Response time distribution analysis of semantic and response interference in a manual response Stroop task

Nabil Hasshim[1,2], Michelle Downes[2], Sarah Bate[1], and Benjamin A. Parris[1]

[1]Department of Psychology, Bournemouth University, United Kingdom

[2]School of Psychology, University College Dublin, Ireland

Correspondence:
Nabil Hasshim
School of Psychology
Newman Building
University College Dublin
Belfield
Dublin 4
Ireland
Email: nabil.hasshim@gmail.com

Raw data, analysis files, and supplementary material are publicly available at

https://osf.io/4d6ta/

Abstract

Previous analyses of response time distributions have shown that the Stroop effect is observed in the mode (μ) and standard deviation (σ) of the normal part of the distribution, as well as its tail (τ). Specifically, interference related to semantic and response processes have been suggested to specifically affect the mode and tail of respectively. However, only one study in the literature has directly manipulated semantic interference, and none manipulating response interference. The present research aims to address this gap by manipulating both semantic and response interference in a manual response Stroop task, and examining how these components of Stroop interference affect the response time distribution. Ex-Gaussian analysis showed both semantic and response conflict to only affect τ. Analysing the distribution by rank-ordered response times (Vincentizing) showed converging results as the magnitude of both semantic and response conflict increased with larger response times. Additionally, response conflict appeared earlier on the distribution compared to semantic conflict. These findings further highlight the difficulty in attributing specific psychological processes different parameters (i.e. μ, σ, and τ). The effect of different response modalities on the makeup of Stroop interference is also discussed.

Keywords: Stroop interference, response conflict, response time distribution, manual responding

Response time distribution analysis of semantic and response interference in a manual response Stroop task

A fundamental question for selective attention research is at which stage, or stages, of information processing controlled attention is utilised to select relevant information and ignore irrelevant information (Duncan, 1980; Lupker & Katz, 1981). Probably the most popular selective attention task in psychological research is the Stroop task (Klein, 1964; Stroop 1935), where the objective of responding to the colour that the word is printed in is made difficult when the word itself spells out the name of a different colour. Popular models of the Stroop task (e.g., Cohen, Dunbar, & McClelland, 1990; Roelofs, 2003) have attributed interference to the response output stage. However, there has been a recent upsurge in studies investigating whether Stroop interference stems from earlier in the processing stream, such as at the level of semantics, and how it can be reliably measured and dissociated from response conflict (e.g., Augustinova & Ferrand, 2012, 2014; Augustinova, Flaudias, & Ferrand, 2010; Augustinova, Silvert, Spatola, & Ferrand, 2017; De Houwer, 2003; Goldfarb & Henik, 2007; Hasshim & Parris, 2014, 2015; Parris, 2014; Risko, Schmidt, & Besner, 2006; Schmidt & Cheesman, 2005; Schmidt, Hartsuiker, & De Houwer, 2018; Shichel & Tzelgov, 2017). These studies typically compare the performance (usually using response time) of classic incongruent trials which are thought to involve both semantic and response conflict, to another type of incongruent trial which involves semantic but not response conflict (see Sharma & McKenna, 1998, for examples of such stimuli).

In addition to introducing such incongruent trial variants, some researchers have analysed response time (RT) distributions to complement the standard practice of analysing mean RTs. However, it has been argued that attributing specific cognitive processes uniquely to parameters of theoretical distributions should be done with great caution (Matzke & Wagenmakers, 2009). Even though the ex-Gaussian model fits RT distributions neatly, it is data driven and atheoretical (Luce, 1986), and the parameters do not correspond well with a Wald distribution or diffusion modelling (Matzke & Wagenmakers, 2009). Nevertheless, Heathcote et al. (1991) noted that the ex-Gaussian function can and should be used in a purely descriptive manner and as recent attempts support, it would be a useful addition to the literature to show how variants of the Stroop paradigm affect different components of the ex-Gaussian distribution (Steinhauser & Hubner, 2009; White, Risko, & Besner, 2016).

