The Role of Contingency and Correlation in the Stroop Task

Nabil Hasshim[1] & Benjamin A. Parris[2]

[1]De Montfort University & [2]Bournemouth University

Author Note

Ethical approval was obtained from the research ethics committee of De Montfort University faculty of Health and Life Sciences (Ref: 3427). All materials, raw data, analysis, approved stage 1 protocol and documents related to this project are publicly available at https://osf.io/tfeq6/

Orcid:

Nabil Hasshim: https://orcid.org/0000-0001-6375-8554

Benjamin Parris: https://orcid.org/0000-0003-2402-2100

Corresponding author:

Nabil Hasshim: nabil.hasshim@dmu.ac.uk

DOI: 10.1177/17470218211032548

Abstract

Facilitation (faster responses to Congruent trials compared to Neutral trials) in the Stroop task has been a difficult effect for models of cognitive control to explain. The current research investigated the role of word-response contingency, word-colour correlation, and proportion congruency in producing Stroop effects. Contingency and correlation refers to the probability of specific word-response and word-colour pairings that are implicitly learnt while performing the task. Pairs that have a higher probability of occurring are responded to faster, a finding that challenges top-down attention control accounts of Stroop task performance. However studies that try to experimentally control for contingency and correlation typically do so by increasing the proportion of incongruent trials in the task, which cognitive control accounts posit affects interference control via the top-down biasing of attention. The present research focused on whether facilitation is also affected by contingency and correlation while additionally looking at the effect of proportion congruency. This was done in two experiments that compared the typical design of Stroop task experiments (i.e., having equal proportions of Congruent and Incongruent trials but also contingency and correlational biases) to: a) a design that had unequal congruency proportions but no contingency or correlation (Experiment 1), and b) a design where the correlation is biased but proportion congruency and contingency were not (Experiment 2). Results did not support the hypotheses that contingency or correlation affected facilitation. Interference was almost halved in the alternative design of Experiment 2, demonstrating an effect of contingency learning in typical measures of Stroop interference.

Keywords: *Cognitive control, Contingency learning, Facilitation, Stroop task*

DOI: 10.1177/17470218211032548

The Role of Contingency and Correlation in Stroop Task Facilitation

The Stroop task (Klein, 1964; Stroop, 1935), probably the most widely used paradigm in selective attention research (MacLeod 1992), requires participants to respond to the 'ink' colour of individual words, while ignoring what the word spells out. The efficiency in performing the task is influenced by the property of the word, with the classic finding being that responses are fastest when the word and colour are *Congruent* (e.g., the word 'blue' printed in blue), and slower when they are *Incongruent* (e.g., 'blue' printed in green). When the word does not evoke a colour (e.g., 'table' presented in red) or is made up of a string of letters or symbols (e.g., 'xxxx' in blue, or '#####' in red), the time taken to respond to these *neutral* trials is typically between that of Congruent and Incongruent trials. The difference in performance between Congruent and Neutral trials is often taken as a measure of *facilitation*, while the difference between neutral and Incongruent trials is taken as a measure of *interference* (see MacLeod, 1991, and Parris, Hasshim, Wadsley, Augustinova & Ferrand, submitted, for comprehensive reviews).

An interesting and consistent finding in the Stroop literature for which models have attempted to account is that facilitation effects are less stable, less reliable and are generally much smaller than interference (e.g., Glaser & Glaser, 1982; Lindsay & Jacoby, 1994; MacLeod, 1998; Augustinova, Parris & Ferrand, 2019; but see Melara & Algom, 2003), and can be absent or even reversed (e.g., Dalrymple-Alford, 1972; Goldfarb & Henik, 2007; Kalanthroff & Henik, 2013).

Converging information hypothesis of facilitation

Extant models of Stroop performance (e.g., models by Cohen, Dunbar, & McClelland, 1990; Melara & Algom; 2003; Phaf, Van der Heijden, & Hudson, 1990; Roelofs 2003) describe facilitation and interference as stemming from the same

DOI: 10.1177/17470218211032548

mechanisms. These models posit that information from the colour and word dimensions converge on Congruent trials and diverge on Incongruent trials. Converging and diverging of information results in facilitation and interference respectively. Thus, in this view, the information from the word dimension of a Congruent trial aids in stimulus processing and thus improves task performance on that trial since it converges with information from the colour (i.e., both word and colour provide evidence towards the same response). Smaller facilitation effects are accounted for within the context of the parallel distributed processing models of Stroop task performance (e.g. Cohen, Dunbar, & McClelland, 1990).

Inadvertent reading

An alternative account to the converging information account of facilitation is the inadvertent reading hypothesis (Dunbar & MacLeod, 1984; Kane & Engle, 2003; MacLeod & MacDonald, 2000). This account postulates that on some trials, participants fail in the goal of ignoring the word and inadvertently respond to the meaning of the word. When this happens on incongruent trials it results in an incorrect response (which are ignored in analyses of correct response latencies), but on Congruent trials, they manifest as a fast correct response since reading is generally faster than colour naming (MacLeod, 1991). As these trials are classified as correct trials, they are then included in the calculation of overall response times for Congruent trials, contributing to the measured facilitation effect. Therefore, individual differences in participants' tendencies for goal failure has an effect on the calculation of congruent trials, but not incongruent trials, which can explain the inconsistency in the measurement of Stroop facilitation in the literature and also the asymmetrical magnitude of the two effects (MacLeod & MacDonald, 2000).

DOI: 10.1177/17470218211032548

The converging information and inadvertent reading accounts of facilitation both provide reasonable accounts of existing data. However, there is reason to believe that facilitation effects might be smaller than originally thought; indeed, there is reason to believe that facilitation effects in their entirety are the result of an experimental confound present in all previous experiments. If this were shown to be true, the above accounts of Stroop facilitation would be redundant.

Colour-response contingency and colour-word correlation

The role of contingency learning in the Stroop task has been highlighted by Schmidt and colleagues (e.g., Schmidt, 2013; Schmidt 2016; Schmidt, Crump, Cheesman, & Besner, 2007; Schmidt & De Houwer, 2012). Contingency learning refers to how the probability of each word-response pairing is implicitly learnt and subsequently used to predict a response upon further encounters with the word. For example, if the word 'green' is more often presented in red, the response to red will be predicted whenever the word 'green' is encountered in the future, facilitating responses when this prediction is correct (Schmidt, et al., 2007). In typical Stroop task designs involving Congruent and Incongruent trials, an equal number of each type of trial is displayed while presenting each possible word-colour combination (Dishon-Berkovits & Algom, 2000). The left half of Figure 1 (labelled "standard design") shows an example of the frequency of each colour and word combination in such a design. Equal numbers of Congruent (italicised numbers) and Incongruent trials leads to a higher frequency of each word being displayed in its corresponding colour compared to another colour (e.g., nine instances of 'yellow' in yellow vs three instances of 'yellow' in blue), further speeding responses to congruent stimuli. Thus, even though the number of Congruent and Incongruent trials are equal, measures of Stroop facilitation will have been unintentionally inflated by contingency effects.

