

Onset complexity and task conflict in the Stroop task

Benjamin A Parris¹ , Nabil Hasshim² , Ludovic Ferrand³ and Maria Augustinova⁴ 

Quarterly Journal of Experimental Psychology
2025, Vol. 78(5) 978–996
© Experimental Psychology Society 2023



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/17470218231214515
qjep.sagepub.com



Abstract

The present study examined the extent to which a key marker of task conflict, negative facilitation, is modified by *onset complexity*. Negative facilitation, slower reaction times (RTs) to congruent stimuli than to non-lexical neutral stimuli in the Stroop task, is thought to reflect competition between the task sets of colour naming and word reading in the Stroop task (also known as task conflict). That is, it reflects competition between whole task sets, over and above any competition between specific responses associated with a stimulus. An alternative account of negative facilitation argues that it reflects the specific phonological processing differences between pronounceable (e.g., congruent) and non-pronounceable (e.g., xxxx) stimuli that are magnified by the specific task contexts that produce negative facilitation (a mostly non-lexical trial context). Here we used *onset complexity* to manipulate pronounceability of the irrelevant words in the Stroop task to test this alternative account. However, before applying manipulations that produce negative facilitation, we initially tested whether there was an effect of onset complexity on Stroop task performance. The results from Experiment(s) 1 (and 3) showed that complex onsets led to larger positive facilitation and congruency effects relative to simple onsets, but did not modify incongruent or neutral-word RTs. Experiment 2 directly tested whether onset complexity modifies negative facilitation and provided strong evidence for no effect of onset complexity, contrary to the alternative account predictions. The implications of the results for task conflict theory, selective attention, and phonological processing in the manual response Stroop task are discussed.

Keywords

Task set; Stroop task; onset; manual response

Received: 5 May 2023; revised: 4 September 2023; accepted: 25 October 2023

The Stroop task (Stroop, 1935) requires participants to focus on one dimension of a stimulus, the colour dimension, while ignoring another dimension, the word dimension. The task produces the Stroop *interference* effect—referring to the fact that identifying the colour that a word is printed in takes longer when the word denotes a different colour (colour-incongruent trials; e.g., the word *red* displayed in blue font) compared with a comparison baseline (e.g., the colour-neutral word *top* displayed in blue font). In addition, words that are congruent with the colour (colour-congruent trials; e.g., the word *red* in red font) commonly result in faster colour-identification times when compared with a colour-neutral baseline condition, producing the Stroop *facilitation* effect (Dalrymple-Alford, 1972; Dalrymple-Alford & Budayr, 1966 see MacLeod, 1991; Parris et al., 2022, for a review).

For most of their history, the magnitudes of Stroop interference and Stroop facilitation were thought to result

solely from the information conveyed by the irrelevant word dimension. That is, the extent to which phonological, semantic, or response information in the irrelevant dimension differed (together referred to as *informational conflict*) or was consistent with that in the relevant dimension was thought to determine the magnitude of Stroop

¹Department of Psychology, Faculty of Science and Technology, Bournemouth University, Poole, UK

²School of Health and Society, University of Salford, Salford, UK

³CNRS, LAPSCO, Université Clermont Auvergne, Clermont-Ferrand, France

⁴Normandie Université, UNIROUEN, CRFDP, Rouen, France

Corresponding author:

Benjamin A Parris, Department of Psychology, Faculty of Science and Technology, Bournemouth University, Talbot Campus, Fern Barrow, Poole BH12 5BB, UK.

Email: bparris@bournemouth.ac.uk

interference and facilitation, respectively (Hasshim & Parris, 2014, 2021; MacLeod, 1991; Parris et al., 2022). However, more recent work has attributed some of Stroop task performance to an additional competition between whole task sets, a form of conflict referred to as *task conflict* (MacLeod & MacDonald, 2000; Monsell et al., 2001). That is, competition between the endogenously activated task set for colour naming and the exogenously activated task set for word reading; a form of task set conflict akin to that observed in task switching studies (Monsell, 2003; Rogers & Monsell, 1995). In line with this idea, several lines of research have attempted to measure task conflict independently of informational conflict (see Littman et al., 2019; Parris et al., 2022, for reviews).