Distributional analysis of RT data

Since RT distributions typically have a positively skewed unimodal shape, information about how experimental manipulations affect the shape of the sample's distribution can be missed out by typical analysis of the mean (Heathcote, Popiel, & Mewhort, 1991; Yap, Balota, Tse, & Besner, 2008). Balota and Yap (2011) emphasised how distributional analysis methods, such as ex-Gaussian analysis, can be easily applied to RT data of cognitive experimental paradigms. This approach involves fitting an ex-Gaussian function to the empirical RT distribution and estimates three parameters corresponding to different components of the distribution. The parameters μ and σ correspond to the mean and standard deviation of the normal (Gaussian) portion of the distribution respectively, and τ reflects the mean and standard deviation of the exponential component (Balota & Yap, 2011). In addition to differences in mean RT provided by means analysis, information from the ex-Gaussian parameters reflect how the change in overall mean RT came about: by shifting the RT distribution (μ), changing its scale (σ), or affecting its skew (τ) (Heathcote et al. 1991; Tse, Balota, Yap, Duchek, & McCabe, 2010). This potentially allows for a better understanding of the nature of the effect under investigation. For example, Spieler et al. (1996) showed the slowing down of RTs for participants with mild Dementia of the Alzheimer's Type was isolated in τ, suggesting that it was not due to general slowing of cognitive processes, which would be reflected in μ, but instead due to slow responses on some trials, resulting in an increase in the skew of the distribution, which they attribute to decreased goal-focus.

Vincentizing (Vincent, 1912) is another popular way of looking at effects on an RT distribution. It is a non-parametric technique that involves rank ordering each participant's RT data within each condition and plotting the mean of the deciles (10% quantiles) across participants. Unlike the ex-Gaussian analysis, Vincentile plots are based on raw RT distributions and do not assume an explicit function for the shape of the distribution (Ratcliff, 1979) which is one criticism for estimating parameters as is done in ex-Gaussian analysis (see Rouder & Speckman, 2004; and Matzke & Wagenmakers, 2009). The Vincentile plots show how different parts of the RT distribution are affected by each manipulation and, unlike ex-Gaussian analysis, will reveal whether the magnitude of the effect differs between relatively faster and slower responses[1].

Ex-Gaussian parameters and components of Stroop interference

In studies that have applied ex-Gaussian analysis to Stroop task performance, Stroop interference (the difference in performance between incongruent and baseline trials) has been shown to affect all three parameters (Heathcote et al., 1991; Spieler, Balota, & Faust, 1996). Additionally, White et al. (2016) utilised a semantic associates version of the Stroop task (a popular manipulation used in the literature to isolate semantic conflict) and observed significant effects only in μ

---

[1] This does not imply effects affecting different stages (early vs late) of processing.

(similar to semantic priming effects reported in Balota et al, 1998). This result suggests that the semantic conflict component of Stroop interference only affects μ, and implies that Stroop interference effects observed in the other two parameters (τ and σ) stemmed from the non-semantic component. This is consistent with Spieler et al.'s (1996) suggestion that response conflict is captured in τ. However studies directly measuring components of Stroop interference are not common with White et al. (2016) being the only one thus far manipulating semantic conflict.

The present study

The aim of the present study was to directly observe the effects of semantic and response conflict on the RT distribution when they are dissociated in the same Stroop task. This has not been done in previous distributional analyses of Stroop interference as inferences on the effects its component parts have been using studies on Stroop interference as a whole or comparing interference with task conflict. Ex-Gaussian analysis was used to show the effects of semantic and response conflict on each parameter. Vincentizing was also used as a non-parametric compliment to the ex-Gaussian analysis to graphically show the effects of semantic and response conflict throughout the RT distribution.

The present paradigm utilised the response set effect (RSE); a popular measure of response conflict/competition (e.g., Klein, 1964; Milham et al., 2001; Risko, Schmidt, & Besner, 2006; Sharma & McKenna, 1998; also see MacLeod, 1991, for a review). The RSE is the difference in performance between standard incongruent trials where the irrelevant word is in the set of possible responses, and incongruent trials where the colour that the irrelevant word spells out does not belong to the set of possible responses (i.e., the irrelevant word spells out a colour that is never presented, and thus is never a valid response). These are referred to as *response set* (RS) and *non-response set* (NR) trials respectively. The non-RSE component of Stroop interference is measured by the difference between NR trials and Neutral-word (NE) trials (where the task-irrelevant word is not associated to a colour, e.g., ‘table’) and is taken as a measure of semantic conflict (Klein, 1964; Sharma & McKenna, 1998). If semantic and response conflict affect unique parts of the RT distribution with a manual response as they do with a vocal response, semantic conflict would be in reflected in μ, while response conflict in τ, and this will likely represent a general shift and increased skewness in the distribution respectively.

