see Botvinick et al., 2001, for a pathway-based computational model; see Melara & Algom, 2003, for the global-correlation based Tectonic Theory). This means that after participants (implicitly) learn the PC of the list, an attentional setting is applied uniformly throughout the list with different settings corresponding to each list (e.g., a setting that attenuates or exaggerates processing of the word in the MI versus MC list, respectively). Local accounts, in contrast, do not attribute the effect to adjustments that are based on the global probability of conflict. One such account, the item-specific control account (Bugg et al., 2008; Blais & Bunge, 2010; see Blais, et al., 2007, for an item-specific computational model of the list-wide PC effect), posits that participants instead learn the probability of encountering conflict for each item (e.g., word; Jacoby, Lindsay, & Hessels, 2003), and modulate attention post-stimulus onset (reactively) based on the PC of the item. Another such account, the contingency account (Schmidt & Besner, 2008; see also Schmidt, Crump, Cheesman, & Besner, 2007), posits that participants learn stimulus-response contingencies for each word and produce the high contingency response post-stimulus onset (i.e., say “red” when RED is encountered in a MC list because the word RED is most frequently paired with the color red in this context). The traditional list-wide PC manipulation cannot adjudicate among the three accounts because it employs only biased items (i.e., items to which the PC manipulation is applied).

To disentangle the mechanisms, recent studies implemented novel variants of the traditional list-wide PC design. The key feature of these variants was inclusion of a distinct set of words (and their associated colors; referred to hereafter as *critical items*[2]) in the MC and MI list that were intermixed with biased items but were matched in frequency and PC (e.g., 50% congruent in both lists) and therefore contingency. These critical items enabled evaluation of the contribution of global control independent of item-specific control and contingency learning (Bugg et al., 2008). The item-specific control and contingency accounts predict equivalent congruency effects for these critical items across the MC and MI lists (i.e., no list-wide PC effect); in contrast, the global control account predicts that a list-wide PC effect should be observed for critical items. That is, if attention is modulated based on the global probability of conflict, then the modulation should affect all items within a list not just those that are biased. Accordingly, the congruency effect should be smaller for critical items in the MI list than the MC list.

Initial studies found no list-wide PC effect for critical items in line with local accounts (Bugg et al., 2008; Blais & Bunge, 2010). Importantly, however, subsequent research identified a crucial factor that seemed to preclude engagement of the global control mechanism in those studies, namely use of a small set of biased items that enabled participants to minimize interference for most items in the list by relying on stimulus-response associations (Bugg & Chanani, 2011). Consistent with the Associations as Antagonists to Top Down Control account (Bugg, 2014), when the size of the biased item set was increased (from two to four items) such

[2] Some studies refer to such items as "transfer" items and/or refer to biased items as "inducer" items.

that participants could not reliably predict responses on incongruent trials, evidence for global control was consistently observed—the Stroop effect was smaller for critical items in the MI list compared to the MC list (Bugg, 2014). There are now multiple reports from different labs of list-wide PC effects for critical items, supporting the global control account (Bugg & Chanani, 2011; Gonthier et al., 2016; Hutchison, 2011; see also Bugg et al., 2011; Funes et al., 2010; Torres-Quesada et al., 2013; Wühr et al., 2015). We re-analyzed three such datasets (Bugg, 2014, Experiments 1a & 2b; Gonthier et al., 2016, Experiments 1a & 1b; Hutchison, 2011) for purposes of examining whether a different global mechanism (temporal learning) provides a compelling alternative account of the list-wide PC effect (we use the term *list-wide PC effect* here and hereafter exclusively to refer to a list-wide PC effect for frequency and PC matched [i.e., critical] items that cannot be explained by item-specific control or contingency learning).

The Temporal Learning Account

Currently, the list-wide PC literature is grappling with another account that posits an alternative mechanism that is purported to operate globally, and thus may be able to explain the list-wide PC effect. The account is referred to as temporal learning (Schmidt, 2013a; 2013b). In contrast to the global control account, it does not posit a role for context-sensitive modulations of attention in the list-wide PC effect. Instead, it recasts the list-wide PC effect as an effect of participants learning and then adopting global response rhythms (i.e., retrieving expectancies about when to respond) in each list. According to Schmidt (2013a; 2013b), the MC list produces a globally fast rhythm of responding that selectively benefits responses on fast (i.e., congruent) trials in that context, producing a large congruency effect (incongruent – congruent RT). In contrast, the MI list produces a globally slow rhythm of responding that selectively benefits responses on slow (i.e., incongruent) trials in that context, producing a small congruency effect.[3] For a visual depiction of this account, see Figure 1.

A major source of evidence for the temporal learning account stems from Schmidt (2013a), who used RT on trial *n* - 1 as a proxy for participants learning a temporal deadline of when to respond in a particular list. This was based on his findings of a main effect of prior RT on current RT and a significant (negative) prior RT by congruency interaction (the faster the prior RT, the larger the congruency effect) using linear mixed-effects modeling. Most importantly for present purposes, Schmidt assessed the congruency by PC interaction (the statistical manifestation of the list-wide PC effect); if temporal learning fully explains the list-wide PC effect, this interaction should become non-significant because once temporal learning is accounted for, the purported signature of control should no longer be observed. However, Schmidt did not find that the congruency by PC interaction was eliminated. Instead he found only that the interaction coefficient declined (from $\beta = .059$ without indices of temporal learning to $\beta = .051$ with indices of temporal learning), a pattern he assumed to reflect use of a poor and noisy measure of temporal learning (i.e., prior RT), an assumption that we consider in *Section 3*.

Schmidt (2013a) concluded that the critical contribution of his study was the "cautionary demonstration that the list-level PC effect cannot be taken as strong evidence for conflict adaptation without further controls." (p. e82320). The implication is that temporal learning represents a serious alternative to the global control account of the list-wide PC effect, and in so much as researchers want to make claims based on global control (i.e., global [pathway-based]

[3] However, in the same report, Schmidt (2013a) found that the effect of the list-wide PC manipulation was primarily in the incongruent trials, and noted that his initial explanation was "a bit oversimplified…The reason for a larger effect for incongruent relative to congruent trials in both the modelling and participant data probably has to do with the fact that temporal learning has more time to affect processing on incongruent trials." (p. e82320). Thus, it is unclear whether the temporal learning account predicts movement in both congruent and incongruent trials or an effect primarily, or more strongly, on the incongruent trials.

conflict adaptation; Botvinick et al., 2001), they must incorporate additional experimental controls to account for temporal learning. As reviewed above, this is not unfamiliar territory for researchers investigating the list-wide PC effect (or cognitive control more generally; see e.g., Awh, Belopolsky, & Theeuwes, 2012). Indeed, we and others have advocated for design changes that enable researchers to isolate global control from other mechanisms that could masquerade as global control (e.g., Blais & Bunge, 2010; Blais et al., 2007; Bugg, 2014; Bugg, 2017; Bugg & Chanani, 2011; Bugg et al., 2008; Hutchison, 2011). However, these recommendations were supported by strong and consistent evidence demonstrating that mechanisms such as item-specific control and contingency learning could indeed produce the list-wide PC effect independent of global control (e.g., Blais & Bunge, 2010; Blais et al., 2007; Bugg et al., 2008). It is not yet clear if the evidence for temporal learning is similarly strong or consistent—examining this question is of critical importance should the temporal learning account be considered a serious alternative to the global control account and guide future research on the list-wide PC effect (e.g., design, choice of "necessary" controls, etc.). This article addresses this question, and thereby fills a critical gap in the literature.

Our specific goals were twofold: first, we set out to examine the temporal learning account using multiple datasets in which the list-wide PC effect was observed and attributed to a global control mechanism. These datasets were collected in different labs using different experimental designs but had in common the inclusion of critical items within a list-wide PC paradigm. In service of this goal, we replicated the analyses from Schmidt (2013a) based on the Hutchison (2011) dataset, and we applied those same analyses to two other datasets (Bugg, 2014; Gonthier et al., 2016). Our second goal was evaluating and expanding the analyses conducted by Schmidt to provide a more comprehensive and rigorous test of the temporal learning account.

Our general approach is detailed below, and then specific analyses and interpretation follow in three separate sections addressing: (1) Effects of RT transformations, (2) Evaluation of the change in the critical congruency by PC interaction, and (3) Measurement of the construct of temporal rhythm. We conclude with a section outlining implications for the study of list-wide PC effects, including recommendations for future research.

General Approach

We used three datasets for our analyses[4]. The first was Hutchison (2011), which Schmidt (2013a) analyzed previously. In the study by Hutchison, 226 participants[5] completed 180 trials each in either an MC or one of two MI list contexts (single versus mixed filler items; collapsed for the purposes of these analyses). Critical items (N = 120 per participant) were the same across both lists and consisted of a set of MC items (67% congruent) and two sets of MI items (high- and low-contingency, both 33% congruent). The second dataset we analyzed was Experiments 1a and 2b combined from Bugg (2014), in which a total of 72 participants completed 320 trials each. She too held constant the PC of critical items (50% congruent in her study, N = 96 trials per participant) and compared the congruency effect between groups for whom list-wide PC differed (MC or MI). The third dataset was Experiments 1a and 1b combined from Gonthier et al. (2016), in which a total of 93 participants completed a within-groups PC design, including 384 trials in an MC list and 384 trials in an MI list. As in Bugg, the PC of critical items was 50% congruent (N = 96 trials in each PC list context per participant). In all three studies, there was a significant congruency by PC interaction for critical items (i.e., list-wide PC effect). These three studies constitute a substantial portion of the evidence for the global control account.

---

[4] Data from Hutchison (2011) are available upon request to the author; data from Bugg (2014) and Gonthier et al. (2016) are on the Open Science Framework site: https://osf.io/z2rmw and https://osf.io/b9zyv/, respectively.

[5] These are the same participants analyzed by Schmidt (2013a), who reported N = 230. This discrepancy is due to a reporting error in his original paper (J. Schmidt, personal communication, August 4, 2017).

As a preliminary step, RT data for each of the three datasets were treated as in Schmidt's (2013a) re-analysis of Hutchison (2011). First, current and prior RTs were inverse transformed (-1000/RT) for normalization. The inverse transformation helped normalize the RT from all three datasets (see Figure 2 for the q-q plots). Trials on which the current or prior RT was faster than 300ms were eliminated. Also, trials on which the current or prior (n – 1) response was errant were eliminated, as were filler (non-critical, or inducer) trials. Finally, trials on which prior color or word matched the current color or word were excluded. Congruency and PC were dummy coded. For congruency, congruent trials were coded as 0 and incongruent trials were coded as 1. For PC, trials from MI lists were coded as 0 and trials from MC lists were coded as 1. Prior RT was centered with the grand mean to avoid a spurious correlation with current RT.

Following Schmidt (2013a), we conducted linear mixed-effects modeling using the Lme4 package (Bates, Mächler, Bolker, & Walker, 2014) in R software. For the primary analyses of each dataset, two models were tested. In both, subject number and color-word combinations (as well as experiment, for Bugg, 2014, and Gonthier et al., 2016) were included as random effect factors. Congruency, PC, and prior trial RT were included as fixed effect factors. First, the *simple model* tested the effects of congruency, PC, and the interaction of congruency and PC (RT ~ congruency + PC + congruency:PC + (1|subject) + (1|color-word combinations)). The simple model is considered a baseline for the subsequent model. Second, the *temporal learning model* was examined where the temporal learning indices were added to the simple model (RT ~ congruency + PC + congruency:PC + prior RT + priorRT:congruency + (1|subject) + (1|color-word combinations)). The main interest was the change in the congruency by PC interaction coefficient (β) from the simple model to the temporal learning model. Specifically, an elimination of the interaction coefficient would constitute strong evidence that temporal learning, and not global control, is responsible for the list-wide PC effect. At present, only a reduction in the interaction coefficient has been observed in prior work (Schmidt, 2013a).

Replication of Schmidt's (2013a) Primary Analyses

As shown in the "transformed RT" column in Table 1, the simple model showed a significant interaction between congruency and PC for all three datasets, β = .057[6] for Hutchison (2011), β = .099 for Bugg (2014), and β = .064 for Gonthier et al. (2016). In the temporal learning model, the indices of temporal learning (prior RT and the prior RT by congruency interaction) were significant, and the beta corresponding to the congruency by PC interaction decreased to β = .049 for Hutchison, β = .084 for Bugg, and β = .058 for Gonthier et al., reductions of 0.008, 0.015, and 0.006, respectively. Replicating Schmidt (2013a), in all datasets the reductions were small[7], and critically, the congruency by PC interaction remained significant in the temporal learning model. These patterns motivated the analyses reported in the next three sections that provide a more comprehensive and rigorous test of the temporal learning account.