For task conflict to occur, at least two task sets must compete for activation. This means that the entire collection of control settings/task parameters that programme the system to perform one task (e.g., word reading—consisting of visual analysis, letter/grapheme identification, lexical identification [semantic processing], phonological processing) would compete for activation with the entire collection of control settings that programme the system to perform another task (e.g., classify a colour—consisting of visual analysis, colour identification, semantic processing, phonetic encoding). The idea is that “whole task sets compete, over and above any competition between specific responses associated with a stimulus” (Monsell et al., 2001, pp. 139–140). That is, in addition to and over and above any stimulus-response associations. In the context of the Stroop task, three findings have been interpreted as evidence for “task conflict”: (1) anterior cingulate activation on congruent trials (MacLeod & MacDonald, 2000), (2) slower colour-naming responses to neutral words (e.g., the word “house”) than to non-lexical (i.e., non-pronounceable) stimuli (e.g., “xxx”); see Augustinova et al., 2019), and (3) slower responses to congruent stimuli than to non-lexical stimuli, also known as reverse or negative facilitation (e.g., Goldfarb & Henik, 2007). In the present study, we focus on the latter.

Indeed, to measure task conflict in the Stroop task, studies have often compared performance on trials on which the irrelevant dimension is a word (e.g., a colour word) to trials on which the irrelevant dimension is non-lexical (e.g., a sequence of repeated letters, such as xxx or www; e.g., Goldfarb & Henik, 2007). As congruent stimuli contain words (e.g., the word red in red), they should involve task conflict whereas non-lexical stimuli should not, meaning that colour naming responses should be longer for stimuli that have irrelevant real words than those that have irrelevant non-lexical stimuli. However, when comparing congruent and repeated-letter string trials, congruent trials are commonly responded to more quickly, producing positive, not negative facilitation (e.g., Brown, 2011; Klein, 1964; Monsell et al., 2001), which might represent a challenge to the task conflict account.

Goldfarb and Henik (2007) reasoned that this is because in most Stroop tasks, there is a sufficiently large proportion of lexical stimuli to activate task conflict control. In other words, constant exposure to real words in the Stroop task puts the *task conflict controller* on high alert, ensuring that it is active, and that task conflict is kept low. The activation of task conflict control means that positive facilitation can be expressed in the RT data. Goldfarb and Henik (2007) further reasoned that increasing the proportion of non-lexical neutral trials (e.g., repeated-letter strings) would create the expectation for a low task conflict context, thereby reducing task conflict control. This would result in the exposure of task conflict via its unique behavioural expression, *negative facilitation*, where congruent stimuli are responded to more slowly than repeated-letter stimuli. The results from their Experiment 1 supported this prediction (see also, e.g., Entel et al., 2015; Entel & Tzelgov, 2018, 2020; Kalanthroff et al., 2015).

In a subsequent experiment, Goldfarb and Henik (2007, Experiment 2) replaced repeated-letter trials with neutral word trials (i.e., the Hebrew word for *building*) and showed that negative facilitation was now absent. Thus, they not only reported negative facilitation for the first time, but also showed its extinction. As a result, they introduced the notion of a task conflict controller, a part of a system of cognitive control that is deployed to reduce or prevent task conflict (see also Kalanthroff et al., 2018; and Littman et al., 2019, for a mini review). Also, and importantly, this line of research has provided insights into the nature of selective attention and has given rise to a new potential form of impairment that might explain symptoms of clinical disorders, such as obsessive-compulsive disorder (OCD; Kalanthroff et al., 2017).