DOI: 10.1177/17470218211032548

Another factor that is confounded in such a design is colour-word *correlation* (Dishon-Berkovits & Algom, 2000; Melara & Algom, 2003; Algom & Chajut, 2019). The idea of correlation is similar to the concept of contingency, with the distinction being that while the latter refers to word-response pairings, the former refers to the pairings between the colours and words. However, the mechanisms underlying the two are different. Contingency learning posits that the specific word-response association results in faster responses, a purely associative learning account. On the other hand, the correlation account refers to the perceived probability of each irrelevant word would appear in each available colour. If the perceived probability of a word appearing in any of the available colours is not random then a colour-word correlation is created (Dishon-Berkovits & Algom, 2000). In such cases, the identity of the word can become a reliable source of information in predicting what the colour will be. Thus, it would be beneficial for the cognitive system to allocate some attentional resource to the word reading task, instead of ignoring the word dimension as instructed. As a result a large Stroop effect ensues.

Indeed, Dishon-Berkovits and Algom (2000) repeatedly showed that the Stroop effect (difference between Incongruent and Congruent) was eliminated in designs where the two dimensions making up the target were randomly selected (i.e., zero correlation). Along with other studies like that of Schmidt and Besner (2008), this is evidence that Stroop interference can be explained by the design of the tasks which confound correlation and contingency. The present study aims to extend this by focusing on facilitation, which has not been a focus of studies in the literature and also using the classic colour-word Stroop task (Dishon-Berkovits & Algom, 2000 used variants of the Stroop task, the spatially separated word-word task and the picture-word task).

DOI: 10.1177/17470218211032548

[Figure 1 somewhere here]

As mentioned earlier, one way for empirical studies to control contingency and correlation effects is by ensuring each word-colour combination occurs equally often (e.g., De Houwer, 2003; Dishon-Berkovits & Algom, 2000; Hasshim & Parris, 2014; 2015; Schmidt & Cheesman, 2005). The right half of Figure 1 (labelled “alternative design”) shows an example frequency table of such a manipulation. However, this also means that the number of Congruent and Incongruent trials would not be equal (e.g., there are two times more Incongruent trials than Congruent trials). Thus, it is unclear whether this imbalance in the number of Congruent and Incongruent trials in the experiment would have any influence on any measured effect (this idea, *Proportion Congruency* is discussed in the next section). In two of the studies that included a suitable neutral condition to enable the computation of Stroop facilitation, and controlled for response contingency, Hasshim and Parris (2014) reported no facilitation effects in a manual response paradigm, while Hasshim and Parris (2015) reported facilitation effects using an oculomotor response paradigm.

Proportion congruency

When the design of the Stroop task is manipulated as described above, other interrelated factors are affected as well. Proportion congruency refers to the proportion of trial types making up an experiment. As described earlier, researchers typically strive to have equal numbers of each trial type in their experiments, which then affects contingency and correlation. Controlling for contingency and correlation as shown in the alternative design depicted in Figure 1 would lead to an unequal number of each trial type in the experiment.

Having different proportion of trial types is in fact one key manipulation in studies demonstrating strategic control of attention (e.g., Cheesman & Merikle, 1986;

DOI: 10.1177/17470218211032548

Lindsay & Jacoby, 1994; Logan & Zbrodoff, 1979; West & Baylis, 1998). In these studies, the list-wise proportion congruency is manipulated by administering blocks made up mostly Congruent or mostly Incongruent trials. These studies demonstrate that the Stroop effect is larger when the block is made up of mostly (typically ~80%) Congruent trials compared to blocks with mostly Incongruent trials. The explanation given for this phenomenon was one of strategic top-down control. When most trials encountered are Congruent, attentional resources are biased towards word reading since the word is predictive of the correct response. This results in faster responses to Congruent trials and inflates the measurement of the Stroop effect in those blocks.

However, it has been argued that the resultant biased contingency of such blocks might explain the proportion congruency effect (see Bugg, Jacoby, & Channani, 2011; Schmidt & Besner, 2008; and Schmidt & Lemercier, 2018, for in-depth discussions of this issue). A popular paradigm in exploring this involves the use of two sets of stimuli with different word-response contingencies (item-level proportion congruency). For example Blais and Bunge (2010); Bugg et al. (2011); and Bugg, Jacoby, and Toth (2008), had one set of stimuli where contingency was controlled, while manipulating the global proportion of Congruent and Incongruent trials in the task by varying the number of such trials in a second set of stimuli. The results from these studies indeed show that item-level proportion congruency can account for the proportion congruency effect, suggesting that the effect is not due to a general task-level shift in attentional control.

Lorentz et al. (2016) had a similar research question to the current study and explored the effect of contingency on facilitation by utilising different baselines in two sets of stimuli in the same procedure. Congruent and Incongruent trials had a corresponding set of neutral trials, which matched their different contingencies, as

DOI: 10.1177/17470218211032548

their baseline and they showed that contingency indeed influences facilitation. Compared to their respective contingency-matched neutral conditions Lorentz et al. showed that both facilitation (~40ms) and interference (65ms) were significant. However, Lorentz et al. did not compare the contingency controlled effects to the more common non-contingency controlled design and so the magnitude of the contingency effect is not known. Moreover, they had participants respond with a vocal response in their study which produces larger Stroop effects (Augustinova et al., 2019; Parris et al., submitted) and it is therefore unknown what the magnitude of the contingency effect is with either response type and whether facilitation effects will remain with a manual response. Furthermore, their design did not control for colour-response correlation.