Section 1: RT Transformations[8]

Reaction time is one of the most commonly-used metrics of cognition, yet it also has the notoriously challenging statistical property of strong positive skew. One common way researchers have handled this skew, particularly within the linear mixed-effects modeling approach, is to perform an RT transformation (e.g., inverse RT: -1000/RT) that makes it more normal. The primary evidence that Schmidt (2013a) cited as support for the temporal learning

[6] This beta value, and the beta corresponding to this interaction within the temporal learning model, are both slightly different from the betas Schmidt (2013a) reported (although our values were within .002 of his). This may be due to minor differences in rounding or models between his SPSS analyses and our R analyses.
[7] The tables in the current manuscript include 3 decimal places so that the changes are apparent when visually inspecting the results.
[8] We also examined the role of RT filtering because it varies across studies. Using multiple different RT cutoffs yielded the same overall patterns of small and inconsistent beta weight change in the congruency by PC interaction.

account (i.e., the decrease in the congruency by PC interaction with the addition of indices of temporal learning) was taken from analyses on transformed RT (as reported in the *Replication of Schmidt's [2013] Primary Analyses* section above). Although the transformation helps normalize RT (including in the three present datasets, see Figure 2 for the q-q plots), researchers have called attention to the problems associated with this transformation (Balota, Aschenbrenner, & Yap, 2013, Lo & Andrews, 2015). For example, the transformation alone can change the nature of observed interactions, making them smaller, eliminating them, or even reversing them (Balota et al., 2013). Furthermore, these transformations change the nature of the variable being explored; what was an analysis of raw RT (response time, what researchers are typically formulating predictions about, as in the case of Schmidt's (2013a) predictions) becomes an analysis of response *rate* (a different DV) once transformed to inverse RT. As Lo and Andrews (2015) and Robidoux (2017) pointed out, researchers should take this into account when justifying a transformation that is appropriate for their predictions.

As an initial step toward exploring the potential influence of the RT transformation on the results reported by Schmidt (2013a) and the corresponding results from the other two datasets, we performed the same "replication" analyses reported above but instead used untransformed (raw) RT (see the "untransformed RT" column in Table 1). As in Schmidt (2013a) and the replication analyses performed on transformed RT, prior RT was significant for all datasets. However, for untransformed RT, the prior RT by congruency interaction was not significant for Bugg (2014), and it was in the opposite direction than expected per the temporal learning account for Hutchison (2011) and Gonthier et al. (2016). The temporal learning account predicts a *negative* interaction, indicating that the congruency effect is larger the faster the prior RT, but we found a positive interaction in both cases. Most importantly for present purposes, for two of the three datasets, the results for the critical congruency by PC interaction in untransformed RT were inconsistent with the results in transformed RT: for Hutchison (2011) and Gonthier et al. (2016), the beta coefficient for the interaction in untransformed RT *increased* (by 3.665 and 2.769, respectively)[9] from the simple model to the temporal learning model (counter to the temporal learning account). Only in the Bugg (2014) dataset were the results in untransformed RT consistent with the transformed RT analyses, in that the congruency by PC interaction beta for untransformed RT decreased (by 1.125) in the temporal learning model. The increase in the interaction beta for two of the three datasets in untransformed RT contrasts with the decreasing beta observed for all three in transformed RT.

This inconsistency highlights the interpretational challenges of relying either on transformed RT data, or on untransformed RT data that are skewed. It is arguably more appropriate to utilize generalized linear mixed effects modeling (GLMM) when dealing with skewed RT data, as advocated by Lo and Andrews (2015). They acknowledged the challenge researchers face in wanting to formulate hypotheses based on untransformed RT (cf. Robidoux, 2017) but having to use transformed RT out of mathematical necessity, and they demonstrated that GLMM offers a reasonable alternative. GLMM is similar to linear mixed effects modeling, except it allows the researcher to specify the distribution of RTs, thus avoiding the pitfalls of violating the assumption of normality by using positively skewed untransformed RT or relying on potentially misleading transformed RT.

We adopted the GLMM approach to analyzing untransformed RTs for present purposes of evaluating the temporal learning account. Following Lo and Andrews (2015), we specified the

[9] Note that these are raw RTs, so although they appear to be much larger changes than the changes observed for transformed RT, they are still quite small considering the raw RT scale. The one dataset (Bugg, 2014) with results supporting the temporal learning account for both data types actually showed the smallest change.

RT distribution as a gamma distribution because they showed it is a good fit for untransformed RTs and the link function as a one-to-one, identity link (i.e., direct link from the operationalization to the theoretical variable). In addition, we used a random intercept model to keep the same random effect structure as Schmidt (2013a)[10]. The residuals plot of this model showed random scatter instead of any systematic relations, indicating that this was an appropriate choice of RT distribution (see Figure 3). For each of the three datasets, we contrasted the simple model to the temporal learning model to see if the critical congruency by PC interaction decreased between the two models (see Table 2). The change from the simple model to the temporal learning model again was quite small and inconsistent, and the interaction remained significant in the temporal learning model across all three datasets. For the Bugg (2014) dataset, the interaction beta decreased (by 1.271). However, for the Hutchison (2011) and Gonthier et al. (2016) datasets, the interaction beta increased (by 2.673 and 2.481, respectively). Thus, we do not see a coherent pattern of influence of temporal learning despite using a technique that is arguably superior to transforming data in handling skew. It also merits mention that the congruency by prior RT interaction (a signature of temporal learning; Schmidt, 2013a) was not significant in the GLMM analyses for the Bugg (2014) dataset, $p > .05$, and it was again in the opposite direction as predicted by the temporal learning account for Hutchison (2011) and Gonthier et al. (2016). Collectively, the analyses within this section converge in demonstrating that the inverse RT transformation itself influenced some of the critical effects that were cited as evidence for temporal learning (Schmidt, 2013a; 2013b). In contrast, the signature of global control (i.e., congruency by PC interaction) remained significant in the temporal learning model for all three datasets regardless of the type of analysis performed (linear mixed-effects modeling

[10] We did attempt a full random slope and intercept model, but the GLMM model often failed to converge (when it did converge, results were similar to the random intercept model).

or GLMM) or type of data examined (transformed or untransformed RT) demonstrating that some effects are robust to of the influence of the transformation. In all analyses hereafter, we apply the GLMM approach on untransformed data (which we refer to in shorthand as "untransformed RTs") in concert with the linear mixed effect modeling approach on transformed data ("transformed RTs").

Section 2: Evaluating change in the critical congruency by PC interaction

Central to addressing the question of whether the list-wide PC effect reflects temporal learning and not global control is the change in the beta estimate for the critical congruency by PC interaction from the simple model to the temporal learning model. To date, appraisal of this change as being sufficient to entertain an interpretation based on temporal learning has been founded merely on visual inspection of beta weights (i.e., presence of a numerical reduction in the beta from the simple to the temporal learning model in the Hutchison, 2011, dataset; Schmidt, 2013a). However, this general approach is potentially problematic given issues with the practical interpretability of betas, and the fact that visual inspection is inherently vulnerable to subjectivity. In this case the decrease in the beta reported by Schmidt for the congruency by PC interaction appeared to be small (from $\beta = .059$ in the simple model to $\beta = .051$ in the temporal learning model for the Hutchison, 2011, dataset) and the interaction was not eliminated in the temporal learning model, patterns that we replicated in the Bugg (2014) and Gonthier et al. (2016) datasets. A better approach is to test whether the change is statistically significant, or otherwise meaningful (e.g., proportion of variance explained). We recognize that there is possibly no perfect method for doing so because all methods are limited in some way. Thus, we sought converging evidence across four methods with the assumption that a consistent pattern of results could not be attributed to limitations of any single approach (these analyses will be covered in turn, but for an overall summary of each analysis and whether it provided evidence for or against the temporal learning account, see Table 3). The four methods were: (1) considering the percentage of variance explained by the congruency by PC interaction, with and without indices of temporal learning in the statistical model, (2) comparing model fit AIC and BIC statistics for the congruency by PC interaction with and without taking into account indices of temporal learning, (3) examining the three-way interaction of congruency by PC by prior RT, and (4) using analyses to statistically compare the change in R^2 for the congruency by PC interaction with and without taking into account indices of temporal learning.

First, we looked at the percentage of variance explained by the congruency by PC interaction, before and after the addition of the indices of temporal learning into the statistical model (see Table 4) [11]. Although this measure is also vulnerable to subjective interpretation, like beta change, it arguably provides a more meaningful sense of how much the critical interaction changed with the introduction of the indices of temporal learning. A decrease in the percentage of variance explained by the critical interaction would support the temporal learning account. The percentage of variance explained by the congruency by PC interaction sometimes decreased (Bugg, 2014, transformed RT and Gonthier et al., 2016, transformed and untransformed RT). However, in other cases it *increased* from the simple model to the temporal learning model (Hutchison, 2011, transformed and untransformed RT, and Bugg, 2014, untransformed RT),

[11] One might observe upon examination of Table 4 that the magnitude of the congruency by PC interaction, and the percentage of variance explained by that critical interaction, are sometimes small (although always significant). The reader should be reminded that the list-wide PC effect indexed by this interaction is independent of any contribution of other factors (contingencies, item-specific effects, etc.) due to the use of critical items; moreover, the statistical analyses additionally control for other factors (e.g., color or word repetitions) that may further weaken the magnitude of the list-wide PC effect.

counter to the account. Thus, the changes in percentage of variance explained were inconsistent and at least partially incompatible with the temporal learning account.[12]

Next, we took a somewhat different approach to modeling change in the congruency by PC interaction *without* indices of temporal learning (as in the simple model) and *with* indices of temporal learning (as in the temporal learning model) by using AIC and BIC model fit statistics. We considered model fit statistics because they are easy to interpret (reductions in AIC/BIC values indicate better model fit), and reward parsimony by considering the number of parameters in the model. One criticism of the inclusion of prior RT is that adding any extra variable that produces a robust main effect may decrease the betas of other main effects or interactions in the model. The indices of temporal learning (e.g., prior RT) may be just such variables, as they are strongly correlated with current RT.[13] This potential issue is mitigated by using AIC/BIC values because adding variables penalizes the model fit. The specific approach we took was to compare two changes in AIC/BIC values across 4 different models (see Table 5). The first change was from model 1 (PC, congruency) to model 2 (PC, congruency, congruency by PC), to isolate the congruency by PC interaction *without* the influence of indices of temporal learning. The second change was from model 3 (PC, congruency, prior RT, prior RT by congruency) to model 4 (PC, congruency, prior RT, prior RT by congruency, congruency by PC) to isolate the congruency by PC interaction *with* the influence of indices of temporal learning. This approach of incrementally

[12] A potential peculiarity of this analysis is that in several cases, the beta for the critical interaction decreased (including for transformed RT from Hutchison, 2011, and untransformed data from Bugg, 2014), but the percentage of variance explained by that interaction increased. One possible explanation for this is that the minute and seemingly random differences are attributable to noise in the data. The other possibility is that although the beta decreased, the estimate of the interaction got more precise in the temporal learning model (due to the decrease in standard error that occurs when the overall model is improved, as often occurs when additional factors are added to the model); in fact, standard errors *did* decrease for that interaction term from the simple model to the temporal learning model.

[13] The temporal learning model did have significantly higher R^2 values than the simple model, as demonstrated by χ^2 tests, all *p*s < .05. This is consistent with the suggestion that adding variables that have a high correlation with the dependent measure increases model fit; however, this approach does not provide information about what happens to the critical congruency by PC interaction from the simple to temporal learning model.

adding parameters and comparing model fit statistics is commonly used in other disciplines (e.g., see Raftery, 1995).

Considering first transformed RT, for all three datasets, the AIC and BIC values for the isolated congruency by PC interaction got better (i.e., decreased) when temporal learning indices were included (decreased by 9 for AIC, 10 for BIC in Hutchison, 2011, and by 15 for AIC, 16 for BIC in Bugg, 2014, and by 13 for AIC and BIC in Gonthier et al., 2016); changes of strong to very strong magnitude according to Raftery (1995). These findings strongly support the temporal learning account. However, this conclusion is tenuous because several of the analyses using untransformed RT showed a different pattern. That is, the AIC and BIC values for the isolated congruency by PC interaction got poorer (i.e., increased) or stayed the same when temporal learning indices were included (increased by 6 for AIC and BIC in Hutchison, 2011, and increased by 9 for AIC and BIC in Gonthier et al., 2016). In contrast, the Bugg (2014) untransformed data showed a pattern consistent with the transformed data: a better fit (decreasing AIC/BIC values) when temporal learning indices were in the model versus not, but the effect size was much smaller—magnitude of only 3 and 4, a weaker "positive" effect size (Raftery, 1995). These analyses indicate that the best model fit for the isolated congruency by PC interaction includes the temporal learning indices for all three datasets for transformed RT, but does not include these indices for two of the three datasets for untransformed RT.