Method

Participants

30 participants were recruited from the student population of Bournemouth University in exchange for £12. Ethical approval was obtained from the Bournemouth University Research Ethics Committee.

Apparatus

Stimuli were presented on a PC using Experiment Builder software (SR Research Ltd.) with responses recorded via pressing the 1, 2, and 3 keys on a keyboard, which corresponded to one of the three possible colour responses. Participants were tested individually and positioned approximately 60 cm from the computer screen resulting in the stimuli having a vertical visual angle of 0.95° and horizontal visual angle of 1.91° to 3.82° (depending on word length).

Design

The experiment was a within participants design with three trial types: Neutral (NE), Non-Response set (NR) and Response set (RS) trials. Participants completed the experiment over two sessions conducted on different days. [2]

Stimuli

There were two versions of the experiment to ensure any observed results were not due to the particular colours used as response and non-response set colours. In version one, the response set colours that were assigned button responses were *white*, *blue*, and *orange*, while version two used *black*, *pink*, and *green*. The NR colour words were *red, gold,* and *green* for version one, and *white, blue,* and *orange* for version two. The words *laugh, soon,* and *away* were used for the NE trials in both versions. In both versions, stimuli were presented on a light grey background. The versions administered were counterbalanced among participants. Each response set colour was matched as closely as reasonably possible on word frequency and length, to a non-response colour and neutral word based on the English Lexicon Project database norms (Balota et al., 2007)

---

[2] An additional manipulation of presentation format was originally included in the design as Hasshim and Parris (2018) showed larger RSE when trials from the same trial type were blocked together. However, the analysis of the mean RT data showed that this effect did not statistically replicate, even though the pattern of results were similar to the previous findings. Thus the data from the two presentation formats have been combined throughout this manuscript and this manipulation is not discussed further. Means and ex-Gaussian analyses with presentation format as a factor are presented in the supplementary material and showed similar results to the analysis of the combined data.

Procedure

Each session started with a practice block of 24 letter strings (i.e.. ###, ####, #####, ######) trials displayed in all three colours before moving on to the experimental blocks. During each session, 27 blocks of 40 trials were presented making a total of 2160 experimental trials in both sessions, made up of 720 trials from each trial type.

On each trial, a dark grey fixation cross appeared at the centre of a light grey screen for 500ms, followed by the Stroop stimuli which stayed visible until a response was made. On incorrect responses, visual feedback, in the form of the word ‘Incorrect’, was displayed 1° above the centre of the screen for an additional 1500ms. A 100ms blank screen concluded each trial.

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

Results

Incorrect responses (2.81%) and responses not within 200 ms and 2500 ms (a further 0.45%) were excluded from analyses. The average percentage of valid responses among participants were 96.73% (*SD* = 1.80). All inferential statistical analyses were conducted in JASP Version 0.8.6 (JASP Team, 2018). Bayes factors were calculated for all pairwise comparisons to determine the ratio of evidence for the alternate and null hypotheses using the online Dienes Bayes factor calculator (https://medstats.github.io/bayesfactor.html). All Bayes factors, *B*, reported represent the evidence for H1 relative to H0; to find the evidence for H0 relative to H1, take 1/*B*. $B_{N(0,\,x)}$ denotes a Bayes factor in which the prior was modelled as a normal distribution with a mean of 0 and SD of *x*.

In calculating Bayes factors, a prior distribution (model of H1), which quantifies the likelihood of effect sizes, has to be specified. Following the recommendations of Dienes (2018), the prior distributions were set as normal distributions with the most likely value of 0, with the *SD* of the distribution scaled to the expected raw effect size based on the literature of similar effects. Typically a directional theory is reflected by using a half-normal distribution, but in the present analysis two-tailed distributions were used since the literature of studies investigating semantic and response conflict in Stroop interference have shown effects in both directions. Compared to a one-tailed distribution, this model has a lower likelihood of all positive effect sizes and does not assume an effect in the opposite direction is impossible. This also reflects how the varied findings in the literature lowers the expectation of observing an effect.