Recently, however, the extent to which task sets and their control determine cognitive performance has been questioned. In the context of task switching, some of the processing costs that were previously attributed to controlled switching between active task sets, have been accounted for with reference to feature-integration biases (see Schmidt et al., 2020). Subsequently, Parris et al. (2023) described current challenges to the task conflict account of negative facilitation in selective attention tasks. They pointed out that some of the findings that were foundational in terms of the original conceptualisation of task conflict have not been subsequently strongly supported and argued that the mechanisms that lead to negative facilitation are not well-understood. In an attempt to highlight the notion that other accounts of the data are at least feasible, Parris et al. (2023) also presented a tentative alternative account of negative facilitation. In essence, the alternative account argued that contrary to Monsell et al.’s (2001) position that “whole task sets compete, over and above any competition between specific responses associated with a stimulus” (Monsell et al., 2001, pp. 139–140), it is in fact competition between the specific phonological/

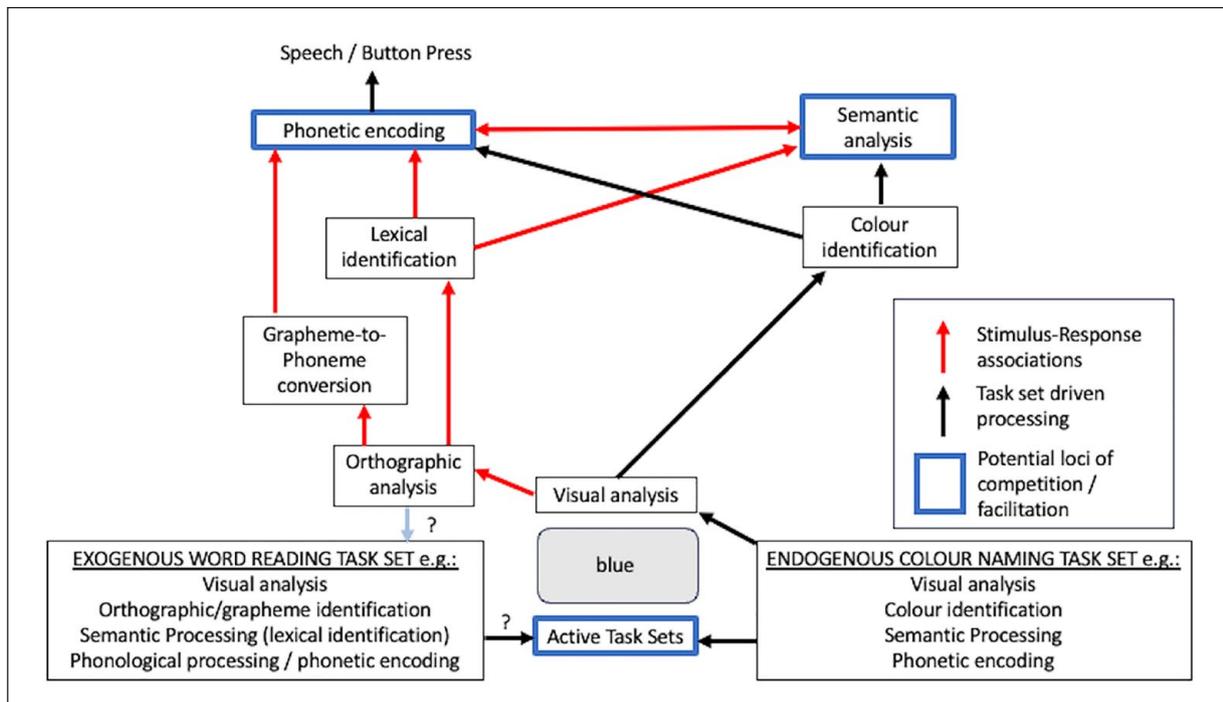


Figure 1. Task conflict theorists argue that “whole task sets compete, over and above any competition between specific responses associated with a stimulus” (Monsell et al., 2001, pp. 139–140). At some, as yet unclear, point in processing, the task set for word reading is activated, creating competition between task sets. That is, independent of, and likely earlier than some of the specific the S–R associations (in red), the task sets (which can be thought of as collections of S–R associations) compete, creating task conflict (a unique marker of which is negative facilitation). In contrast, the alternative account being tested here argues that the S–R associations in red are enough to explain negative facilitation given the differences in the pronounceability of non-lexical (i.e., xxxx) and lexical (e.g., blue) trials, especially in a mostly repeated-letter (e.g., 75%) trial context.

phonetic responses associated with pronounceable irrelevant stimuli and relevant colour names, which leads to negative facilitation (see Parris et al., 2023, Figure 1); competition that is increased by reduced control over word reading induced by the mostly non-lexical trial context.