Next, we turned to the question of statistical significance. As discussed earlier, evidence for the temporal learning account comprised visual inspection of the critical interaction beta, which declined in the temporal learning model (Schmidt, 2013a). This approach, as well as the approaches taken above, still lack statistical analysis of the congruency by PC interaction as prior RT changes. To mitigate this issue, we began with the simplest statistical approach-- examining the three-way interaction (congruency by PC by prior RT; see Table 6). One possibility is that a three-way interaction would be found indicating that the congruency by PC interaction changed significantly as a function of prior RT, which seemingly would lend support to the temporal learning account. However, Schmidt (2013a) has contested this approach, suggesting that the three-way interaction is “hypothesis irrelevant” (p. e82320). At the same time, though, there is precedence for examining the three-way interaction in other time-based accounts, such as the Adaptation to the Statistics of the Environment (ASE) theory (Kinoshita, Mozer, & Forster, 2011; cf. Kinoshita, Forster, & Mozer, 2008; Mozer, Kinoshita & Davis, 2004). The ASE theory attributes PC effects in masked priming to a response initiation mechanism that adapts to the difficulty of recent trials (as reflected in prior RT). Drawing on prior research demonstrating that the conditions that produce a PC effect are those where the RTs on easy (i.e., congruent) items are *more* sensitive to the difficulty of the recent trials (prior RT) than the RTs on hard (i.e., incongruent) items (i.e., prior RT by congruency interaction), Kinoshita et al. (2011) predicted and found a three-way congruency by PC by prior RT interaction. Consistent with the assumptions of the ASE theory, the nature of the three-way interaction was such that the prior RT by PC interaction was observed for congruent but not incongruent trials. Because the three-way interaction is clearly not hypothesis-irrelevant in other timing-based accounts, and it is a simple way to test whether the PC by congruency interaction changes when prior RT is added to the model, examination of the three-way interaction seemed valuable in the present context.

With more than 15,000 observations in the Hutchison (2011) dataset, 5,000 observations in the Bugg (2014) dataset, and 14,000 observations in the Gonthier et al. (2016) dataset, there should be sufficient power to detect a three-way interaction if it exists. The three-way interaction was significant for transformed RT only in the Gonthier et al. (2016) dataset but was significant in all three datasets for untransformed RT. [14] The fact that the three-way interaction is sometimes significant and sometimes not suggests that any effect of prior RT on the list-wide PC effect is inconsistent and/or subtle. However, given that Schmidt (2013a) took issue with this approach, we explored one additional method of statistically evaluating the predictions from the temporal learning account to seek converging evidence for this conclusion.

Last, we conducted analyses to again examine comparisons in the change in R^2 from one model to another; however, now we used 95% confidence intervals to determine whether the change was statistically significant (see Table 7). Like the AIC/BIC analyses above, we constructed models designed to isolate the critical interaction *without* the influence of temporal learning indices [change in R^2 from model 1 (PC, congruency) to model 2 (PC, congruency, congruency × PC)] and to isolate the congruency by PC interaction *with* the influence of temporal learning indices [change in R^2 from model 3 (PC, congruency, prior RT, prior RT by congruency) to model 4 (PC, congruency, prior RT, prior RT by congruency, congruency by PC)][15]. Then we drew confidence intervals around the two changes in R^2 by using a simulation approach in the r2glmm package in R (Jaeger, 2016) to examine whether the change in the isolated congruency by PC interaction was statistically significant. This is a conservative approach whereby non-overlapping CIs (in Table 7, darker gray fill) indicate that the difference between the two changes in R^2 is statistically significant. When the change in R^2 was smaller for

[14] Although Schmidt (2013a) has deemed the three-way interaction hypothesis-irrelevant and thus has not made any predictions about the nature of this interaction, it is notable that three of the four significant three-way interactions observed here reflected a prior RT by PC interaction for incongruent trials, but not congruent trials (Gonthier et al., 2016 transformed and untransformed; Hutchison, 2011, untransformed, although in this case the effect on incongruent trials was a trend [$p < .10$]). This differs from the three-way pattern predicted by and resulting from time-based processes *a la* the ASE account, which shows a prior RT by congruency interaction in congruent, but not incongruent trials. For the only other significant three-way (Bugg, 2014), there was neither an effect on congruent or incongruent trials.

[15] R^2 difference tests and chi-squared likelihood ratio tests both indicated that models including the congruency by PC interaction (models 2 and 4) fit the data better than models without the congruency by PC interaction (models 1 and 3, respectively). This is an indication that the theoretically meaningful congruency by PC interaction, regardless of whether it reflects control or other mechanisms, is an important part of the statistical model of the data.

the model containing temporal learning indices, we interpreted such a result as being consistent with the temporal learning account.

As shown in Table 7, there were two cases in which confidence intervals were overlapping (Hutchison, 2011, transformed RT, and Bugg, 2014, untransformed RT), indicating that adding the temporal learning indices did not significantly change the critical congruency by PC interaction. In all other analyses, (Hutchison, 2011, untransformed RT; Bugg, 2014, transformed RT; and Gonthier et al., 2016, transformed and untransformed RT), the non-overlapping confidence intervals suggested that the *change* in the congruency by PC interaction from a model without to a model with the indices of temporal learning was statistically significant. However, these comparisons indicated inconsistent effects of temporal learning on the critical interaction. Specifically, for the Bugg (2014) and Gonthier et al. (2016) transformed RT analyses, the isolated congruency by PC interaction was smaller with the addition of temporal learning indices, consistent with what the temporal learning account might predict. The opposite was true for Hutchison (2011) and Gonthier et al. (2016) using untransformed RT, in which the isolated congruency by PC interaction was *larger* with the addition of temporal learning indices. These findings indicate that the changes in the critical interaction with the addition of measures purported to reflect temporal learning are small, and only transformed RT analyses provide (albeit inconsistent) statistical support for the temporal learning model.

To summarize this section, multiple approaches to evaluating the meaningfulness and/or significance of the change in the critical congruency by PC interaction pointed to small and inconsistent effects of temporal learning (see Table 3 for summary of all analyses and evidence for/against the temporal learning account). Much of the evidence that was supportive of the temporal learning account occurred in only one or two of the datasets at a time, and usually in transformed RT. However, there were cases in transformed RT, in addition to untransformed RT, where results ran counter to the predictions of the temporal learning account. Collectively, this fuller set of analyses provides at best sporadic support for the temporal learning account of the list-wide PC effect. The inconsistency of the results suggests that we may be capturing noise and random fluctuations in the models rather than true change in the congruency by PC interaction attributable to temporal learning.

Section 3: Measuring temporal rhythm

A key assumption of the temporal learning account is that the temporal rhythm of MC and MI lists can be operationalized as prior RT, that is, the RT on the immediately preceding trial (*n* – 1) centered to the grand mean of all RTs (Schmidt, 2013a). However, Schmidt ultimately concluded that "[p]revious [prior] RT is simply a very poor measure of temporal learning that will only explain a small fraction of the variance due to temporal learning processes" (p. e82320). When discussing the fact that including prior RT reduced but did not eliminate the list-wide PC effect, he reminded the reader that prior RT may be a "very weak proxy of temporal expectancy" and specifically noted that "temporal expectancies are likely based on more than just the previous trial" and "noisiness in temporal expectancies will further reduce the explanatory power of previous RT" (pp. e32320-e32321). The implication is that a better measure of temporal learning may provide stronger support for the temporal learning account—that is, such a measure may fully eliminate the congruency by PC interaction when added to the model. To evaluate this possibility, we developed an alternative way of operationalizing temporal learning based on multiple prior RTs centered to the grand mean rather than a single prior RT. We posited that the average of *three* prior RTs ($M_{\text{three prior RTs}}$)[16], an

[16] We created this measure based on three trials back in part because of a suggestion in Schmidt (2013a): "previous response time can be regarded only as a weak proxy of temporal expectancy (e.g., participants inevitably account for

“extended” measure of temporal learning, would be less noisy and therefore a potentially better measure than prior RT (cf. deBettencourt, Norman, & Turk-Browne, 2017, who found the average of three prior RTs to be a stable and predictive measure of performance in a different paradigm). This accords with Schmidt’s (2013b) sentiment that temporal expectancy likely takes into account more than just the previous trial.

First, we asked the question of whether an extended temporal learning model that utilizes $M_{\text{three prior RTs}}$ provides a better model fit than the original temporal learning model that was based on a single prior RT. Because these models are non-nested, the likelihood ratio test was not appropriate. Instead we statistically compared the extended and original temporal learning models based on simulations using the r2glmm package (Jaeger, 2016) in R. To accommodate the new operationalization, we had to filter trials on which any of the three prior trials had RTs less than 300ms or errant responses, eliminating on average, 10% of trials from each dataset. In all cases (for transformed RT and untransformed RT in all three datasets), this analysis showed that the R^2 for the extended model was greater than that of the original temporal learning model (see Table 8), suggesting that the extended model has a better model fit compared to the original model.

Having demonstrated that a model including the extended measure of temporal learning provides a better fit of the data than the original temporal learning model, we then examined whether the list-wide PC effect would be eliminated or (numerically) more strongly reduced when the extended measure of temporal learning is added to the simple model. To address this

more trials than just the most recent one)”, p.619. However, recognizing that three trials may seem arbitrary and to be comprehensive, we also explored up to five trials back (i.e., two, four, and five in addition to the three reported here). The results were somewhat variable when incorporating more RTs back, but the overall conclusions did not differ from the analysis using three trials back (and in every case, the congruency by PC interaction remained significant after adding the prior RT measure(s)). Thus, we report three trials back, as a balance between parsimony and reliability.

question, we again replicated Schmidt's (2013a) primary analyses albeit replacing the original indices of temporal learning (prior RT and the prior RT by congruency interaction) with the extended indices ($M_{\text{three prior RTs}}$ and the $M_{\text{three prior RTs}}$ by congruency interaction) and again adjusting the filtering to accommodate the new operationalization (see Table 9).

For transformed RT, the beta for the critical congruency by PC interaction decreased from the simple model to the extended temporal learning model in all three datasets (decreases of 0.001 for Hutchison, 2011, 0.023 for Bugg, 2014, and 0.011 for Gonthier et al., 2016), as expected per a temporal learning account. These magnitudes were nominally greater than the decreases observed from the simple model to the original temporal learning model for the Bugg (2014) and Gonthier et al. (2016) datasets.[17] Most importantly, however, for all three datasets the beta changes were again relatively small and the congruency by PC interaction remained highly significant in the extended temporal learning model. For untransformed RT, the results were mixed. For Bugg (2014) the critical interaction beta decreased by 1.974 from the simple to the extended temporal learning model but for the Hutchison (2011) and Gonthier et al. (2016) datasets, the beta increased (magnitudes of 0.081 and 3.344, respectively). These patterns are similar to those we observed when comparing the simple model to the original temporal learning model (see *Section 1*).[18] Most critically, in no case was the critical interaction eliminated, including for the Bugg (2014) dataset that provided the most consistent evidence (across transformed and untransformed RT) in favor of the extended temporal learning model.

In the extended temporal learning operationalization presented above, the influence of prior trials was equally weighted. It is possible, however, that a better operationalization of temporal learning weights the influence of preceding trials in proportion to their distance from

[17] This was true regardless of whether the original filtering or the extended filtering was applied to the model.
[18] Again, this was true regardless of whether the original filtering or the extended filtering was applied to the model.

the current trial (cf. Mozer, Kinoshita, & Shettel, 2007, for a sophisticated computational model of the exponentially-decaying influence of prior RT on several aspects of cognitive performance).[19] To address this possibility, we examined a weighted extended RT composite to approximate an exponentially decaying function of the influence of prior RTs, with the first trial back contributing the most weight (60%), the second trial back contributing less weight (30%), and the third trial back contributing the least weight (10%). As shown in Table 8, this weighted extended temporal learning model also provided a better model fit to the data than the original temporal learning model (Schmidt, 2013a) though it fit the data less well than the (unweighted) extended temporal learning model presented above (in terms of R^2, see Table 9). Of most importance, the results from this weighted extended temporal learning model were similar to the extended temporal learning model: for transformed RT, the beta for the critical congruency by PC interaction decreased from the simple model to the weighted extended temporal learning model in all three datasets (decreases of 0.014 for Hutchison, 2011, 0.020 for Bugg, 2014, and 0.010 for Gonthier et al., 2016; as predicted by a temporal learning account, see Table 9). For untransformed RT, consistent with the extended temporal learning model analyses, the beta for the critical congruency by PC interaction decreased from the simple model to the weighted extended temporal learning model for only Bugg (2014), a decrease of 1.465, whereas it increased for Hutchison (2011) and Gonthier et al. (2016), by 2.198 and 3.177, respectively. In sum, in all three datasets and for both transformed and untransformed RT analyses, the beta changes were again relatively small and critically, the congruency by PC interaction remained highly significant even after the weighted extended temporal learning measure was added to the analyses.