For analyses measuring mean non-RSE (non-response – neutral) and RSE (response set – non-response) effects, the prior distributions were scaled to 15ms and 32ms respectively. These values corresponded to the average raw effect sizes reported in Experiment 1 of Hasshim and Parris (2018), which is the experiment with the most similar design to the current study. For analyses involving ex-Gaussian parameters the model for the prior distribution were scaled to 14ms, the effect size of the significant semantic conflict of μ reported in White et al. (2016).

Means analysis

Table 1 shows the descriptive statistics of each trial type in experiment. A repeated measures ANOVA showed a significant effect of trial type [$F(2,58) = 35.05$, $p < .001$, $\eta_p^2 = 0.547$], while pairwise comparisons showed significant non-RSE [$t(29) = 4.56$, $p_{bonf} < .001$, $d = 0.832$, $B_{N(0,15)} = 4250.78$] and RSE [$t(29) = 5.10$, $p_{bonf} < .001$, $d =$

0.931, $B_{\mathrm{N}(0,15)} = 43604.22$]. This indicated that both semantic and response conflict were observed.

Table 1: Descriptive statistics of mean reaction time (in ms) of each trial type

	Neutral	Non-response	Response set
Mean	594	607	627
Std. Deviation	74	80	80
Minimum	457	452	481
Maximum	780	811	815

Ex-Gaussian analysis

Ex-Gaussian parameters were obtained by using the QMPE software (Heathcote, Brown, & Cousineau, 2004). Data from all 30 participants for all 3 trial types were fitted together, with the number of quantiles set to 623, which is one less than the smallest number of valid responses for any participant in any condition. All parameter estimations were within the recommended output criterion (QMPE v2.18 Technical Manual). This resulted in parameter estimates (μ, σ, and τ) for each trial type and repeated measures ANOVAs were conducted for each of the three parameters. Descriptive statistics of each parameter estimate are detailed in Table 2.

Table 2: Descriptive statistics (in ms) of each trial type in each parameter

		Neutral	Non-response	Response set
mu	Mean	415	412	412
	Std. Deviation	42	40	43
	Minimum	280	284	269
	Maximum	522	512	518
sigma	Mean	50	48	50
	Std. Deviation	11	10	12
	Minimum	30	28	20
	Maximum	73	71	76
tau	Mean	179	195	216
	Std. Deviation	56	68	69
	Minimum	75	79	97
	Maximum	303	346	351

Mu

The effect of trial type was non-significant [$F(2,58) = 1.90$, $p = .159$, $\eta_p^2 = 0.061$]. Bayes factors for non-RSE and RSE were $B_{N(0,14)} = 0.360$ and $B_{N(0,14)} = 0.151$. This suggests that μ did not index response or semantic conflict.

Sigma

The effect of trial type was non-significant [$F(2,58) = 0.694$, $p = .504$, $\eta_p^2 = 0.023$]. Bayes factors for non-RSE and RSE were $B_{N(0,14)} = 0.172$ and $B_{N(0,14)} = 0.140$. This indicates that σ did not index response or semantic conflict.

Tau

The effect of trial type was statistically significant [$F(2,58) = 33.47$, $p < .001$, $\eta_p^2 = 0.536$]. Post-hoc pairwise comparisons showed both non-RSE [$t(29) = 4.14$, $p_{bonf} < .001$, $d = 0.757$, $B_{N(0,14)} = 771.64$] and RSE [$t(29) = 4.82$, $p_{bonf} < .001$, $d = 0.879$, $B_{N(0,14)} = 11901.07$], indicating that τ indexed both semantic and response conflict.

Vincentizing

Figure 1 shows a visual depiction of the mean RT of each trial type at each quantile. The curves suggest that the magnitude of both non-RSE and RSE increase with increasing RTs.