An alternative account of negative facilitation

Parris et al. (2023) pointed out that the observance of negative facilitation is determined by whether the baseline condition is non-lexical (e.g., repeated-letter strings) or lexical (colour-neutral words). Task conflict proponents argue that the difference between these two types of stimuli is the presence of task conflict on neutral word trials and its absence on repeated-letter trials (hence their large proportion puts the task controller on low alert). Indeed, some researchers (including ourselves) have even used the difference in reaction times (RTs) between these two types of stimuli as a measure of task conflict (Augustinova et al., 2018, 2019; Ferrand et al., 2020; Kinoshita et al., 2017, 2018). However, Parris et al. (2023) argued these stimuli also differ in *pronounceability*. Recent work indicates that pronounceable letter strings result in the sub-lexically generated phonology,

which then interferes with the segment-to-frame association processing in articulation planning, also known as phonological encoding (Kinoshita et al., 2017). Given that repeated-letter Stroop trials are free of this pronounceability cost, they might be responded to faster than congruent Stroop trials, which are burdened by the pronounceability cost, thereby producing negative facilitation (see Figure 1 for a visual representation of the two accounts of negative facilitation). Under this account, negative facilitation is not due to the fact that a large portion of repeated-letter string trials puts the task conflict controller on low alert allowing the expression of task conflict. Rather, a large portion of repeated-letter string trials reduces control over word reading and enhances the pronounceability cost of the infrequently presented pronounceable items, including congruent stimuli, to the point that congruent items are responded to slower than items that are free of pronounceability cost. Of course, if the pronounceability cost is accentuated, why would the pronounceability benefit offered by congruent trials not also be accentuated, producing more positive facilitation? Parris et al. (2023) pointed out that having a large proportion of neutral trials in the Stroop task, whether they are neutral words, repeated-letter strings, or shapes, has consistently resulted in the absence of positive facilitation. In

other words, there is something about a mostly neutral trial context that prevents positive facilitation (see the “General Discussion” for a fuller discussion of this issue). Thus, the conditions of the experiment remove the congruency benefit and congruent words become equivalent to neutral words. Under this account, there is no need for a task controller and instead the burden of control falls to the controller that is already responsible for reducing the impact of the phonetic code of the irrelevant stimulus during Stroop task performance.

A prediction that follows from this alternative account is that any factor that modifies the pronounceability cost will modify negative facilitation. Therefore, the present study examined the extent to which negative facilitation is modified by pronounceability, using *onset complexity* to manipulate pronounceability. Initially, we investigated whether there was an effect of onset complexity on Stroop task performance in a mostly neutral word trial context. We then changed the mostly neutral word context to a mostly non-lexical, repeated-letter trial context that was expected to induce negative facilitation.

Onset complexity

The onset of a word is the initial phonological unit that contains the initial consonant or blend of consonants. Simple onsets are those that consist of a single consonant followed by a vowel. The colour words *red* and *purple* are examples of colour words that have a simple onset. Complex onsets are those that consist of two or more consonants. The colour words *blue* and *green* have complex onsets.

Onset complexity has been shown to affect word reading latency. For example, Frederiksen and Kroll (1976) reported that words with complex onsets are read aloud more slowly than words with simple onsets. However, Kawamoto and Kello (1999) reported that words with complex onsets are read aloud more quickly than are words with simple onsets—exactly the opposite result. Kawamoto and Kello hypothesised that the inconsistency in their results was due to how response latency was measured. Rastle and Davis (2002) replicated Kawamoto and Kello’s (1999) results when response latency was based on hand-marking digitised responses, but reported an opposite effect when a simple voice-key was used, replicating Frederiksen and Kroll (1976). Moreover, Rastle and Davis (2002) reported no effect of onset complexity when an integrator voice-key was used. These results suggest that the effect of onset complexity on word reading latency can be determined by how acoustic latency is measured (Rastle & Davis, 2002; see also Kinoshita, 2000 vs. Schiller, 2008; for contrasting effects of onset complexity in the masked onset priming effect). However, given the clear benefits of visually inspecting the acoustic waveform and hand-marking onsets, the best evidence to date suggests a complexity

benefit (words with complex onsets are read more quickly than those with simple onsets; ~9 ms in the work of Rastle & Davis (2002) and ~5 ms in the work Kawamoto & Kello (1999). Nevertheless, it is noteworthy that in the Stroop task, the word is not actually read aloud. Instead, a colour is named aloud, as in the vocal Stroop task, or it is classified by button-press, as in the manual Stroop task. This permits consideration of onset complexity without the attendant issues associated with measuring acoustic latency.