[19] We are grateful to Sachiko Kinoshita and James Schmidt for this suggestion.

In conclusion, it may be true that prior RT is not an optimal measure of temporal learning; however, using *evidentially* more robust measures of temporal learning (simple or weighted mean of three prior RTs [see also Footnote 16] rather than one prior RT) did not yield results bolstering the temporal learning account.[20] That is, the congruency x PC interaction still only decreased slightly (and actually increased in several cases) and was always highly significant in the extended and weighted temporal learning models. Therefore, we conclude that even with a better operationalization of temporal learning, the evidence for temporal learning remains weak and inconsistent. It is of course still possible that a superior measure of temporal learning exists that might provide solid evidence for the temporal learning account. For example, one might argue that we simply did not get the weights "right" in the weighted temporal learning model. However, considering the collective set of analyses we performed (extended temporal learning measure, weighted extended temporal learning measure, and analyses referenced in Footnote 16), we in fact evaluated evidence for temporal learning for first trial weights ranging from 20% to 100%. For this reason, and because the temporal learning account does not make a precise prediction about the weights, another viable candidate measure is not obvious. It is also noteworthy that the technique of using just one prior RT (n – 1) as an index of time-based learning processes has been used in the context of the ASE theory to examine the locus of PC effects in other domains (e.g., Kinoshita et al., 2011), and one prior RT has proven *not* to be too noisy or inaccurate to observe predicted interaction effects of prior RT on a PC effect in that literature. All things considered, the clearest interpretation is that prior RT (or a better measure

[20] For comprehensiveness, we also examined the efficacy of the extended operationalization of temporal learning (unweighted, since it explained more variance than the weighted version) by conducting all analyses reported in *Section 2* after replacing prior RT with the mean of the three prior RTs. Results were largely consistent with those obtained using the original temporal learning model, and like the analyses we report in full in this section, did not provide further support for the temporal learning account.

of temporal learning) does not eliminate the congruency by PC interaction in the Stroop task because temporal learning fails to explain the list-wide PC effect or has only a very small influence on this effect.

General Discussion

This article aimed to fill an important gap in the literature by critically evaluating the temporal learning account of the list-wide PC effect (Schmidt, 2013a; 2013b). To the extent that evidence for temporal learning is strong and consistent, this account could potentially supplant the global control account. Analyses were performed on three published datasets in which list-wide PC effects were observed on critical trials, consistent with the global control account (Bugg, 2014; Hutchison, 2011; Gonthier et al., 2016). These analyses included replications of those originally reported by Schmidt (2013a) for the Hutchison (2011) dataset, and multiple additional analyses for purposes of providing a rigorous and comprehensive test of the temporal learning account across all three datasets (summary in Table 3). We found that the evidence for the temporal learning account was neither strong nor consistent. Moreover, the list-wide PC effect remained highly significant in all attempts to control for temporal learning, including with use of better measures than one prior RT. Therefore, we do not currently consider the temporal learning account to be a robust competitor to the global control account of the list-wide PC effect. The most serious limitations of the temporal learning account that our analyses uncovered are as follows:

a) Conclusions regarding the critical signature of the list-wide PC effect, the congruency by PC interaction, were based solely on modeling inverse-transformed RT data from Hutchison (2011). While this transformation normalizes RT, others have demonstrated it is potentially problematic (Balota et al., 2015; Lo & Andrews, 2015). In line with these concerns, across three datasets (Bugg, 2014; Gonthier et al., 2016; Hutchison, 2011), we found disparate results when untransformed RT (based almost exclusively on the GLMM approach) was used as an alternative to the inverse transformation. As Table 3 illustrates, analyses of untransformed RT using GLMM were clearly less consistent with the temporal learning account than analyses of transformed RT; however, analyses of transformed RT also yielded inconsistent results;[21]

b) The change in the congruency by PC interaction, when prior RT was added to the (simple) model as an index of temporal learning, was small and inconsistent across datasets and data analysis types (transformed or untransformed RT). For some datasets the interaction effect decreased, consistent with the temporal learning account, but in other cases the effect did not change, or even increased. Importantly, there was not a single instance in which the congruency by PC interaction was eliminated when indices of temporal learning were added;

c) The temporal learning account of the list-wide PC effect was in part based on observing a change in the interaction (beta) when indices of temporal learning were added to the simple model (Schmidt, 2013a). There was not an objective assessment of the meaningfulness and/or statistical significance of the change in the congruency by PC interaction. We tackled this concern via multiple approaches as described in *Section 2*. Across analyses the change was neither consistently significant nor consistently in the direction predicted by the temporal learning account; and

---

[21] Furthermore, although not the primary focus of the present analyses because this two-way interaction alone does not inform theoretical mechanisms underlying the list-wide PC effect, it was also the case that the prior RT by congruency interaction was either non-significant or in the opposite direction as predicted by a temporal learning account when examined using untransformed RT GLMM analyses.

d) The temporal learning account was developed based on the assumption that temporal rhythm could be operationalized as prior RT, which Schmidt (2013a) ultimately deemed a poor and noisy measure that limited the ability to observe an elimination of the list-wide PC effect once prior RT was added to the model. We developed two new (extended) measures of temporal learning based on a simple or weighted average of three prior RTs that we anticipated might be better and less noisy (cf. Schmidt, 2013b). Although the extended and weighted extended temporal learning models fit the data better than the original temporal learning model, critically, there was still not a single instance in which the congruency by PC interaction was eliminated when the extended or weighted extended indices of temporal learning were added to the simple model.

These observations are not intended to suggest that we found no evidence in support of the temporal learning account. As is apparent from Table 3, some results were consistent with the account. For instance, considering transformed RT, the beta for the congruency by PC interaction did decrease after accounting for temporal learning indices and the AIC/BIC statistics supported the temporal learning account in all three datasets, suggesting that temporal learning contributes to the list-wide PC effect. However, the changes in the critical interaction were small and not always statistically significant. Consequently, we cannot reject the possibility that even in these cases the change in the critical interaction may be capturing noise and random fluctuations in the model rather than true signal. It is therefore our conclusion that the global control account of the list-wide PC effect is not currently undermined by the temporal learning account.

Moving Forward: Modeling Approaches

We have criticized the temporal learning account of the list-wide PC effect on empirical and conceptual grounds. However, that is not to suggest that temporal learning or other timing-related processes are wholly irrelevant to performance in a list-wide PC paradigm but rather to emphasize that there is very clearly a strong list-wide PC effect that survived systematic attempts to control for such processes. Indeed, what was remarkably consistent amidst a great deal of inconsistency was the list-wide PC effect. Across all three datasets, both data types (transformed and untransformed), and all analyses, the congruency by PC interaction was highly significant and never eliminated after controlling for temporal learning, consistent with the global control account.

Our criticisms of the temporal learning account by no means suggest that the baby should be thrown out with the bath water. We recognize that future research may identify a time-based account that does in fact undermine the global control account of the list-wide PC effect provided that the following key patterns can be demonstrated: 1) consistent evidence for an effect of the time-based process on the congruency effect (i.e., significant main effect of prior RT [or another measure] on current RT and significant prior RT by congruency interaction that are in the direction anticipated by the account), 2) consistent evidence that the change in the critical congruency by PC interaction (from the simple to temporal learning model) is statistically significant, and 3) consistent evidence for a reduction in the beta associated with the list-wide PC effect (i.e., congruency by PC interaction) to a non-significant level (i.e., no list-wide PC effect) when indices of time-based processes are added to the model. Were only the first and second criteria supported, this would imply that the time-based process does make a significant contribution to the list-wide PC effect, but the list-wide PC effect nevertheless survives after controlling for that process. Without additionally demonstrating support for the third criterion, it cannot be concluded that the global control account is invalid. Evidence supporting all three criteria would constitute strong motivation for researchers to change the list-wide PC design or routine analytic strategies to potentially enable examination of global control independent of temporal learning.

One might ask whether the ASE theory (e.g., Kinoshita et al., 2011; cf. Kinoshita et al., 2008; Mozer et al., 2004) represents a time-based account of the list-wide PC effect that meets these criteria. As described earlier, the ASE also uses prior trial RT to represent its time-based process, however, prior RT is a proxy for the difficulty of recent trials. The assumption is that response initiation is adapted to difficulty, which can vary across list contexts. Given that we modeled prior RT to test the temporal learning account (Schmidt, 2013a; 2013b) but did not find evidence suggesting a strong and consistent contribution of prior RT to the list-wide PC effect, one might be inclined to conclude that the current results may also be taken as evidence against the ASE theory. However, we are reluctant to draw this conclusion because the ASE theory is an account of PC effects in masked priming. That is, the ASE theory was proposed as an alternative to an extant account that was based on the utility of the prime, which was presumed to be higher in a mostly related (congruent) context. As Kinoshita et al. (2011) pointed out, it is unclear how prime utility could affect performance in a masked priming paradigm. The ASE theory is not an account of PC effects in unmasked priming, where a prime utility explanation is perfectly plausible. The Stroop task is more similar to unmasked than masked priming in that the word and color (or picture) are fully visible on each trial and therefore the congruency of a given trial can be registered by participants. As such, it is unclear whether the ASE theory would attempt to explain the list-wide PC effect.

Future attempts to examine potential time-based accounts of the list-wide PC effect may also benefit from the following observation gleaned from the present study. An account that is based largely on a single approach (e.g., examining change in interaction beta from simple to temporal learning model) to modeling data from a single study may lead to premature conclusions. If one examines Table 3, it is apparent that a more comprehensive and rigorous approach to analyzing the transformed data, or analysis of the untransformed data, of Hutchison (2011) would not have yielded consistent support for the temporal learning account of the list-wide PC effect. In striking contrast, if a researcher had analyzed only transformed data from Gonthier et al. (2016), s/he would have found consistent support for the temporal learning account.[22] Yet, our analyses of all three datasets indicate that this conclusion too would have been weakly informed. With respect to future modeling attempts aimed at examining contributions of time-based processes to the congruency by PC interaction, we concur with Balota et al. (2014) in suggesting that "…at the very least, one needs to explore both the raw data and the transformed data when making inferences about how variables combine to influence performance." (p. 1570).

Moving Forward: Experimental Approaches

One might criticize the current study by arguing that we have over-emphasized the modeling results and ignored relevant experimental work on the temporal learning account. However, we would argue that we have not misrepresented the quality of the current evidence

[22] In contrast to Bugg (2014) and Hutchison (2011), Gonthier et al. (2016) used a picture-word Stroop task and a within-subject manipulation of list-wide PC. There is no *a priori* reason why picture-word Stroop should be more subject to temporal learning influences, but temporal learning may have a more robust effect in a within-subjects design wherein participants perform both MC and MI lists and are relatively faster (slower) in one list. If future research confirms this speculation, then for researchers who are interested in targeting global modulations of control with use of the list-wide PC manipulation, a clear suggestion would be to avoid within-subject PC manipulations.

favoring the temporal learning account by taking this approach because there are serious limitations with extant experimental work testing the temporal learning account (as detailed below). Furthermore, considering *all* the relevant experimental evidence and not solely those studies designed to test the temporal learning account, there are findings that are accommodated by the global control account but are not readily explained by temporal learning processes. For example, Entel, Tzelgov, and Bereby-Meyer (2014) and Bugg, Diede, Cohen-Shikora, and Selmeczy (2015) showed list-wide PC effects for lists equated in PC but differing in instructions provided in advance of the lists. In this case, global control worked to configure attention in a context-sensitive manner; in contrast, it is unclear how temporal rhythms would explain this pattern because all item types were equated across conditions (i.e., there were no biased trials to set differential rhythms across lists). Additionally, Wendt et al. (2012) found that the global likelihood of conflict on flanker trials in MC and MI lists affected performance on interspersed visual search trials (comprising different stimuli and responses). The key pattern supporting a global control account was that search times were longer when the search target appeared in a position corresponding to the flanker stimuli relative to the target stimulus, and this slowing was exacerbated in the MI list where participants were presumably engaging a global control setting involving the perceptual filtering of (mostly conflicting) flanker stimuli. For the temporal learning account to explain this pattern, it would have to assume rhythms are task-general (i.e., individuals automatically apply the rhythm from one task to another regardless of the degree of stimulus/response overlap) and can reflect very different mean RTs, in addition to specifying which task is dominant in establishing the rhythms within a given list, which it presently does not do. As a final example, Hutchison (2011) found modulations of the list-wide PC effect as a function of factors related to control (i.e., working memory capacity and PC of the specific items within the list) and not speed, that are accommodated by a global control account (see Hutchison, 2011, p. 858), but are not predicted or well explained by a temporal learning account.