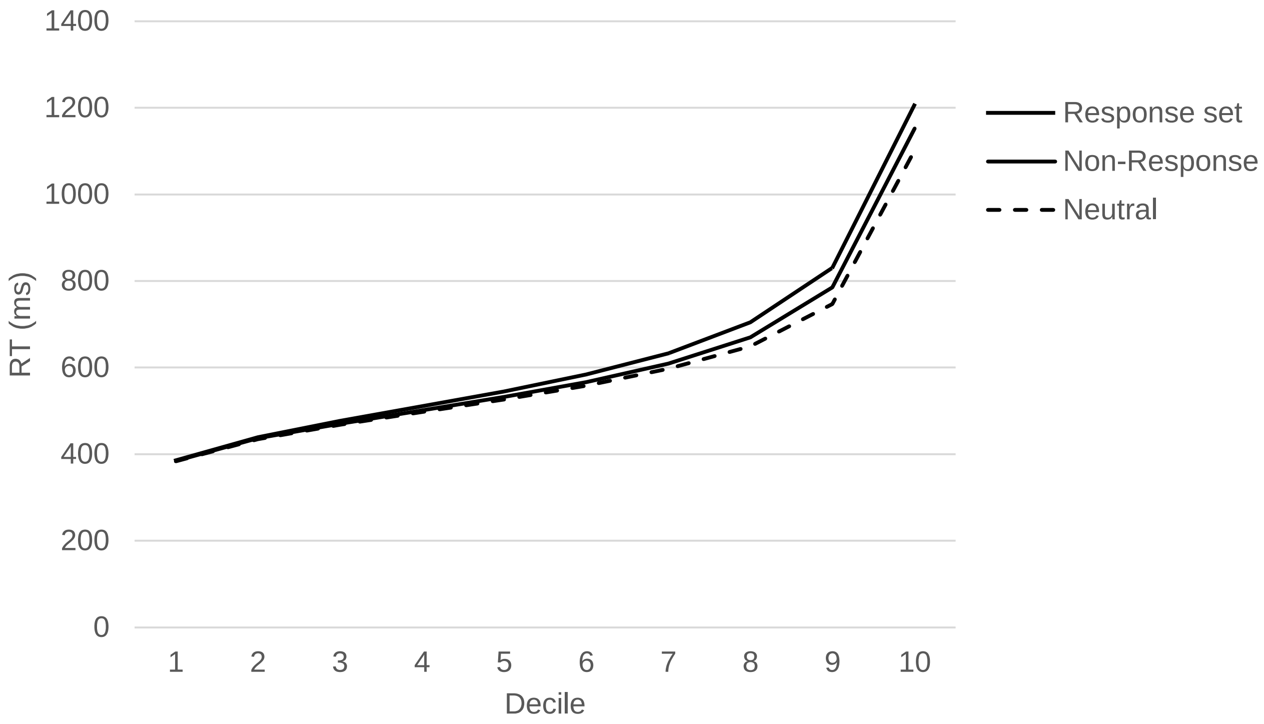

Figure 1: Mean RT of each trial type at each decile

To test this observation, a 3 (trial type) x 10 (bins) ANOVA was conducted, which showed a statistically significant interaction [$F(18,522) = 22.39$, $p < .001$, $\eta_p^2 = 0.436$]. Additionally, inferential tests of non-RSE and RSE effects at each bin were conducted and displayed in Tables 3 and 4 respectively. This is typically not done with vincentized data, but collapsing trials across presentation formats resulted in enough trials for the tests to be well powered.

Table 3: Inferential tests of non-RSE (NR – NE) at each decile

Decile	*M* (ms)	*SE*	*t*	*df*	*p*	Cohen's *d*	$B_{N(0,15)}$
1	1	1.51	0.531	29	0.599	0.097	0.276
2	2	1.33	1.44	29	0.162	0.262	0.246
3	3	1.86	1.44	29	0.161	0.263	0.341
4	3	2.05	1.59	29	0.123	0.290	0.467
5	6	2.03	2.92	29	0.007*	0.533	8.78
6	8	2.60	3.11	29	0.004*	0.567	18.41
7	12	3.41	3.48	29	0.002**	0.635	69.94
8	21	5.39	3.91	29	< .001**	0.713	291.34

9	38	8.35	4.59	29	< .001**	0.838	1506.34
10	50	13.26	3.76	29	< .001**	0.687	35.36

**statistically significant with Bonferroni-Holm correction*
***statistically significant at single-step Bonferroni corrected α of .0025*