Finally, as with all manipulations, such techniques have limitations. The use of two sets of stimuli within an experiment necessitates the use of more response options than is typical (e.g., Bugg, 2014 used 8 colours while Lorentz et al., 2016 used 9 colours) which might not be practical for manual response tasks as remembering all the colour-button mappings induces greater memory load, and affects task performance. Furthermore, this technique reduces the number of trials that are used in measuring the effect of interest which reduces statistical power (Braem et al., 2019)

The current study

The current research aimed to study the effects of word-response contingency and word-colour correlation in a straightforward way. Experiment 1 compared the magnitude of facilitation in a standard Stroop task design against an alternative design where each word had an equal probability of appearing in each colour (see Figure 1). In the alternative design there were no contingencies between the words

DOI: 10.1177/17470218211032548

and responses and word-colour correlation was zero, but there were twice as many Incongruent trials as Congruent trials.

Experiment 2 compared the standard design to another alternative design in which the colour word of Incongruent trials appeared in only one specific colour throughout the task, matching the word-response contingency of Congruent trials (see Figure 2). This means that although there was a positive word-colour correlation, there was no word-response contingency and the number of Congruent and Incongruent trials were the same (a similar design was used in Hasshim & Parris, 2018). The two experiments allow for the investigation of the intertwining effects of word-response contingency, word-colour correlation, and proportion congruency on facilitation.

[Figure 2 somewhere here]

Theoretical implications

The main question being asked is what proportion of the Stroop facilitation effect is a by-product of the design of Stroop experiments which confound word-response contingency and word-colour correlation. The results will inform theoretical accounts of this long-established effect. Should there be no facilitation when contingency is controlled it would support the notion that facilitation is not a failure of cognitive control per se but a consequence of the computation of the statistical properties of the experimental context (Algom & Chajut, 2019; Schmidt 2013; 2019).

If facilitation effects are observed even when contingency and correlation are manipulated, we will have a better foundation from which to judge accounts of Stroop facilitation. The converging information and inadvertent reading hypotheses have different predictions as to how facilitation and interference effects manifest

DOI: 10.1177/17470218211032548

throughout the RT distributions (Roelofs, 2010), which was tested with the data obtained from this study.

Besides being informative in the ongoing discussion on how much the Stroop effect is a reflection of cognitive flexibility and associative learning, the two experiments might potentially be useful in understanding stimulus driven learning processes. As detailed earlier, the two stimulus driven accounts of interest work by slightly different mechanisms. The contingency learning account postulates a pure associative learning mechanism of specific word-response pairings while in the word-colour correlation account, performance is affected by individuals' perception that some pairings occur more often than the other possible pairings. Comparing the results of the two experiments will allow for the comparison of the two accounts.

Experiment 1

Method

Participants

60 individuals participated in Experiment 1. Participants were undergraduate psychology students and received course credit for their participation. Data from 10 participants were excluded as they did not meet the accuracy threshold of 90% correct answers overall as specified in the data exclusion criteria.

Prior to data collection, a target of 44 participants was set for the experiment. Data collection sessions were advertised until the day this number was reached. Additional participants who signed up on the last day were still able to participate.

The sample size was estimated based on the effect sizes (δs of 0.59, 0.63, and 0.78) obtained in the measures of facilitation of Hasshim and Parris (2015), using the jpower module of jamovi software (The jamovi project, 2019). The

DOI: 10.1177/17470218211032548

minimum desired power was specified as 0.9, with a minimally interesting effect size (δ) of 0.5, and type I error rate (α) of 0.05.

Furthermore the statistical power of the analyses will also be improved by having more trials per experiment compared to Hasshim and Parris (2015). Increasing the number of trials improves statistical power in psychophysics experiments like the Stroop task (Rouder & Haaf, 2018). The number of trials to be used was determined by the time it typically takes participants complete similar experiments in the lab and for each session to not exceed 30 minutes.

Apparatus

The experiment was programmed using PsychoPy software (Peirce et al., 2019) and data collection was conducted online via Pavlovia.org. Participants were instructed to perform the task using a desktop or laptop computer only. Responses were recorded via participants pressing the G, H, and J keys on their keyboard, which corresponded to one of the three possible colour responses.

Design

The experiment employed a 3 (Congruent, Neutral, & Incongruent) x 2 (standard design & alternative design) within participants design. Each participant went through blocks of trials from either the standard or alternative design (see Figure 1) first, before going through blocks from the other design. The order of this was randomly determined.

On each trial the properties of the target stimuli (its word and colour) was generated corresponding to the numbers in Figures 1 (e.g., in standard design blocks the number of each trial type will follow that of the left panel). Facilitation was

DOI: 10.1177/17470218211032548

calculated by the difference between Neutral and Congruent trials, while interference was calculated by the difference between Incongruent and Neutral trials.

Stimuli

Two sets of stimuli were used, with each set containing three colour words and three neutral words (see Table 1 for the words and colours used in each set and the lexical properties of the words). Participants encountered stimuli from one set in the first half of the experiment (either standard or alternative design) and the other set in the second half, the order of which was randomised.

To check that there are no carryover effects from the first half the experiment, supplementary analyses splitting participants by the order of presentation design were conducted to make sure that the pattern of results is consistent throughout the experiment.

[Table 1 somewhere here]

Procedure

Participants went through 6 blocks of trials as follows: a practice block of 24 trials, two experimental blocks of 135 trials each from one of the designs, another practice block of 24 trials, and two experimental blocks of 135 trials each from the second design. The resulting number of trials in each experiment was 588 (48 practice and 540 experimental trials). Practice trials consisted of hash symbols (e.g., ###, ######) displayed in the three response colours. Each of the experimental blocks consisted of Congruent, Incongruent, and Neutral trials.

On each trial, a grey fixation cross appeared at the centre of a black screen for 500ms, followed by the Stroop stimuli which stayed visible for 2500ms or until a response was made. If no response, or an incorrect response was made within

DOI: 10.1177/17470218211032548

2500ms, an additional feedback screen in the form of the text ‘incorrect’ or ‘no response’ was shown for 1500ms. The feedback was in black text over a grey background. A 1000ms blank black screen concluded each trial.

A break was administered after each block with participants allowed to take as much time as they wanted (minimum of 5 seconds) before initiating the next block by pressing the space bar.

Data exclusion criteria

Only correct responses >200 ms were analysed as fast responses are assumed to be anticipatory. Since the task is relatively easy and similar research conducted in our lab have shown participants’ performance to typically be ~95% accurate, data from any participants where < 90% of trials are valid were excluded. Error rates are not one of the main dependent variables of interest but were similarly analysed and reported.