It is further worth noting that an effect of onset complexity on word naming latencies in either direction has associated accounts: longer RTs for complex onsets have been accounted for by the notion of a longer activation for the first phoneme of a complex onset than for the single phoneme in a simple onset (MacKay, 1987). Shorter latencies for complex onsets have been accounted for by the notion of there being a larger number of competitor words at the initial phoneme position (O’Seaghdha et al., 1992; Taraban & McClelland, 1987) for simple onset stimuli (e.g., *sand* would have neighbours *band*, *hand*, *land*, *wand*)—here complexity is beneficial. Nevertheless, what is important for the purpose of the present experiments is that both accounts are based on the notion that the effect of onset complexity occurs at the stage at which the pronunciation is determined (Kawamoto & Kello, 1999). Therefore, in line with our initial reasoning about pronounceability cost in the Stroop task (see Parris et al., 2023) it remains plausible that onset complexity also determines the amount of this cost. As far as we are aware, no study has yet investigated this type of effect (but see Berent & Marom, 2005, who showed a skeletal congruency effect in the Stroop task that reflected the overlap of phonological frames between word and colour name). Thus, the aim of the present study was to fill this void and to investigate whether there is an onset complexity effect in the Stroop task (Experiment 1) before going on to explore the effect of onset complexity on negative facilitation (Experiment 2). To be clear, the alternative account does not predict that either simple or complex onsets are the sole cause of negative facilitation; if onset complexity modifies performance, negative facilitation will be bigger in one condition than in the other, but it should be present in both because relative to a non-pronounceable baseline, pronounceable stimuli will incur a processing cost.

Experiment 1

The aim of the present experiment was to investigate how onset complexity affects Stroop task performance. To investigate this issue initially and independently of factors thought to induce task conflict, the present experiment employed a mostly neutral word context (see Goldfarb & Henik, 2007, Experiment 2). As recent evidence indicates that phonological processing does occur with manual responses (Parris et al., 2019; but see Kinoshita et al., 2017; Kinoshita & Mills, 2020), the participants used

manual key-presses to identify the colour of different Stroop stimuli. The irrelevant dimension of these stimuli consisted of irrelevant colour words with simple (i.e., red and purple) and complex (i.e., blue and green) onsets. The onset complexity of colour-neutral words was further manipulated, such that, words with complex (e.g., dry) and simple (e.g., dog) onsets were used. If onset complexity affects pronunciation of any letter string, a difference will be observed between the neutral trials with simple vs. complex onsets. If, however, the locus of the onset complexity effect is at the level of activation of competitor responses (i.e., affects informational conflict), onset complexity will interact with Stroop effects. In the present experiment, we also employed two different sets of complex neutral and simple neutral words to avoid observed effects being due specifically to the choice of neutral words. Half the participants were exposed to one set, half to the other.

Method

Design. A 4 (word type: incongruent vs. congruent vs. neutral complex vs. neutral simple) \times 2 (onsets: complex vs. simple) repeated-measures design with RTs as the dependent variable.

Participants. Following Brysbaert and Stevens' (2018) recommendation for at least 1,600 trials per condition, in both experiments presented in this article, we aimed to initially recruit at least 67 participants. Moreover, test sensitivity of the interactions and main effects were assessed through Bayes factors and the optional stopping rule was applied for interactions (Rouder, 2014) after the initial 67 were recruited with the maximum number of participants limited to 90 participants given the funds available. In total, 76 participants were recruited through the online testing platform, Testable (testable.org). Participants had to fulfil the following criteria to participate: (1) be a verified mind, (2) have an approval rate above 95%, (3) be at least 18 years of age and below the age of 60, (4) have English as their first language (British or American), and (5) could not have participated in previous Stroop studies run on the host account. It was also requested that the Testable Minds platform matched the number of male and female participants. All participants were paid US\$4.50. Of the 75 participants, four were removed for reporting a language other than English as their first language and three were removed for having > 70% errors in any one of the experimental conditions. Of the remaining 69 participants, 36 were male. The ages ranged from 19 to 60 with an average age of 31.9 years (standard deviation [*SD*] = 10.3).