Table 4: Inferential tests of RSE (RS – NR) at each decile

Decile	*M* (ms)	*SE*	*t*	*df*	*p*	Cohen's *d*	$B_{N(0,32)}$
1	1	1.76	0.793	29	0.434	0.145	0.075
2	3	1.69	1.61	29	0.118	0.294	0.194
3	6	2.14	2.82	29	0.009*	0.514	3.38
4	10	2.44	4.14	29	< .001**	0.755	367.58
5	12	2.63	4.74	29	< .001**	0.864	5580.35
6	17	3.72	4.69	29	< .001**	0.857	6055.55
7	24	5.38	4.39	29	< .001**	0.801	1888.35
8	35	7.32	4.76	29	< .001**	0.869	10277.70
9	45	9.42	4.75	29	< .001**	0.868	9171.84
10	56	14.31	3.89	29	< .001**	0.710	221.21

**statistically significant with Bonferroni-Holm correction*
***statistically significant at single-step Bonferroni corrected α of .0025*

The comparisons seem to support the assertions from the ex-Gaussian analysis, in that both non-RSE and RSE seem to increase with increasing RTs, with mean raw effect size getting larger at later bins.

Discussion

The results show that while the non-RSE and the RSE were observed in mean RT data as expected, ex-Gaussian analyses showed that both effects were captured only in τ, with the Bayes factors calculated favouring evidence for the null in μ and σ. The vincentile plots further suggested that the both effects were strongest in the tail of the RT distribution. This is a departure from studies that have linked semantic and response conflict to μ and τ respectively (Spieler, Balota, & Faust, 1996; White, Risko, & Besner, 2016).

The role of response modality

One way that the present research deviates from studies in the literature that use ex-Gaussian analysis is that those studies required participants to respond vocally, by naming the colour aloud (a *vocal response*). Studies that have employed *manual responses*, where participants respond by pressing a button on a keyboard or response pad, have observed results similar to the that of the present research. Namely, Aarts, Roelofs, and van Turennout (2009); Parris, Dienes, and Hodgson (2013); and Steinhauser and Hübner (2009) observed Stroop interference to be significant only in τ[3], in contrast to the studies employing a vocal response. It should be noted that these studies have mainly been concerned with measuring task conflict (conflict between the task sets of word reading and colour naming), and none considered the role of semantic conflict. It is possible that the choice of vocal or manual response influences the makeup of Stroop interference and the ex-Gaussian component and/or distributional location in which response and semantic conflict are observed.

Conclusions

Not only do the results suggest that performance in the Stroop task is qualitatively different depending on the mode of response, but the processes involved in overcoming interference might also differ. Observing non-RSE in τ but not μ may mean that unlike vocal responses, semantic conflict in a manual response Stroop

---

[3] Steinhauser and Hübner (2009) interpreted that response conflict was mainly indexed in μ and task conflict in τ, but their switching paradigm was designed to measure task conflict, with their measurement of response conflict being the effect of congruency. A similar measurement of response conflict was used by Aarts, Roelofs, and van Turennout (2009), which involved congruent trials. Although this makes these results not directly comparable, we interpret their conditions that corresponded closest to our definition of interference to only affect τ.

task affects the tail of the distribution only. Another interpretation is that observing interference only in τ shows that only response-based conflict is involved during manual responses, as suggested by Sharma and McKenna (1998). Roelofs' (2003) model of Stroop interference predicts that so-called "semantic conflict" arises as a result of semantic connections between the non-response set colours and the response set colours. That is, it is only due to the activation a response set colour receives indirectly from the irrelevant non-response set Stroop word that interference arises on non-response set trials: in other words, all conflict is response conflict. This may be true when participants respond with manual responses. Indeed, our data suggest that response conflict on response set trials appears earlier than semantic conflict on non-response set trials indicating an extra processing step before interference arises in the latter.

Although the present research does not allow us to make any conclusive statements on the processes involved in semantic and response conflict, it adds to the current concerns in attributing cognitive processes to specific ex-Gaussian parameters. Although possible interpretations of the data have been brought up, more research needs to be done before it is clear which interpretation should be favoured, if at all. The present study also highlights the need to take response modalities into account when studying components of Stroop interference.

Declaration of conflicts of interest:
The Authors declare that there is no conflict of interest.

Funding:
This work was supported by the Experimental Psychology Society Small Grants awarded to the first and last author.

References

Aarts, E., Roelofs, A., & van Turennout, M. (2009). Attentional control of task and response in lateral and medial frontal cortex: brain activity and reaction time distributions. *Neuropsychologia, 47*(10), 2089-2099.

Augustinova, M., & Ferrand, L. (2012). Suggestion does not de-automatize word reading: Evidence from the semantically based Stroop task. *Psychonomic Bulletin & Review*, *19*(3), 521-527.