Regarding the extant experimental work that was designed to directly test the temporal learning account, two approaches have been pursued: (1) manipulating timing in tasks without conflict or control demands to examine whether an effect analogous to a list-wide PC effect is observed (Schmidt, 2014), and (2) controlling the temporal rhythms across MC and MI contexts and tracking changes in the PC effect (Schmidt, 2017). Adopting the first approach, Schmidt (2014) manipulated the proportion of high- versus low-contrast trials between lists in a letter identification task and found a pattern similar to the list-wide PC effect: letters in the high-contrast condition (faster, analogous to an MC condition) produced larger contrast effects than did letters in the low-contrast condition (slower, analogous to an MI condition). Based on these patterns, Schmidt (2014) concluded that a “conflict adaptation account is entirely ruled out.” (p.375). However, we would argue that while such patterns provide potential evidence for a role of temporal learning in producing effects that manifest similarly to control phenomena, just because a pattern of data looks like another pattern, it does not mean that the underlying mechanisms are necessarily the same. In other words, a global control mechanism may operate in a task that includes conflict such as the Stroop task but not in a conflict-free task (indeed, there would be no input to guide its operation).

Adopting the second approach, Schmidt (2017) used a prime-probe direction task (participants press a button corresponding to the direction of the probe which can match [congruent] or mismatch [incongruent] the prime). Temporal rhythm was controlled by having participants withhold responses to probes for a set time on select trials to equate the timing of congruent and incongruent trials. When participants were subjected to a very short wait condition, a congruency effect was observed, and the typical list-wide PC effect was found. In contrast, when participants were subjected to a longer wait condition, there was not a congruency effect (i.e., congruent and incongruent response times were equated), and the list-wide PC effect was eliminated. As Schmidt (2017) acknowledged, while this finding supports the temporal learning account, there are also several limitations. One is that theories about global control were developed based on performance in the Stroop task (i.e., list-wide PC effects on critical items) and not the prime-probe task he employed (and this criticism applies equally to the evidence from Wendt et al., 2012, reviewed above). Second, introducing a long wait time may have hindered conflict detection and/or the application of cognitive control. To this point, there was no congruency effect in the longer wait condition that eliminated the list-wide PC effect. This may suggest that the typical signals of conflict that inform global control are minimal or absent in that condition. Still, we see experiments like this as a strong step in the right direction and believe future experiments aimed at examining the role of temporal learning in the list-wide PC effect in Stroop tasks will be theoretically informative.

One clear advantage of experimentation is that it offers a straightforward means of examining predictions of the temporal learning account without having to rely on complicated models, which themselves suffer from the limitations discussed herein (interpretability, reliance on transformations that may change the nature of what is being studied [DV] and interactions of interest, assessing statistical significance of changes). A second advantage is that experimentation circumvents problems pertaining to the construct validity of temporal learning indices. That is, even if future research were to obtain the modeling evidence detailed above that would provide strong statistical support for the temporal learning account, it would still be vulnerable to the criticism that prior RT (or an alternative measure) may be reflecting some process other than temporal learning that could produce a similar set of effects. As noted earlier, Kinoshita et al. (2011) interpreted prior RT as an indicator of prior trial difficulty. Without independent validation of these constructs, it seems that an account of prior RT (or another measure) based on temporal learning is on no stronger footing than one based on some other process, including learning-related (e.g., Abrahamse, Braem, Notebaert, & Verguts, 2016) or other (e.g., cautiousness; extent of word processing) processes that may inform or coincide with cognitive control.

For these reasons, continued experimentation seems fruitful for seeking converging evidence and potentially characterizing, as has been done in the item-specific PC literature (see e.g., Bugg, Jacoby, & Chanani, 2011; Bugg, 2015), conditions under which mechanisms other than control such as temporal learning may contribute to the list-wide PC effect. We currently have several experiments in progress to further assess the validity of the temporal learning account, specifically addressing some of the concerns raised above with respect to prior experimental work: We are using the Stroop task and controlling temporal rhythm to look for changes in indices of global control.

Conclusion

The contingency account (Schmidt & Besner, 2008) and item-specific control account (Blais et al., 2007; Bugg et al., 2008) were influential in spurring developments in the cognitive control literature with respect to experimental design (e.g., controlling contingencies or item-specific PC) and theory (e.g., regarding mechanisms that underlie list-wide PC effects). These developments were grounded in firm evidence demonstrating that contingency and/or item-specific control could produce the list-wide PC pattern independent of any contribution of global control. The current study considered whether the temporal learning account may be similarly influential, potentially affirming Schmidt's (2013a) conclusion that the critical contribution of his study was "the cautionary demonstration that the list-level PC effect cannot be taken as strong evidence for conflict adaptation without further controls." (p. e82320). We applied multiple statistical methods to the transformed and untransformed data from three previously published studies that had yielded a list-wide PC effect and found that a) the statistical evidence for the temporal learning account was neither strong nor consistent, and b) there was a highly significant list-wide PC effect that survived multiple, systematic attempts to control for temporal learning. Therefore, we conclude that the temporal learning account is not currently a robust competitor to the global control account of the list-wide PC effect. Accordingly, we cannot justify recommending that researchers adopt additional controls to account for temporal learning when investigating list-wide PC effects. Examination of the list-wide PC effect for critical items (as is currently the gold standard) effectively controls for those mechanisms that clearly have been demonstrated to be viable confounds if left uncontrolled (i.e., contingency learning and item-specific control).

References

Abrahamse, E., Braem, S., Notebaert, W., & Verguts, T. (2016). Grounding cognitive control in associative learning. *Psychological Bulletin*, *142*(7), 693. doi: 10.1037/bul0000047

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437-443.

Balota, D. A., Aschenbrenner, A. J., & Yap, M. J. (2013). Additive effects of word frequency and stimulus quality: the influence of trial history and data transformations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(5), 1563.

Bates, D.; Maechler, M.; Bolker, B.; Walker, S. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7. 2014. http://CRAN.R-project.org/package=lme4>

Blais, C., & Bunge, S. (2010). Behavioral and neural evidence for item-specific performance monitoring. *Journal of Cognitive Neuroscience*, *22*(12), 2758-2767. doi:10.1162/jocn.2009.21365

Blais, C., Harris, M. B., Guerrero, J. V., & Bunge, S. A. (2012). Rethinking the role of automaticity in cognitive control. *The Quarterly Journal of Experimental Psychology*, *65*(2), 268-276. doi: 10.1080/17470211003775234

Blais, C., Robidoux, S., Risko, E. F., & Besner, D. (2007). Item-specific adaptation and the conflict-monitoring hypothesis: A computational model. *Psychological Review*, *114*(4), 1076-1086. 10.1037/0033-295X.114.4.1076Botvinick, 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.

Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In A. R. A. Conway, C.

Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 76–106). Oxford, England: Oxford University Press.

Bugg, J. M. (2012). Dissociating levels of cognitive control: The case of Stroop interference. *Current Directions in Psychological Science, 21*, 302–309. doi:10.1177/0963721412453586

Bugg, J. M. (2014). Conflict-triggered top-down control: Default mode, last resort, or no such thing? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(2), 567-587. doi:10.1037/a0035032

Bugg, J. M. (2015). The relative attractiveness of distractors and targets affects the coming and going of item-specific control: Evidence from flanker tasks. *Attention, Perception, & Psychophysics*, *77*(2), 373 – 389.

Bugg, J. M. (2017). Context, conflict, and control. In T. Egner (Ed.) The Wiley handbook of cognitive control. (pp. 79 – 96). Chichester, West Sussex, UK: John Wiley & Sons.

Bugg, J. M., & Chanani, S. (2011). List-wide control is not entirely elusive: Evidence from picture-word Stroop. *Psychonomic Bulletin & Review, 18,* 930- 936. doi:10.3758/s13423-011-0112-y

Bugg, J. M., Jacoby, L. L., & Chanani, S. (2011). Why it is too early to lose control in accounts of item-specific proportion congruency effects. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 844 – 859.

Bugg, J. M., & Crump, M. J. C. (2012). In support of a distinction between voluntary and stimulus-driven control: A review of the literature on proportion congruent effects. *Frontiers in Psychology: Cognition, 3*, 1-16. doi: 10.3389/fpsyg.2012.00367

Bugg, J. M., & Diede, N. T. (2017). The effects of awareness and secondary task demands on Stroop performance in the precued lists paradigm. *Acta Psychologica.* doi: 10.1016/j.actpsy.2016.12.013

Bugg, J. M., Diede, N. T., Cohen-Shikora, E., & Selmeczy, D. (2015). Expectations and experience: Dissociable bases for cognitive control? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 1349-1373. doi: 10.1037/xlm0000106

Bugg, J. M., Jacoby, L. L., & Toth, J. P. (2008). Multiple levels of control in the Stroop task. *Memory & Cognition*, *36*(8), 1484-1494.

Bugg, J. M., McDaniel, M. A., Scullin, M. K., & Braver, T. S. (2011). Revealing list-level control in the Stroop task by uncovering its benefits and a cost. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 1595 – 1606. doi: 0.1037/a0024670

Cheesman, J., & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *40*(4), 343.

deBettencourt, M.T., Norman, K.A., Turk-Browne, N.B. (2017). Forgetting from lapses of sustained attention. *Psychonomic Bulletin and Review.* doi: 10.3758/s13423-017-1309-5

De Pisapia, N., & Braver, T. S. (2006). A model of dual control mechanisms through anterior cingulate and prefrontal cortex interactions. *Neurocomputing*, *69*(10-12), 1322-1326.

Dishon-Berkovits, M., & Algom, D. (2000). The Stroop effect: It is not the robust phenomenon that you have thought it to be. *Memory & Cognition*, *28*(8), 1437-1449.

Egner, T., & Hirsch, J. (2005). The neural correlates and functional integration of cognitive control in a Stroop task. *Neuroimage*, *24*(2), 539-547.

Entel, O., Tzelgov, J., & Bereby-Meyer, Y. (2014). Proportion congruency effects: instructions may be enough. *Frontiers in Psychology*, *5,* 1108.

Funes, M. J., Lupiáñez, J., & Humphreys, G. (2010). Analyzing the generality of conflict adaptation effects. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(1), 147.

Glaser, M.O., & Glaser, W.R. (1982). Time course analysis of the Stroop phenomenon. *Journal of Experimental Psychology: Human Perception & Performance*, 8, 875-894.

Gonthier, C., Braver, T. S., & Bugg, J. M. (2016). Dissociating proactive and reactive control in the Stroop task. *Memory & Cognition*, *44*(5), 778-788.

Hutchison, K. A. (2011). The interactive effects of list-based control, item-based control, and working memory capacity on Stroop performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 851–860. doi:10.1037/a0023437

Jacoby, L. L., Lindsay, D. S., & Hessels, S. (2003). Item-specific control of automatic processes: Stroop process dissociations. *Psychonomic Bulletin & Review, 10,* 638–644. doi:10.3758/BF03196526

Jaeger, B. (2016). R2glmm: Computes R squared for mixed (multilevel) models. *R package version 0.1, 2*.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*(1), 47-70.

Kinoshita, S., Forster, K. I., & Mozer, M. C. (2008). Unconscious cognition isn't that smart: Modulation of masked repetition priming effect in the word naming task. *Cognition, 107*, 623– 649. doi:10.1016/j.cognition.2007.11.011

Kinoshita, S., Mozer, M. C., & Forster, K. I. (2011). Dynamic adaptation to history of trial difficulty explains the effect of congruency proportion on masked priming. *Journal of Experimental Psychology: General*, *140*(4), 622.

Lindsay, D. S., & Jacoby, L. L. (1994). Stroop process dissociations: The relationship between facilitation and interference. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 219-234.

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*.

Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology*, *12*(4), 523-553.

Logan, G. D., & Zbrodoff, N. J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & Cognition*, *7*(3), 166-174.

Logan, G. D., Zbrodoff, N. J., & Williamson, J. (1984). Strategies in the color-word Stroop task. *Bulletin of the Psychonomic Society*, *22*(2), 135-138.

Lowe, D. G., & Mitterer, J. O. (1982). Selective and divided attention in a Stroop task. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *36*(4), 684-700.

Melara, R. D., & Algom, D. (2003). Driven by information: a tectonic theory of Stroop effects. *Psychological Review*, *110*(3), 422.

Mozer, M. C., Kinoshita, S., & Davis, C. (2004). Control of response initiation: Mechanisms of adaptation to recent experience. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the Twenty Sixth Annual Conference of the Cognitive Science Society* (pp. 981–986). Hillsdale, NJ: Erlbaum.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological methodology*, 111-163.

Robidoux, S. (2017, March 19). Geometry and Harmony: The Consequences of Transforming Data [Blog post]. Retrieved from http://serjerobidoux.blogspot.com/2017/03/geometry-and-harmony-consequences-of.html.