Analysis plan

Within each experiment, a statistically significant (p <.05) difference between the facilitation effects between the two designs and a Bayes factor larger than 3 would be taken as support for the hypothesis that the measurement of facilitation is influenced by experiment design. Otherwise, it would be concluded that the hypothesis was not supported by the data, with a Bayes factor smaller than 0.33 (evidence for the null is 3 times that of the alternate hypothesis) indicating that correlation and contingency did not impact facilitation.

Facilitation was calculated for each participant by subtracting their mean response time (RT) on Congruent trials from their mean RT on Neutral trials, while interference was calculated from subtracting the mean RT of Neutral trials from that

DOI: 10.1177/17470218211032548

of Incongruent trials. An omnibus 2 (Stroop effects: facilitation & interference) x 2 (design: standard & alternative) ANOVA was conducted with the main effects of the Stroop effects indicating whether the facilitation and interference effects were observed. A statistically significant interaction would suggest that the different designs affect the measurement of Stroop effects, and this was explored in the following planned comparison.

Since the main research question was whether facilitation is significantly reduced in the alternative design compared to the standard design, a *t*-test was conducted comparing the size of facilitation between the standard and alternative design conditions. To complement the frequentist *t*-test comparison, a Bayes factor (BF) was calculated with the using the Dienes Bayes factor calculator (Dienes 2011, 2014), with the prior distribution defined as a half-normal distribution with a maximum probability at 0ms and standard deviation of 23ms. The value of the standard deviation is based on the raw saccade latency effect size of facilitation in Hasshim and Parris (2015). Using the estimated sample size and abovementioned prior distribution, along with the previous study's standard error of 6.61ms, a sensitive Bayes factor (>3 in favour of the theory) was estimated with a raw effect size of at least 12.8ms. In addition to each calculated BF, robustness regions were also reported to show the range of raw effect sizes where this criteria would be met. This would illustrate whether conclusions drawn from the BF is sensitive to the priors chosen.

If the facilitation effect, as typically observed in the literature using the standard design, was influenced by the confounding correlation and/or contingency in the design, then a statistically significant effect would be expected, showing the facilitation effect to be smaller or absent in the alternative design compared to the

DOI: 10.1177/17470218211032548

standard design. This result would suggest that task design influences the measurement of facilitation and that it should be something future studies need to consider. However, it is unclear from the current literature what the effect of proportion congruency is independent of contingency and correlation. If proportion congruency effects are due to contingency (Schmidt, 2018), which has been controlled in the alternative design then there should be no effect of proportion congruency. This means that smaller facilitation effects would be due to the lack of contingency and correlation.

A non-significant difference between the two designs would suggest that facilitation is not influenced by correlation and contingency. However, a less parsimonious possibility is that proportion congruency and the combined effects of correlation and contingency are of equal strengths and have opposing effects.

Exploratory analyses (Stroop interference)

Although not the main research question, there was the opportunity to explore whether contingency and correlation has an effect on Stroop interference. These analyses were exploratory, and the outcomes did not affect the conclusions drawn from the main analysis. To answer this question, the same analyses were conducted as before, but with the calculated interference effects instead of facilitation. For the frequentist analyses, the alpha level was halved (.025) to account for increased Type 1 error in multiple comparisons. To calculate Bayes factors of the pairwise comparisons a half-normal distribution with a maximum probability at 0ms and standard deviation of 18ms was used as the priors. The value for the standard deviation was taken from the raw effect size of similar trial types reported in Hasshim and Parris (2015).

DOI: 10.1177/17470218211032548

Additionally, the effect of Stroop facilitation and interference on the response time (RT) distribution was also explored. Roelofs (2010) suggested that the inadvertent reading and converging information hypotheses predicted that facilitation affected the response time distribution differently. According to the converging information hypothesis, facilitation occurs on most trials and manifests as a general speeding up of RTs. Thus, when comparing the RT distributions of Congruent and Neutral trials, the shapes will be similar, but the entire distribution of Neutral trials will be shifted closer to that of that of the faster Congruent trials . Conversely, the inadvertent reading hypothesis states that facilitation results from a small number of Congruent trials that have very short RTs. This would then result in the RT distribution of Congruent and Neutral trials to be different on the faster RTs, and converge when the RTs are slower.

Roelofs (2010) applied Vincentized averaging on his data and observed that the effects of facilitation can be seen throughout the distribution, in line with the converging information hypothesis. The data from the current study were rank-ordered and grouped into 20% quantiles, and the mean RT of each quantile plotted to allow for a visual depiction of the effects of facilitation and interference throughout the RT distribution.

To formally test these observations, the first and last quantiles of the facilitation effect were be compared within each experiment. The inadvertent reading hypothesis would predict a larger effect in the first quantile, while the converging information hypothesis predicts that the effects at the two quantiles will be comparable.

DOI: 10.1177/17470218211032548

Results

The mean RTs and error rates for each of the conditions are shown in Table 2[1]. The omnibus 2 (Stroop effects: facilitation & interference) x 2 (design: standard and alternative) ANOVA revealed that the interaction was non-significant [$F(1,49) = 0.006$, $p = .940$, $\eta_p^2 < .001$]. The planned comparisons showed that the difference between the facilitation effects from the standard (M = 30ms, SD = 35ms) and alternative (M = 18 ms, SD = 49ms) designs was non-significant [t (49) = 1.37, p = .178, d = .193]. The Bayes factor calculated with the previously specified prior distribution was B = 1.45, indicating that the evidence was insensitive. Sensitivity analysis postulating a maximum effect of up to seven times the plausible effect showed insensitive BFs of less than 3 throughout, meaning that the same inconclusive result would have been obtained, which indicates that the interpretation was not sensitive to the prior distribution chosen. Assuming the same standard error, a BF supporting the hypothesis of the alternate design reducing facilitation effects would only be obtained if the reduction was greater than 16ms.

The pre-specified analysis involving Stroop interference (incongruent – neutral) showed that the interference effects between the standard (M = 65ms, SD = 62ms) and alternative (M = 54 ms, SD = 52ms) designs was non-significant [t (49) = 1.11, p = .275, d = .156]. The Bayes factor obtained using the specified priors was insensitive (BF = 1.29). Sensitivity analysis showed that an insensitive BF would have been obtained with a prior distribution scaled up to 95ms (a range of raw effect

[1] The supplementary materials report the RT analyses done separately for participants who went through the two orders of presentation (standard or alternative designs first) for both experiments. The results were largely similar to the original analyses, with the alternative designs showing smaller, but non-significant, facilitation effects in both experiments. The pattern for interference were also similar with the effects in Experiment 1 being numerically smaller, but non-significant. In Experiment 2, the pattern for interference effect was also similar, although only significant in the participants who did the alternative design first, while the effect did not reach significance (p = .070) in the participants who did the standard design first.