Augustinova, M., & Ferrand, L. (2014). Automaticity of word reading: Evidence from the semantic Stroop paradigm. *Current Directions in Psychological Science*, *23*(5), 343-348.

Augustinova, M., Flaudias, V., & Ferrand, L. (2010). Single-letter coloring and spatial cuing do not eliminate or reduce a semantic contribution to the Stroop effect. *Psychonomic Bulletin & Review*, *17*(6), 827-833.

Augustinova, M., Silvert, L., Spatola, N., & Ferrand, L. (2018). Further investigation of distinct components of Stroop interference and of their reduction by short response-stimulus intervals. *Acta Psychologica*, 189, 54-62.

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. *Current Directions in Psychological Science*, *20*(3), 160-166.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., ... & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39*(3), 445-459.

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), 332-361.

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., & Mclatchie, N. (2018). Four reasons to prefer Bayesian analyses over significance testing. *Psychonomic Bulletin & Review*, *25*(1), 207-218

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review*, *87*(3), 272-300.

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., & 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.

Heathcote, A., Brown, S., & Cousineau, D. (2004). QMPE: Estimating Lognormal, Wald, and Weibull RT distributions with a parameter-dependent lower bound. *Behavior Research Methods, Instruments, & Computers*, *36*(2), 277-290.

Heathcote, A., Popiel, S. J., & Mewhort, D. J. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin*, *109*(2), 340-347.

JASP Team (2018). JASP (Version 0.8.6)[Computer software].

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

Lupker, S. J., & Katz, A. N. (1981). Input, decision, and response factors in picture–word interference. *Journal of Experimental Psychology: Human Learning and Memory*, *7*(4), 269-282.

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

Matzke, D., & Wagenmakers, E. J. (2009). Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review*, *16*(5), 798-817.

Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., & Kramer, A. F. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Cognitive Brain Research*, *12*(3), 467-473.

Parris, B. A. (2014). Task conflict in the Stroop task: When Stroop interference decreases as Stroop facilitation increases in a low task conflict context. *Frontiers in Psychology*, *5*, 1182.

Parris, B. A., Dienes, Z., & Hodgson, T. L. (2013). Application of the ex-Gaussian function to the effect of the word blindness suggestion on Stroop task performance suggests no word blindness. *Frontiers in Psychology*, *4*, 647.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86*(3), 446-461.

Risko, E. F., Schmidt, J. R., & Besner, D. (2006). Filling a gap in the semantic gradient: Color associates and response set effects in the Stroop task. *Psychonomic Bulletin & Review*, *13*(2), 310-315.

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

Rouder, J. N., & Speckman, P. L. (2004). An evaluation of the Vincentizing method of forming group-level response time distributions. *Psychonomic Bulletin & Review*, 11(3), 419-427.

Schmidt, J. R., & Cheesman, J. (2005). Dissociating stimulus-stimulus and response-response effects in the Stroop task. *Canadian Journal of Experimental Psychology 59*(2), 132-138.

Schmidt, J. R., Hartsuiker, R. J., & De Houwer, J. (2018). Interference in Dutch–French Bilinguals. *Experimental Psychology 65*(1), 13-22

Shichel, I., & Tzelgov, J. (2018). Modulation of conflicts in the Stroop effect. *Acta Psychologica, 189*, 93-102.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(2), 461-479.

Steinhauser, M., & Hübner, R. (2009). Distinguishing response conflict and task conflict in the Stroop task: evidence from ex-Gaussian distribution analysis. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(5), 1398-1412.

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

Tse, C. S., Balota, D. A., Yap, M. J., Duchek, J. M., & McCabe, D. P. (2010). Effects of healthy aging and early stage dementia of the Alzheimer's type on components of response time distributions in three attention tasks. *Neuropsychology*, *24*(3), 300-315.

White, D., Risko, E. F., & Besner, D. (2016). The semantic Stroop effect: An ex-Gaussian analysis. *Psychonomic Bulletin & Review*, *23*(5), 1576-1581.

Yap, M. J., Balota, D. A., Tse, C. S., & Besner, D. (2008). On the additive effects of stimulus quality and word frequency in lexical decision: evidence for opposing interactive influences revealed by RT distributional analyses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(3), 495-513.