Schmidt, J. R. (2013a). Temporal learning and list-level proportion congruency: Conflict adaptation or learning *when* to respond? PLoS ONE, *8*, e82320. doi: 10.1371/journal.pone.0082320

Schmidt, J. R. (2013b). Questioning conflict adaptation: Proportion congruent and Gratton effects reconsidered. *Psychonomic Bulletin & Review*, *20*, 615 – 630. doi: 10.3758/s13423-012—0373-0

Schmidt, J. R. (2014). List-level transfer effects in temporal learning: Further complications for the list-level proportion congruent effect. *Journal of Cognitive Psychology, 26,* 373–385. doi: 10.1080/20445911.2014.896367

Schmidt, J. R. (2017). Time-out for conflict monitoring theory: Preventing rhythmic biases eliminates the list-level proportion congruent effect. *Canadian Journal of Experimental Psychology/Revue canadienne de Psychologie Expérimentale, 71*(1), 52.

Schmidt, J. R., & Besner, D. (2008). The Stroop effect: why proportion congruent has nothing to do with congruency and everything to do with contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(3), 514-523. doi: 10.1037/0278-7393.34.3.514

Schmidt, J. R., Crump, M. J. C., Cheesman, J., & Besner, D. (2007). Contingency learning without awareness: Evidence for implicit control. *Consciousness and Cognition, 16*, 421–435. doi:10.1016/j.concog.2006.06.010

Shor, R. E. (1975). An auditory analog of the Stroop test. *Journal of General Psychology*, *93*, 281.

Torres-Quesada, M., Funes, M. J., & Lupiáñez, J. (2013). Dissociating proportion congruent and conflict adaptation effects in a Simon–Stroop procedure. *Acta psychologica*, *142*(2), 203-210.

West, R., and Baylis, G. C. (1998). Effect of increased response dominance and contextual disintegration on the Stroop interference effect in older adults. *Psychology and Aging*, 13, 206–217, doi: 10.1037/0882-7974.13.2.206

Wühr, P., Duthoo, W., & Notebaert, W. (2015). Generalizing attentional control across dimensions and tasks: Evidence from transfer of proportion-congruent effects. *The Quarterly Journal of Experimental Psychology*, *68*(4), 779-801.

Table 1

Analyses of Congruency × PC interaction for the Simple Model and Temporal Learning Model, Using Transformed and Untransformed RT, in the Hutchison (2011), Bugg (2014), and Gonthier et al. (2016) Datasets

			Transformed RT, LME				Untransformed RT, LME			
		Variable	Est.	*SE*	*t*	*p*	Est.	*SE*	*t*	*p*
Hutchison (2011)	Simple model	Intercept	-1.663	0.027	-61.666	<.001	633.058	15.500	40.843	<.001
		Congruency	0.237	0.024	9.935	<.001	112.683	15.029	7.498	<.001
		PC	-0.006	0.028	-0.215	.830	-4.087	13.114	-0.312	.756
		Congruency × PC	0.057	0.009	6.202	<.001	25.171	5.217	4.825	<.001
	Temporal Learning model	Intercept	-1.673	0.025	-66.796	<.001	630.349	14.820	42.534	<.001
		Congruency	0.247	0.023	10.698	<.001	115.300	14.785	7.798	<.001
		PC	0.011	0.024	0.485	.628	0.597	11.555	0.052	.959
		Prior RT	0.193	0.008	23.897	<.001	0.113	0.008	13.521	<.001
		Congruency × PC	0.049	0.009	5.345	<.001	28.836	5.215	5.530	<.001
		Congruency × Prior RT	-0.101	0.012	-8.308	<.001	0.049	0.013	3.655	<.001
Bugg (2014) 1a & 2b	Simple model	Intercept	-1.698	0.035	-48.545	<.001	610.789	15.827	38.592	<.001
		Congruency	0.197	0.025	7.995	<.001	83.171	12.766	6.515	<.001
		PC	-0.052	0.041	-1.269	.209	-19.068	16.948	-1.125	.264
		Congruency × PC	0.099	0.013	7.458	<.001	38.258	5.951	6.429	<.001
	Temporal Learning model	Intercept	-1.709	0.032	-53.394	<.001	607.900	14.906	40.783	<.001
		Congruency	0.205	0.025	8.239	<.001	83.885	12.961	6.472	<.001
		PC	-0.031	0.036	-0.860	.393	-13.476	14.936	-0.902	.370
		Prior RT	0.208	0.015	14.234	<.001	0.135	0.015	9.012	<.001
		Congruency × PC	0.084	0.013	6.318	<.001	37.133	5.957	6.233	<.001
		Congruency × Prior RT	-0.153	0.020	-7.827	<.001	-0.024	0.021	-1.167	.243
Gonthier et al. (2016) 1a & 1b	Simple model	Intercept	-1.488	0.026	-58.121	<.001	699.477	15.539	45.013	<.001
		Congruency	0.154	0.021	7.337	<.001	87.699	13.525	6.484	<.001
		PC	-0.039	0.005	-7.444	<.001	-18.784	3.378	-5.561	<.001
		Congruency × PC	0.064	0.008	8.369	<.001	27.986	4.913	5.696	<.001
	Temporal Learning model	Intercept	-1.496	0.023	-64.046	<.001	696.609	14.729	47.295	<.001
		Congruency	0.158	0.021	7.583	<.001	87.085	13.499	6.451	<.001
		PC	-0.023	0.005	-4.335	<.001	-13.551	3.374	-4.017	<.001
		Prior RT	0.179	0.009	20.515	<.001	0.115	0.010	11.402	<.001
		Congruency × PC	0.058	0.008	7.560	<.001	30.755	4.901	6.266	<.001
		Congruency × Prior RT	-0.080	0.011	-6.934	<.001	0.045	0.014	3.248	.001

Note. Est. = estimate of regression coefficient; *SE* = standard error; *t* = t-value; *p* = significance.

Table 2

Analyses of Congruency × PC interaction for the Simple Model and Temporal Learning Model, Untransformed RT using GLMM, in the Hutchison (2011), Bugg (2014), and Gonthier et al. (2016) Datasets

			Untransformed RT, GLMM (gamma distribution, identity link)			
		Variable	Est.	*SE*	*t*	*p*
Hutchison (2011)	Simple model	Intercept	672.855	2.840	236.934	<.001
		Congruency	84.325	7.299	11.552	<.001
		PC	-1.727	2.532	-0.682	.495
		Congruency × PC	22.316	2.540	8.786	<.001
	Temporal Learning model	Intercept	670.270	1.999	335.283	<.001
		Congruency	87.861	1.953	44.997	<.001
		PC	3.213	2.613	1.230	.219
		Prior RT	0.115	0.008	14.876	<.001
		Congruency × PC	24.989	2.345	10.657	<.001
		Congruency × Prior RT	0.040	0.013	3.082	0.002
Bugg (2014) 1a & 2b	Simple model	Intercept	633.611	7.218	87.788	<.001
		Congruency	74.710	15.707	4.756	<.001
		PC	-18.932	5.976	-3.168	.002
		Congruency × PC	39.090	4.909	7.963	<.001
	Temporal Learning model	Intercept	630.142	7.162	87.984	<.001
		Congruency	74.228	5.965	12.443	<.001
		PC	-13.380	7.292	-1.835	.067
		Prior RT	0.136	0.014	9.677	<.001
		Congruency × PC	37.819	4.422	8.553	<.001
		Congruency × Prior RT	-0.032	0.020	-1.583	.113
Gonthier et al. (2016) 1a & 1b	Simple model	Intercept	725.373	2.411	300.890	<.001
		Congruency	71.795	2.512	28.580	<.001
		PC	-16.550	1.654	-10.000	<.001
		Congruency × PC	30.593	1.993	15.350	<.001
	Temporal Learning model	Intercept	720.472	2.770	260.094	<.001
		Congruency	73.567	2.189	33.613	<.001
		PC	-11.771	1.686	-6.983	<.001
		Prior RT	0.125	0.009	13.239	<.001
		Congruency × PC	33.074	1.933	17.112	<.001
		Congruency × Prior RT	0.034	0.013	2.510	.012

Note. Est. = estimate of regression coefficient; *SE* = standard error; *t* = t-value; *p* = significance.

Table 3

Overview of Key Analyses Examining Temporal Learning Account with Transformed and Untransformed RT, in the Hutchison (2011), Bugg (2014), and Gonthier et al. (2016) Datasets

	Transformed RT, LME			Untransformed RT, GLMM		
	Hutchison (2011)	Bugg (2014) 1a &2b	Gonthier et al. (2016) 1a &1b	Hutchison (2011)	Bugg (2014) 1a &2b	Gonthier et al. (2016) 1a &1b
Beta changes in Congruency x PC	yes	yes	yes	no	yes	no
% of variance explained by Congruency × PC	no	yes	yes	no	no	yes†
AIC / BIC	yes	yes	yes	no	yes	no
Three-way Congruency × PC × Prior RT	no	no	yes*	yes*	yes*	yes*
Adjusted R^2	no	yes	yes	no	no	no
Non-significant Congruency × PC interaction after controlling for temporal learning	no	no	no	no	no	no

Note. ‘yes’ (lighter fill) indicates a given analysis output was consistent with temporal learning account and ‘no’ (darker fill) indicates the output was not consistent with temporal learning account. † In this comparison, the percentage of variance explained by the critical congruency by PC interaction decreased when prior RT was added to the model (supportive of the temporal learning account), but the actual beta of the congruency by PC interaction increased (not supportive of the temporal learning account). * indicates that a significant three-way interaction was observed, and while we interpret a significant interaction here as being generally supportive of the temporal learning account, the direction (form) this interaction should take is unclear so our interpretation (“yes” vs. “no”) does not consider the direction (please see Footnote 14 for detailed information about the direction of the interaction).

Table 4

Analyses of Percentage of Variance Explained by the Congruency × PC Interaction in the Simple Model and Temporal Learning Model Using Transformed and Untransformed RT, in the Hutchison (2011), Bugg (2014), and Gonthier et al. (2016) Datasets

		Transformed RT, LME				Untransformed RT, GLMM			
		Estimate	*t*	*p*	**% of variance explained**	Estimate	*t*	*p*	**% of variance explained**
Hutchison (2011)	Simple	0.057	6.202	<.001	**6.11%**	22.316	8.786	<.001	**9.22%**
	Temporal learning	0.049	5.345	<.001	**6.32%**	24.989	10.657	<.001	**29.47%**
Bugg (2014) 1a &2b	Simple	0.099	7.458	<.001	**7.57%**	39.090	4.909	<.001	**8.68%**
	Temporal learning	0.084	6.318	<.001	**7.42%**	37.819	8.553	<.001	**15.16%**
Gonthier et al. (2016) 1a & 1b	Simple	0.064	8.369	<.001	**7.22%**	30.593	15.35	<.001	**23.02%**
	Temporal learning	0.058	7.560	<.001	**6.81%**	33.074	17.112	<.001	**19.62%**

Note. Estimate = regression coefficient; *SE* = standard error; *t* = t-value; *p* = significance.

Table 5

Analyses of Changes in AIC and BIC Model Fit Statistics by Adding the Congruency × PC Interaction, With and Without Prior RT, in the Hutchison (2011), Bugg (2014), and Gonthier et al. (2016) Datasets

		Fixed effect factors	Transformed RT, LME				Untransformed RT, GLMM			
			AIC	Δ AIC	BIC	Δ BIC	AIC	Δ AIC	BIC	Δ BIC
Hutchison (2011)	Model 1	Cong + PC	4143	36	4189	29	192763	23	192808	15
	Model 2	Cong + PC + **Cong:PC**	4107		4160		192740		192793	
	Model 3	Cong + PC + Prior RT + Prior RT: Cong	3563	27	3624	19	192393	29	192454	21
	Model 4	Cong + PC + Prior RT + Prior RT: Cong **+ Cong:PC**	3536		3605		192364		192433	
Difference in Δ AIC/BIC (Model 4 – Model 3) vs. (Model 2 – Model 1)				-9		-10		+6		+6
Bugg (2014) 1a &2b	Model 1	Cong + PC	360	53	406	47	63789	49	63835	43
	Model 2	Cong + PC + **Cong:PC**	307		359		63740		63792	
	Model 3	Cong + PC + Prior RT + Prior RT: Cong	151	38	210	31	63670	46	63729	39
	Model 4	Cong + PC + Prior RT + Prior RT: Cong **+ Cong:PC**	113		179		63624		63690	
Difference in Δ AIC/BIC (Model 4 – Model 3) vs. (Model 2 – Model 1)				-15		-16		-3		-4
Gonthier et al. (2016) 1a & 1b	Model 1	Cong + PC	-1780	68	-1727	61	167852	49	167905	42
	Model 2	Cong + PC + **Cong:PC**	-1848		-1788		167803		167863	
	Model 3	Cong + PC + Prior RT + Prior RT: Cong	-2251	55	-2183	48	167521	58	167589	51
	Model 4	Cong + PC + Prior RT + Prior RT: Cong **+ Cong:PC**	-2306		-2231		167463		167538	
Difference in Δ AIC/BIC (Model 4 – Model 3) vs. (Model 2 – Model 1)				-13		-13		+9		+9

Note. Cong = congruency effect; Cong:PC = Congruency × PC interaction; Negative signs for difference in Δ AIC/BIC indicate that the degree to which the model fit improved by adding the Congruency × PC interaction is smaller when Prior RT and Prior RT × Congruency are already included in the model, supporting temporal learning account.