DOI: 10.1177/17470218211032548

sizes larger than the observed interference effect in the standard design), which indicates that the interpretation is not sensitive to the chosen prior. Assuming the same standard error, a BF supporting the hypothesis of the alternate design reducing interference effects would only be obtained if the reduction was greater than 17ms.

Error rates

The omnibus 2 (Stroop effects: facilitation & interference) x 2 (design: standard and alternative) ANOVA for error rates revealed that the interaction was non-significant [$F(1,49) = 0.095$, $p = .759$, $\eta_p^2 = .002$]. The main effect of design was statistically significant [$F(1,49) = 10.75$, $p = .002$, $\eta_p^2 = .180$], and the main effect of Stroop effects were not [$F(1,49) = 0.783$, $p = .381$, $\eta_p^2 = .016$]. Post-hoc analysis revealed that error rates were higher in the standard compared to the alternative design [t (49) = 3.28, $p_{holm} = .002$, $d = .464$].

[Table 2 somewhere here]

Experiment 2

Method

Participants

61 participants were recruited from the same population as Experiment 1 and based on the same power analysis. Data from 12 participants were excluded for not meeting the 90% accuracy threshold.

Apparatus and Design

The apparatus used was the same as that of Experiment 1, while the design was also similar, apart from the makeup of trials in the alternative design which followed the example on the right panel of Figure 2. Blocks of trials in this alternative design

DOI: 10.1177/17470218211032548

thus had equal number of Congruent and Incongruent trials, while also controlling for contingency (but introduce correlation).

As with Experiment 1, facilitation was calculated by the difference between Neutral and Congruent trials, while interference was calculated by the difference between Incongruent and Neutral trials.

Stimuli, Procedure, and Data exclusion criteria

The details of the stimuli, procedure and the data exclusion criteria were exactly the same as those of Experiment 1.

Analysis plan

Similar to Experiment 1, an omnibus 2 (Stroop effects: facilitation & interference) x 2 (design: standard & alternative) ANOVA was conducted. A planned t-test between the size of facilitation in the standard and alternative design was also conducted, and its corresponding BF was calculated using the same prior as Experiment 1.

This comparison will further elucidate the effects of task design on the measurement of facilitation. If facilitation were found to be smaller in the alternative design, not only does it suggest that the facilitation effect is influenced by the design of an experiment, but that facilitation observed in studies employing the standard design were due to learnt word-response contingency. If the facilitation effects between the two designs were not statistically different, it would suggest that facilitation is not influenced by task design. However the role of correlation might be a factor as the correlation in the alternative design (contingency coefficient C = 0.58) was even higher than that of the standard design (C=0.31).

DOI: 10.1177/17470218211032548

Results

The omnibus 2 (Stroop effects: facilitation & interference) x 2 (design: standard and alternative) ANOVA revealed that the interaction was statistically significant [$F(1,48) = 6.67$, $p = .013$, $\eta_p^2 = .122$]. The planned comparisons showed that the difference between the facilitation effects from the standard (M = 10ms, SD = 44ms) and alternative (M = 19ms, SD = 38ms) designs was non-significant [t (48) = -1.07, p = .290, d = - .153, BF = 0.18]. This indicates that there was no significant difference between the size of facilitation measured in the two designs with Bayes factors indicating evidence for the null hypothesis. Sensitivity analysis indicated that this interpretation holds for a prior distribution scaled to at least 12 ms (i.e., the BF is >1/3 with prior estimates smaller than 12 ms). As the result of interest is evidence for the alternate hypothesis, this does not affect the interpretation that the alternate design does not reduce facilitation effect[2].

Interference effects between the standard (M = 72ms, SD = 63ms) and alternative (M = 38 ms, SD = 54ms) designs was significant with the Bayes factor indicated evidence for a difference [t (48) = 3.20, p = .002, d = .458, BF = 45.68]. Sensitivity analysis postulating a maximum effect of seven times the plausible effect showed BFs of more than 3 throughout, meaning that the interpretation was not sensitive to the prior distribution chosen.

Error rates

The omnibus 2 (Stroop effects: facilitation & interference) x 2 (design: standard and alternative) ANOVA for error rates revealed that interaction was non-significant [$F(1,48) = 1.218$, $p = .275$, $\eta_p^2 = .025$]. The main effects of design [$F(1,48)$

[2] The BF analysis used a half-normal distribution to reflect the directional hypothesis, akin to a one-tailed test. However, since raw effects indicated a larger value in the alternative design, we repeated the original analysis using a normal distribution, which resulted in a BF of 0.569. This does not change the interpretation that the alternative design did not reduce facilitation.

DOI: 10.1177/17470218211032548

= 0.055, *p* = .815, η_p^2 = .001], and Stroop effects were non-significant as well [*F*(1,48) = 1.063, *p* = .308, η_p^2 = .022].

Response time distribution

The pattern of mean facilitation effects of the rank-ordered RTs (see Figure 3) did not support the inadvertent reading hypothesis as facilitation effects did not decrease through the quantiles. The comparison between the facilitation effects at the first and last quantiles were statistically non-significant for both the designs of Experiment 2 [standard: *t*(48) = -0.476, *p* = .636, *d* = -.068; alternate: *t*(48) = 1.95, *p* = .057, *d* = .279] and in the standard design of Experiment 1 [*t*(49) = 1.25, *p* = .216, *d* = .177]. In the alternative design of Experiment 1, this difference was statistically significant *t*(49) = 2.53, *p* = .015, *d* = .358].

[Figure 3 somewhere here]

General Discussion

The aim of the current study was to explore whether the magnitude of Stroop effects is influenced by the imbalance of stimuli pairings inherent to some common task designs. The primary process of interest was facilitation as it has not been previously explored directly in this context. Experiment 1 compared the standard Stroop task design with one that controlled for colour-word correlation and word-response contingency, which necessitated twice the number of incongruent trials compared to congruent trials. Experiment 2 had equal numbers of each trial-type and controlled for word-response contingency, but not colour-word correlation. Since a direct comparison of the two ways of controlling for contingency effects has not been previously made for interference effects, these were also analysed in secondary analyses.

DOI: 10.1177/17470218211032548

Although visual inspection of the mean RTs of the different conditions (see Figure 4) showed the expected pattern of results, with smaller Stroop effects seen in the two alternative designs compared to the standard design, the inferential statistics comparing the effects of facilitation and interference between the two designs in each experiment showed that only the difference in interference effects for Experiment 2 was statistically significant. For facilitation, the BFs obtained were insensitive in Experiment 1, and even indicated evidence for no difference in Experiment 2. For Stroop interference, the difference between the measured effects in the two designs was not statistically significant in Experiment 1, with the BF obtained being insensitive, while in Experiment 2, interference was significantly smaller in the alternative design, with the BF also indicating evidence for a difference.

[Figure 4 somewhere here]

For the primary aim of studying the effects of facilitation, the RT data in Experiment 2 provide evidence supporting the null hypothesis that larger facilitation effects are not observed in the standard design, suggesting that word-response contingency effects do not affect the measurement of facilitation in manual response Stroop tasks. Our results were however insensitive with regards to the influence of colour-word correlation on performance. We further explored the distribution of RTs to investigate whether inadvertent reading could be the mechanism by which the Stroop facilitation effect occurs. If this is the case, larger facilitation effects would be expected to be observed in trials with faster RTs. This theoretical prediction was not observed. Our results do not therefore support the inadvertent reading hypothesis of Stroop facilitation effects. As depicted in Figure 3 the pattern of results was also not fully consistent with the predictions of the converging information hypothesis either.

DOI: 10.1177/17470218211032548

In the standard design in both experiments, there is a visible decrease in facilitation from quantile four to quantile five, and it is unclear why this might be the case. For a more detailed inspection of the effects on RT distributions, a more formal technique such as ex-Gaussian analysis, which requires much more data points (e.g., see Hasshim, Downes, Bate, & Parris, 2019), would be necessary.

As noted, the planned secondary analyses did reveal a smaller interference effect in the alternative design of Experiment 2 where contingency was controlled compared to the standard design. This effect was not statistically significant in Experiment 1 in which both contingency and correlation were controlled. At first blush, the significant results from the RT analyses of interference effects in Experiment 2 suggests that when word-response contingency is controlled for, Stroop interference is reduced. However, the potential influences of correlation and proportion congruency should be carefully considered since they are intertwined, and their independent influence cannot be easily ascertained. As stated in the introduction, although the alternative design of Experiment 2 controlled for word-response contingency and had equal number of congruent and incongruent trials, it also had an even higher colour-word correlation coefficient compared to the standard design. According to Dishon-Berkovits and Algom (2000) correlation disrupts the selective attention process as it makes the irrelevant word dimension a more reliable source of information, encouraging attention to be focused on it. Thus, increasing correlation would be expected to result in increased interference. However, the opposite was observed in Experiment 2, which suggests that the predicted effect of correlation was not observed. Alternatively, it is possible that the interference-increasing effect of correlation might have been hidden by the interference-reducing effects of contingency; an account that explains the lack of an effect on interference

DOI: 10.1177/17470218211032548

in Experiment 1. It is also possible that the effect of correlation is smaller in conventional colour-word Stroop tasks compared to the Stroop-like tasks used in Dishon-Berkovits and Algom (2000).

Analysis of the error rates showed an effect of task design in Experiment 1 but not Experiment 2, the reverse of what was observed in the RT data. This might suggest that the alternative designs of both experiments did have the predicted effects, but there was a trade-off between responding quickly and accurately, and the effect of task design is only observable in the RT or error data. The finding of an effect of task design in Experiment 1, when contingency and correlation was controlled, but not in Experiment 2 when only contingency was controlled indicates that it is correlation that drives the effect in the error rate data.

In conclusion, the results of the current study show that word-response contingency does not significantly affect the measurement of facilitation in the Stroop task. They do however show that contingency affects Stroop interference and indicate a possible effect of correlation on Stroop task accuracy. The findings provide further support of the idea that bottom-up associative learning processes influence the measurement of Stroop effects (e.g., Schmidt 2019; Algom & Chajut, 2019) and highlights the importance of considering correlation and contingency in task designs in studies that aim to study the processes involved in performing the Stroop task.

DOI: 10.1177/17470218211032548

Supplementary Material

The Supplementary Material is available at: qjep.sagepub.com

Data Accessibility Statement

The data and materials from the present experiment are publicly available at the Open Science Framework website: https://osf.io/tfeq6/

DOI: 10.1177/17470218211032548

References

Algom, D., & Chajut, E. (2019). Reclaiming the Stroop effect back from Control to Input-Driven Attention and Perception. *Frontiers in psychology*, *10*.

Augustinova, M., Parris, B. A., & Ferrand, L. (2019). The loci of Stroop interference and facilitation effects with manual and vocal responses. *Frontiers in Psychology*, *10*.

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

Braem, S., Bugg, J. M., Schmidt, J. R., Crump, M. J., Weissman, D. H., Notebaert, W., & Egner, T. (2019). Measuring adaptive control in conflict tasks. *Trends in cognitive sciences*. *23*(9), 769-783

Brown, T. L. (2011). The relationship between Stroop interference and facilitation effects: Statistical artifacts, baselines, and a reassessment. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(1), 85.

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.

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(3), 844.

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

DOI: 10.1177/17470218211032548

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

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychological Review, 97*(3), 332.

Dalrymple-Alford, E. C. (1968). Interlingual interference in a color-naming task. *Psychonomic Science, 10*(6), 215-216.

Dalrymple-Alford, E. C. (1972). Associative facilitation and interference in the Stroop color-word task. *Perception & Psychophysics, 11*(4), 274-276.

De Houwer, J. (2003). On the role of stimulus-response and stimulus-stimulus compatibility in the Stroop effect. *Memory & Cognition, 31*(3), 353-359.

Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on?. *Perspectives on Psychological Science, 6*(3), 274-290.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5*, 781.

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.

Dunbar, K., & MacLeod, C. M. (1984). A horse race of a different color: Stroop interference patterns with transformed words. *Journal of Experimental Psychology: Human Perception and Performance, 10*(5), 622.

Glaser, M. O., & Glaser, W. R. (1982). Time course analysis of the Stroop phenomenon. *Journal of Experimental Psychology: Human Perception and Performance, 8*(6), 875-894.

DOI: 10.1177/17470218211032548

Goldfarb, L., & Henik, A. (2007). Evidence for task conflict in the Stroop effect. *Journal of Experimental Psychology: Human Perception and Performance, 33*(5), 1170-1176.