Table 6

Analyses of Congruency × PC × Prior RT Interaction using Transformed and Untransformed RT, in the Hutchison (2011), Bugg (2014), and Gonthier et al. (2016) Datasets

		Transformed RT, LME				Untransformed RT, GLMM			
	Variable	Est.	*SE*	*t*	*p*	Est.	*SE*	*t*	*p*
Hutchison (2011)	Intercept	-1.673	0.025	-66.724	<.001	669.745	3.356	199.564	<.001
	Congruency	0.246	0.023	10.536	<.001	87.677	3.151	29.681	<.001
	PC	0.011	0.024	0.485	.628	3.099	3.151	0.983	.325
	Prior RT	0.185	0.011	17.579	<.001	60.089	2.629	22.853	<.001
	Congruency × PC	0.048	0.009	5.276	<.001	24.816	2.857	8.686	<.001
	Congruency × Prior RT	-0.091	0.016	-5.745	<.001	5.852	2.330	2.511	.012
	PC × Prior RT	0.020	0.016	1.214	.225	-3.742	2.686	-1.393	.164
	Congruency × PC × Prior RT	-0.026	0.025	-1.048	.295	9.555	3.974	2.404	.016
Bugg (2014) 1a & 2b	Intercept	-1.708	0.032	-53.402	<.001	631.124	5.206	121.240	<.001
	Congruency	0.204	0.025	8.181	<.001	73.781	5.149	14.329	<.001
	PC	-0.031	0.036	-0.859	.393	-12.698	5.623	-2.258	.024
	Prior RT	0.189	0.021	8.968	<.001	52.389	3.917	13.374	<.001
	Congruency × PC	0.084	0.013	6.326	<.001	36.371	4.005	9.081	<.001
	Congruency × Prior RT	-0.121	0.028	-4.335	<.001	-14.760	4.483	-3.292	.096
	PC × Prior RT	0.038	0.029	1.285	.199	7.306	4.382	1.667	<.001
	Congruency × PC × Prior RT	-0.063	0.039	-1.622	.105	-11.740	4.985	-2.355	.019
Gonthier et al.(2016) 1a & 1b	Intercept	-1.496	0.023	-63.874	<.001	719.642	2.410	298.662	<.001
	Congruency	0.156	0.021	7.444	<.001	74.667	2.374	31.452	<.001
	PC	-0.023	0.005	-4.357	<.001	-10.748	1.952	-5.506	<.001
	Prior RT	0.163	0.012	13.706	<.001	66.943	2.132	31.393	<.001
	Congruency × PC	0.057	0.008	7.530	<.001	29.953	2.349	12.755	<.001
	Congruency × Prior RT	-0.032	0.016	-1.972	.049	28.011	2.265	12.366	<.001
	PC × Prior RT	0.031	0.016	1.980	.048	0.100	2.122	0.047	.963
	Congruency × PC × Prior RT	-0.096	0.023	-4.184	<.001	-39.918	2.727	-14.639	<.001

Note. Est. = estimate of regression coefficient; *SE* = standard error; *t* = t-value; *p* = significance.

Table 7

Analyses of Changes in R^2 by Adding the Congruency × PC Interaction, With and Without Prior RT, in the Hutchison (2011), Bugg (2014), and Gonthier et al. (2016) Datasets

		Fixed effect factors	Transformed RT, LME		Untransformed RT, GLMM	
			Adj.R^2 [95 % CI]	Difference [95 % CI]	Adj.R^2 [95 % CI]	Difference [95 % CI]
Hutchison (2011)	Model 1	Cong + PC	0.123 [0.114, 0.133]	**0.00070** [0.00058, 0.00091]	0.099 [0.090, 0.108]	**0.00021** [0.00016, 0.00036]
	Model 2	Cong + PC + **Cong:PC**	0.124 [0.114, 0.133]		0.099 [0.090,0.108]	
	Model 3	Cong + PC + Prior RT + Prior RT: Cong	0.165 [0.155, 0.175]	**0.00070** [0.00044, 0.00104]	0.130 [0.120, 0.140]	**0.00112** [0.00099, 0.00135]
	Model 4	Cong + PC + Prior RT + Prior RT: Cong + **Cong:PC**	0.166 [0.156, 0.176]		0.131 [0.121, 0.141]	
Bugg (2014) 1a &2b	Model 1	Cong + PC	0.143 [0.126,0.160]	**0.00563** [0.00532, 0.00620]	0.131 [0.115, 0.148]	**0.00454** [0.00433, 0.00503]
	Model 2	Cong + PC + **Cong:PC**	0.148 [0.132,0.166]		0.135 [0.119,0.153]	
	Model 3	Cong + PC + Prior RT + Prior RT: Cong	0.183 [0.165, 0.201]	**0.00407** [0.00353, 0.00479]	0.156 [0.139, 0.174]	**0.00534** [0.00499, 0.00596]
	Model 4	Cong + PC + Prior RT + Prior RT: Cong + **Cong:PC**	0.187 [0.169, 0.206]		0.161 [0.145, 0.180]	
Gonthier et al. (2016) 1a & 1b	Model 1	Cong + PC	0.098 [0.089, 0.108]	**0.00291** [0.00283, 0.00310]	0.082 [0.073, 0.091]	**0.00151** [0.00148, 0.00167]
	Model 2	Cong + PC + **Cong:PC**	0.101 [0.092, 0.111]		0.083 [0.075, 0.092]	
	Model 3	Cong + PC + Prior RT + Prior RT: Cong	0.137 [0.126, 0.147]	**0.00251** [0.00234, 0.00278]	0.110 [0.101, 0.120]	**0.00284** [0.00273, 0.00307]
	Model 4	Cong + PC + Prior RT + Prior RT: Cong + **Cong:PC**	0.139 [0.129, 0.150]		0.113 [0.103, 0.123]	

Note. The 95% CI of difference in R^2 was calculated based on a simulation approach by using the package r2glmm (Jaeger, 2016) in R. Gray fill indicates comparison between difference of 95% CIs; darker fill indicates non-overlapping CIs and lighter fill indicates overlapping CIs.

Table 8

Analyses of Changes in R^2 for the Original Temporal Learning Model, Extended Temporal Learning Model, and Weighted Extended Temporal Learning Model, Using Transformed and Untransformed RT, in the Hutchison (2011), Bugg (2014), and Gonthier et al. (2016) Datasets

		Transformed RT, LME		Untransformed RT, GLMM	
		Adj.R^2	R^2 Difference From Original TL model	Adj.R^2	R^2 Difference From Original TL model
	Model	[95 % CI]	[95 % CI]	[95 % CI]	[95 % CI]
Hutchison (2011)	Original TL model	0.163 [0.152,0.174]		0.133 [0.123,0.144]	
	Extended TL model	0.219 [0.207,0.231]	0.055^{***} [0.055,0.056]	0.168 [0.157,0.179]	0.035^{***} [0.034,0.035]
	Weighted extended TL model	0.203 [0.192,0.215]	0.040^{***} [0.039,0.041]	0.162 [0.151,0.173]	0.029^{***} [0.028,0.029]
Bugg (2014) 1a &2b	Original TL model	0.198 [0.177,0.222]		0.169 [0.149,0.192]	
	Extended TL model	0.237 [0.215,0.261]	0.038^{***} [0.037,0.039]	0.204 [0.183,0.228]	0.035^{***} [0.033,0.036]
	Weighted extended TL model	0.229 [0.207,0.253]	0.030^{***} [0.029,0.031]	0.196 [0.175,0.220]	0.027^{***} [0.026,0.028]
Gonthier et al. (2016) 1a & 1b	Original TL model	0.136 [0.125,0.147]		0.111 [0.101,0.121]	
	Extended TL model	0.171 [0.160,0.183]	0.035^{***} [0.034,0.036]	0.138 [0.127,0.149]	0.027^{***} [0.026,0.028]
	Weighted extended TL model	0.162 [0.151,0.174]	0.026^{***} [0.026,0.027]	0.132 [0.122,0.143]	0.022^{***} [0.021,0.022]

Note. TL = temporal learning; Original and Extended TL models are non-nested and therefore statistical testing of comparison between those two models was based on simulation using r2glmm package (Jaeger, 2016) in R. Asterisks in Difference column indicate R^2 for Extended or Weighted extended model is greater than that of Original TL model, suggesting better model fit. ***: p <.001, *: p <.05.

Table 9

Analyses of Congruency × PC interaction for the Simple, Extended Temporal Learning, and Weighted Extended Temporal Learning Model, Using Transformed and Untransformed RT, in the Hutchison (2011), Bugg (2014), and Gonthier et al. (2016) Datasets

			Hutchison (2011)				Bugg (2014) 1a & 1b				Gonthier et al. (2014) 1a & 1b			
		Variable	Est.	SE	*t*	*p*	Est.	SE	*t*	*p*	Est.	SE	*t*	*p*
Transformed RT, LME	Simple Model	Intercept	-1.665	0.027	-67.687	<. 001	-1.697	0.035	-48.155	<. 001	-1.488	0.026	-58.190	<. 001
		Congruency	0.234	0.024	9.777	<. 001	0.195	0.025	7.870	<. 001	0.152	0.021	7.168	<. 001
		PC	-0.006	0.028	-0.200	.842	-0.058	0.042	-1.392	.168	-0.040	0.006	-7.372	<. 001
		Congruency × PC	0.053	0.010	5.494	<. 001	0.105	0.014	7.468	<. 001	0.064	0.008	8.008	<. 001
	Extended TL model	Intercept	-1.664	0.027	-67.873	<. 001	-1.714	0.030	-57.663	<. 001	-1.501	0.022	-67.126	<. 001
		Congruency	0.234	0.024	9.827	<. 001	0.208	0.025	8.291	<. 001	0.158	0.021	7.551	<. 001
		PC	-0.004	0.028	-0.141	.888	-0.024	0.031	-0.766	.446	-0.014	0.005	-2.474	.013
		Mean Prior RT	0.218	0.015	14.380	<. 001	0.373	0.022	16.906	<. 001	0.311	0.012	25.058	<. 001
		Congruency × PC	0.052	0.010	5.427	<. 001	0.082	0.014	5.918	<. 001	0.053	0.008	6.690	<. 001
		Congruency× Mean Prior RT	-0.167	0.026	-6.472	<. 001	-0.243	0.027	-9.005	<. 001	-0.135	0.015	-8.971	<. 001
	Weighted Extended TL model	Intercept	-1.677	0.025	-67.388	<. 001	-1.712	0.031	-55.842	<. 001	-1.501	0.023	-65.863	<. 001
		Congruency	0.245	0.025	10.023	<. 001	0.206	0.027	7.704	<. 001	0.158	0.021	7.489	<. 001
		PC	0.021	0.021	1.020	.309	-0.025	0.032	-0.783	.436	-0.016	0.005	-2.890	.004
		Mean Prior RT	0.326	0.011	28.784	<. 001	0.331	0.022	14.987	<. 001	0.275	0.011	24.027	<. 001
		Congruency × PC	0.039	0.010	4.134	<. 001	0.085	0.016	5.349	<. 001	0.054	0.008	6.880	<. 001
		Congruency× Mean Prior RT	-0.146	0.016	-9.275	<. 001	-0.203	0.028	-7.231	<. 001	-0.117	0.014	-8.209	<. 001
Untransformed RT, GLMM	Simple Model	Intercept	669.756	2.564	261.231	<. 001	634.861	7.265	87.390	<. 001	726.272	2.764	262.72	<. 001
		Congruency	84.387	3.245	26.002	<. 001	71.886	9.624	7.469	<. 001	69.930	2.419	28.91	<. 001
		PC	1.148	2.414	0.476	.634	-21.702	8.592	-2.526	.012	-17.045	1.676	-10.17	<. 001
		Congruency × PC	19.892	2.715	7.327	<. 001	43.362	4.994	8.683	<. 001	30.507	2.287	13.34	<. 001
	Extended TL model	Intercept	668.934	2.951	226.711	<. 001	627.749	5.793	108.355	<. 001	719.302	4.449	161.665	<. 001
		Congruency	84.599	3.247	26.052	<. 001	72.856	5.091	14.311	<. 001	70.170	2.390	29.366	<. 001
		PC	8.053	2.957	2.723	.006	-11.647	5.448	-2.138	.033	-8.608	2.190	-3.931	<. 001
		Mean Prior RT	0.242	0.012	19.836	<. 001	0.254	0.235	10.835	<. 001	0.242	0.014	17.715	<. 001
		Congruency × PC	21.982	2.701	8.137	<. 001	41.602	5.244	7.934	<. 001	33.500	3.520	9.516	<. 001
		Congruency× Mean Prior RT	0.039	0.018	2.179	.029	-0.038	0.031	-1.237	.216	0.032	0.018	1.825	.068
	Weighted Extended TL model	Intercept	664.471	2.269	292.903	<. 001	628.362	6.703	93.737	<. 001	719.676	8.001	89.954	<. 001
		Congruency	89.435	2.424	36.897	<. 001	72.882	6.873	10.604	<. 001	70.230	8.487	8.275	<. 001
		PC	8.117	2.636	3.080	.002	-12.094	7.325	-1.651	.099	-9.304	3.076	-3.025	.003
		Mean Prior RT	0.213	0.001	18.899	<. 001	0.225	0.021	10.496	<. 001	0.212	0.013	15.849	<. 001
		Congruency × PC	22.090	2.293	9.635	<. 001	41.897	4.925	8.507	<. 001	33.684	4.818	6.992	<. 001
		Congruency× Mean Prior RT	0.052	0.017	2.979	.003	-0.030	0.029	-1.039	.299	0.036	0.018	1.968	.049

Note. Est. = estimate of regression coefficient; *SE* = standard error; *t* = t-value; *p* = significance; Mean Prior RT = average of three prior reaction times.