Hasshim, N., Downes, M., Bate, S., & Parris, B. A. (2019). Response time distribution analysis of semantic and response interference in a manual response Stroop task. *Experimental Psychology, 66*(3), 231-238.

Hasshim, N., & Parris, B. A. (2014). Two-to-one color-response mapping and the presence of semantic conflict in the Stroop task. *Frontiers in psychology, 5*, 1157.

Hasshim, N., & Parris, B. A. (2015). Assessing stimulus–stimulus (semantic) conflict in the Stroop task using saccadic two-to-one color response mapping and preresponse pupillary measures. *Attention, Perception, & Psychophysics, 77*(8), 2601-2610.

Hasshim, N., & Parris, B. A. (2018). Trial type mixing substantially reduces the response set effect in the Stroop task. *Acta Psychologica*, *189*, 43-53.

The jamovi project (2019). *jamovi* (Version 0.9) [Computer Software]. Retrieved from https://www.jamovi.org

Kalanthroff, E., & Henik, A. (2013). Individual but not fragile: Individual differences in task control predict Stroop facilitation. *Consciousness and Cognition, 22*(2), 413-419.

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.

DOI: 10.1177/17470218211032548

Klein, G. S. (1964). Semantic power measured through the interference of words with color-naming. *The American Journal of Psychology, 77*(4), 576-588.

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.

Lorentz, E., McKibben, T., Ekstrand, C., Gould, L., Anton, K., & Borowsky, R. (2016). Disentangling genuine semantic Stroop effects in reading from contingency effects: On the need for two neutral baselines. *Frontiers in Psychology*, *7*, 386.

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.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*(2), 163-203.

MacLeod, C. M. (1992). The Stroop task: The" gold standard" of attentional measures. *Journal of Experimental Psychology: General, 121*(1), 12-14.

Macleod, C. M. (1998). Training on integrated versus separated Stroop tasks: The progression of interference and facilitation. *Memory & Cognition, 26*(2), 201-211.

MacLeod, C. M., & MacDonald, P. A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences, 4*(10), 383-391.

Melara, R. D., & Algom, D. (2003). Driven by information: a tectonic theory of Stroop effects. *Psychological Review, 110*(3), 422-471.

DOI: 10.1177/17470218211032548

Parris, B.A., Hasshim, N., Wadsley, M., Augustinova, M., & Ferrand, L. (submitted). A critical review of methods and evidence for levels of processing contributing to colour-word Stroop effects and the implications for the loci of attentional selection

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195-203.

Phaf, R. H., Van der Heijden, A. H. C., & Hudson, P. T. (1990). SLAM: A connectionist model for attention in visual selection tasks. *Cognitive Psychology, 22*(3), 273-341.

Roelofs, A. (2003). Goal-referenced selection of verbal action: modeling attentional control in the Stroop task. *Psychological Review, 110*(1), 88-125.

Roelofs, A. (2010). Attention and facilitation: Converging information versus inadvertent reading in Stroop task performance. Journal of Experimental Psychology: Learning, *Memory, and Cognition, 36*(2), 411.

Rouder, J. N., & Haaf, J. M. (2018). Power, dominance, and constraint: A note on the appeal of different design traditions. *Advances in Methods and Practices in Psychological Science, 1*(1), 19-26.

Schmidt, J. R. (2013). The Parallel Episodic Processing (PEP) model: Dissociating contingency and conflict adaptation in the item-specific proportion congruent paradigm. *Acta Psychologica, 142*(1), 119-126.

Schmidt, J. R. (2016). Proportion congruency and practice: A contingency learning account of asymmetric list shifting effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(9), 1496.

DOI: 10.1177/17470218211032548

Schmidt, J. R. (2019). Evidence against conflict monitoring and adaptation: An updated review. *Psychonomic Bulletin & Review*, *26*(3), 753-771.

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.

Schmidt, J. R., & Cheesman, J. (2005). Dissociating stimulus-stimulus and response-response effects in the Stroop task. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 59*(2), 132.

Schmidt, J. R., Crump, M. J., Cheesman, J., & Besner, D. (2007). Contingency learning without awareness: Evidence for implicit control. *Consciousness and Cognition, 16*(2), 421-435.

Schmidt, J. R., & De Houwer, J. (2012). Contingency learning with evaluative stimuli. *Experimental Psychology, 59* (4), 175.

Schmidt, J. R., & Lemercier, C. (2019). Context-specific proportion congruent effects: Compound-cue contingency learning in disguise. *Quarterly Journal of Experimental Psychology, 72*(5), 1119-1130.

Sharma, D., & McKenna, F. P. (1998). Differential components of the manual and vocal Stroop tasks. *Memory & Cognition, 26*(5), 1033-1040.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology,18*(6), 643-662

West, R., & Baylis, G. C. (1998). Effects of increased response dominance and contextual disintegration on the Stroop interference effect in older adults. *Psychology and Aging, 13*(2), 206.

DOI: 10.1177/17470218211032548

Figure Captions

Figure 1. Makeup of trials in Experiment 1

Note. Italicised numbers indicate Congruent trials. The standard design has equal number of Congruent and Incongruent trials which result in unequal word-response contingency and word-colour correlation. The alternative design has twice as many Incongruent trials as Congruent trials which result in equal contingency and correlation of zero.

Figure 2. Makeup of trials in Experiment 2

Note: The standard design is the same as that of Experiment 1. In the alternative design each colour word appears in only one incongruent colour and its congruent colour resulting in equal numbers of Congruent and Incongruent trials and each colour word having an equal (50%) chance of appearing in one congruent or one incongruent colour. However, since each word appears in only one incongruent colour and not the other, there is a correlation bias in that e.g., only the words white and blue will predict a white response, whereas the word red will not.

Figure 3. Mean facilitation effects at each quantile in both experiments. Vertical axes in milliseconds.

Figure 4. Mean RTs of each trial-type condition in the two experiments. Vertical axes in milliseconds.

DOI: 10.1177/17470218211032548

Standard design

WORD \ COLOUR	yellow	green	red
yellow	*30*	15	15
green	15	*30*	15
red	15	15	*30*
motion	10	10	10
close	10	10	10
due	10	10	10

Alternative design

WORD \ COLOUR	blue	pink	white
blue	*20*	20	20
pink	20	*20*	20
white	20	20	*20*
soon	10	10	10
quiet	10	10	10
times	10	10	10

Figure 1. *Makeup of trials in Experiment 1*

Note. Italicised numbers indicate Congruent trials. The standard design has equal number of Congruent and Incongruent trials which result in unequal word-response contingency and word-colour correlation. The alternative design has twice as many Incongruent trials as Congruent trials which result in equal contingency and correlation of zero.

DOI: 10.1177/17470218211032548

WORD	COLOUR: yellow	green	red
yellow	*30*	15	15
green	15	*30*	15
red	15	15	*30*
motion	10	10	10
close	10	10	10
due	10	10	10

Standard design

WORD	COLOUR: blue	pink	white
blue	*30*	30	0
pink	0	*30*	30
white	30	0	*30*
soon	10	10	10
quiet	10	10	10
times	10	10	10

Alternative design

Figure 2. Makeup of trials in Experiment 2

Note: The standard design is the same as that of Experiment 1. In the alternative design each colour word appears in only one incongruent colour and its congruent colour resulting in equal numbers of Congruent and Incongruent trials and each colour word having an equal (50%) chance of appearing in one congruent or one incongruent colour. However, since each word appears in only one incongruent colour and not the other, there is a correlation bias in that e.g., only the words white and blue will predict a white response, whereas the word red will not.

DOI: 10.1177/17470218211032548

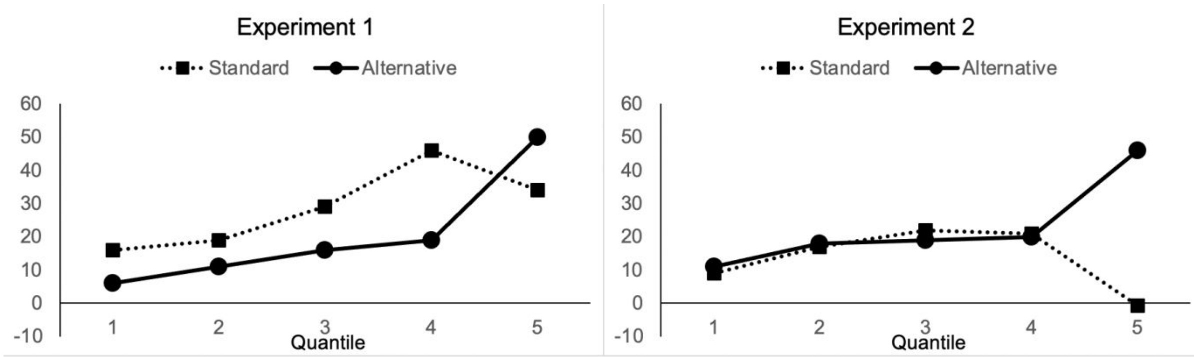

Figure 3. *Mean facilitation effects at each quantile in both experiments. Vertical axes in milliseconds.*

DOI: 10.1177/17470218211032548

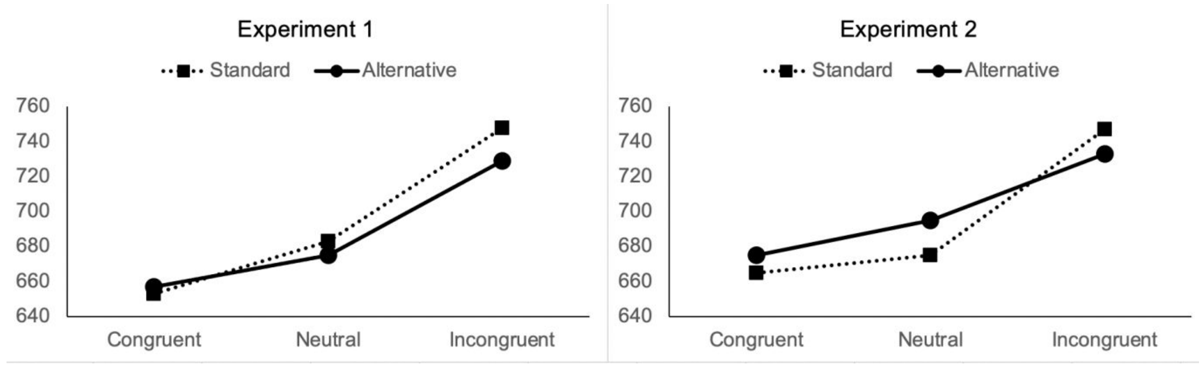

Figure 4. *Mean RTs of each trial-type condition in the two experiments. Vertical axes in milliseconds.*

DOI: 10.1177/17470218211032548

Table 1

Length and frequency details of the word stimuli used, taken from the English Lexicon Project (Balota et al. 2007)

		Word	Length	HAL frequency	log Hal frequency
Set 1	Colour	yellow	6	19319	9.87
		green	5	90773	11.42
		red	3	103819	11.55
	Neutral	motion	6	19183	9.86
		close	5	84927	11.35
		due	3	100775	11.52
Set 2	Colour	blue	4	89005	11.40
		pink	4	13066	9.48
		white	5	149742	11.92
	Neutral	soon	4	90301	11.41
		quiet	5	13086	9.48
		times	5	156832	11.96

DOI: 10.1177/17470218211032548

Table 2

Descriptive statistics of response times in milliseconds (and error rates in %) of all conditions in both experiments

	Standard design			Alternative design		
	Congruent	Neutral	Incongruent	Congruent	Neutral	Incongruent
Experiment 1 (N = 50)						
Mean	653 (2.73)	683 (4.24)	748 (6.27)	657 (4.00)	675 (4.44)	729 (5.12)
SD	128 (2.20)	131 (3.30)	161 (4.12)	134 (3.60)	135 (2.81)	152 (3.16)
Minimum	444 (0.0)	451 (0.0)	453 (0.0)	419 (0.0)	428 (0.0)	455 (0.0)
Maximum	1229 (7.78)	1238 (13.33)	1335 (16.67)	973 (15.0)	1051 (12.22)	1128 (12.50)
Experiment 2 (N = 49)						
Mean	665 (3.90)	675 (4.31)	747 (6.01)	675 (3.38)	695 (4.54)	733 (5.67)
SD	117 (3.64)	118 (3.12)	141 (4.20)	141 (2.84)	135 (3.35)	143 (4.17)
Minimum	498 (0.0)	518 (0.0)	542 (0.0)	482 (0.0)	484 (0.0)	502 (0.0)
Maximum	1065 (12.22)	1125 (11.11)	1186 (16.67)	1060 (11.11)	1083 (12.22)	1079 (17.78)

DOI: 10.1177/17470218211032548