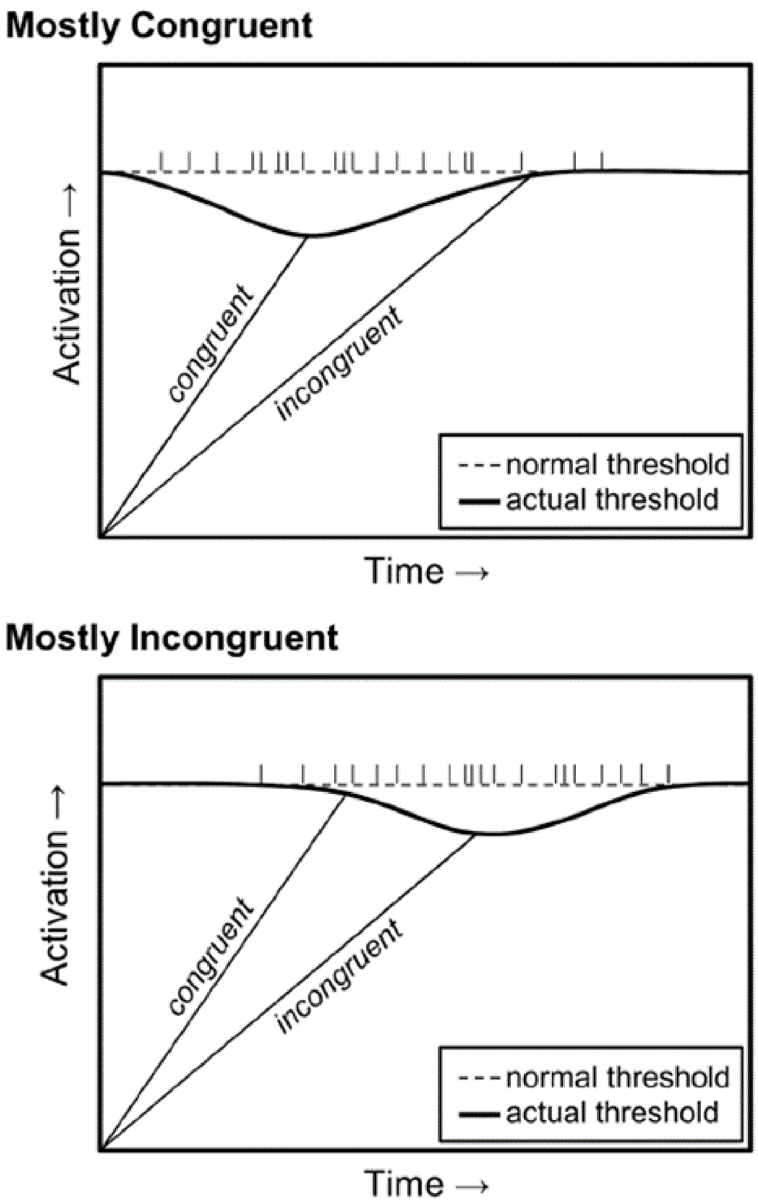

Figure 1. From Schmidt (2013), visual depiction of the temporal learning account.

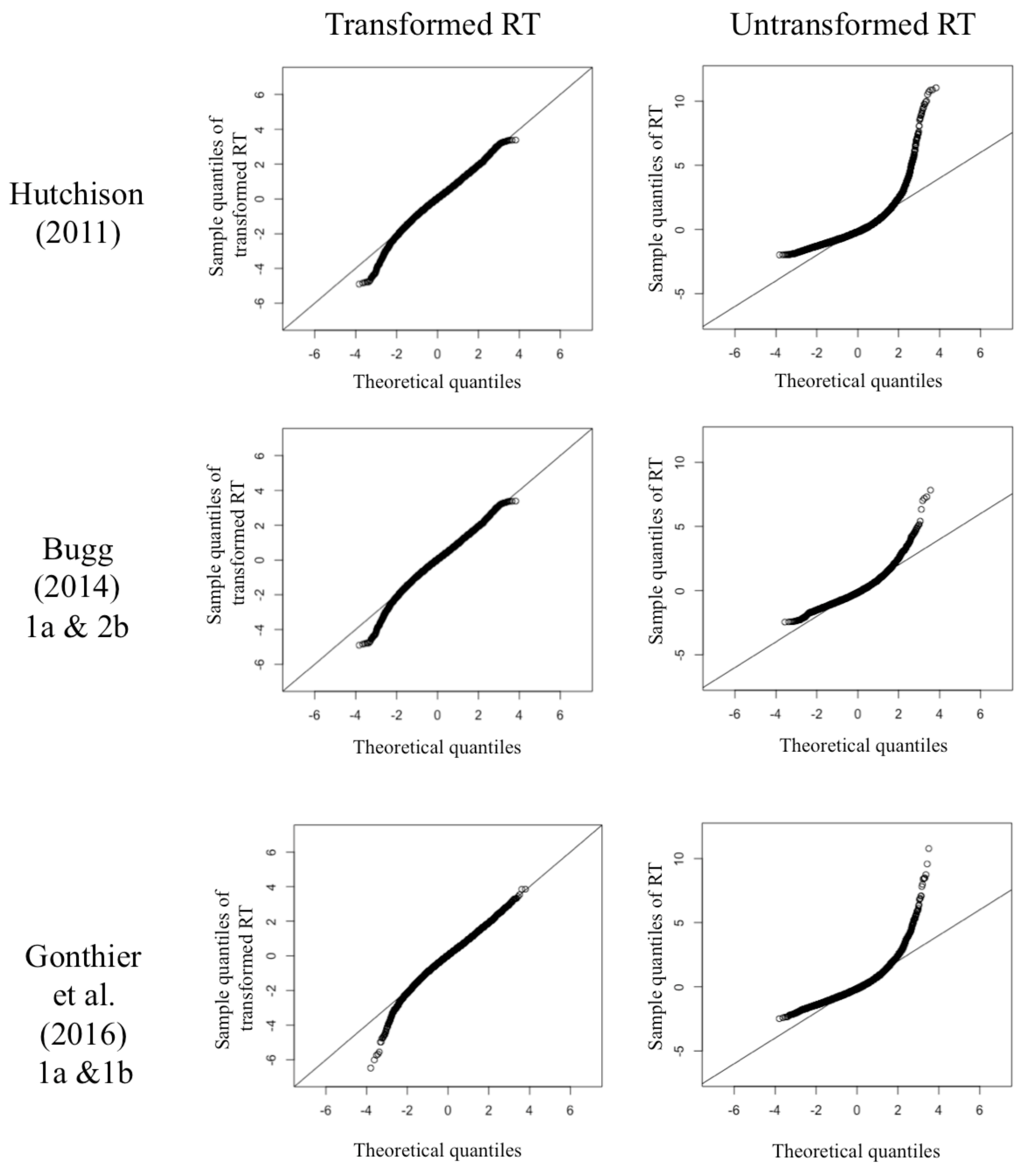

Figure 2. Q-Q plots of untransformed and transformed RT, using a cutoff of RTs > 300, from Hutchison (2011), Bugg (2014), and Gonthier et al. (2016) datasets

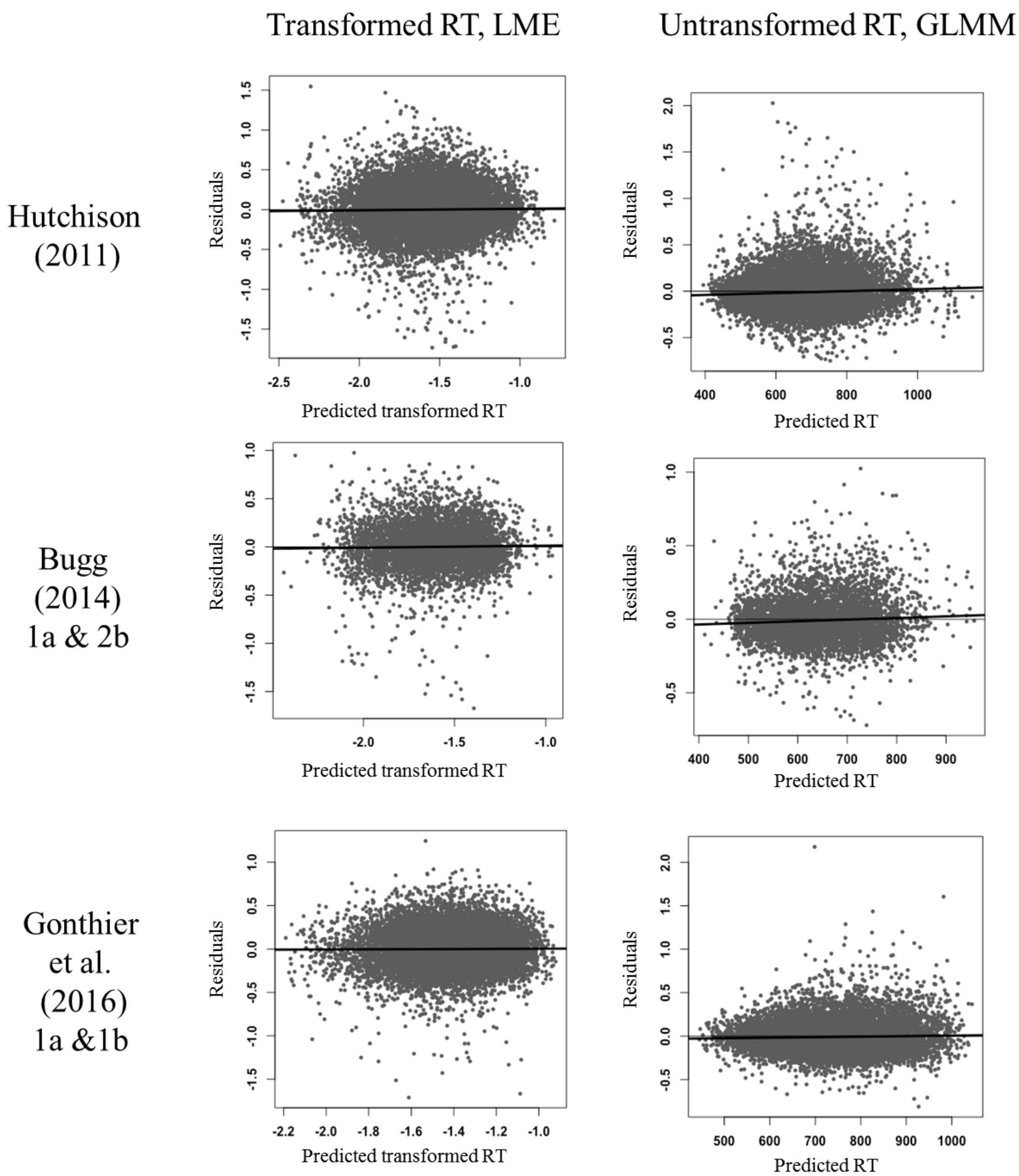

Figure 3. Residual plots of a full model (PC × Congruency × Prior RT with untransformed and transformed RT, using a cutoff of RTs > 300, from Hutchison (2011), Bugg (2014), and Gonthier et al. (2016) datasets.