Stimuli and procedure. The colours red (RGB: 255, 0, 0), purple (RGB: 128, 0, 128), green (RGB: 0, 128, 0), and blue (RGB: 0, 0, 255) were used as the response colours.

These colours were chosen because the names *red* and *purple* have simple onsets, while *green* and *blue* have complex onsets and the two onset types were matched for word length (average 4.5 letters for both). The neutral words were matched to the colour words on the following dimensions: (1) number of letters, (2) number of phonemes, and (3) number of syllables. Complex graphemes, such as “ph” “sh” (two letters but one phoneme) and onsets with the first letters of the colour names R/B/G/P were avoided for the neutral stimuli. There were two types of neutral stimuli; for complex onset colour words there were both complex and simple onset neutral stimuli. Similarly, for the simple onset colour words, there were both complex and simple onset neutral stimuli. A comparison of the two types of neutral words that are matched on psycholinguistic variables permits measurement of potential onset complexity effects for neutral words. The complex onset neutral words matched to the complex onset colours blue and green were *crew* and *train* (complex neutrals for complex [CNC]) and the simple onset neutral words were *cake* and *taste* (simple neutrals for complex [SNC]). Similarly, for the simple onset colour words, there were both complex and simple onset neutral stimuli. The complex onset neutral words matched to the simple onset colours red and purple were *dry* and *frozen* (heretofore complex neutrals for simple [CNS]) and the simple onset neutral words were *dog* and *farmer* (simple neutrals for simple [SNS]; see Table 1 for the lexical characteristics of the stimuli used in Experiment 1).

Furthermore, a second set of neutral stimuli were employed to avoid any observed effects being due to the phonological characteristics of the onsets. In the second set, the CNS words were *fly* and *clever* and the SNS words were *far* and *carbon*. The CNC words were *tree* and *floor* and the SNC words were *tape* and *force*.

Once participants clicked the link to the experiment, they were directed to a reCaptcha challenge. Once completed, participants calibrated their screen to ensure all stimuli were presented the same size for all participants. Participants were then asked to enter their age, gender, nationality and first language. They were then presented with instructions asking them to respond as quickly and as accurately as they could to the colour of the font of the letter string presented at the screen's centre. On each trial, a fixation cross was presented for 2000 ms before the onset of the Stroop stimulus. The Stroop stimuli stayed on the screen until a response was made or until 2,000 ms had elapsed. Participants responded using the c (blue), b (purple), n (green), and v (red) keys on their keyboards. Incorrect responses resulted in feedback showing a large X (size + 5 in Testable), which was presented for 500 ms.

There were 384 trials in total. Forty-eight of the trials were incongruent, 48 were congruent. Of the 48 incongruent trials, 24 had complex onset, 24 had simple onsets; likewise for the congruent items. Of the 288 neutral word

Table 1. Stimulus characteristics in Experiment 1.

| | Colour words | | Neutral—Set 1 | | Neutral—Set 2 | |
|-----------------------------|--------------|--------|---------------|--------|---------------|--------|
| | Complex | Simple | Complex | Simple | Complex | Simple |
| Number of letters | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 |
| Number of phonemes | 3.5 | 3.5 | 3.75 | 3.75 | 3.75 | 4.0 |
| Number of syllables | 1.00 | 1.5 | 1.25 | 1.25 | 1.25 | 1.25 |
| Word frequency ^a | 87.43 | 80.19 | 50.18 | 75.26 | 69.47 | 91.37 |

Note. Stimulus characteristics for congruent stimuli are identical to those for incongruent stimuli. a: Subtitle word frequency (Brysbaert & New, 2009).

stimuli, 144 had simple onsets and 144 had complex onsets. Of the 144 simple onset neutral trials, 72 were matched to the complex onset colour words and 72 were matched to the simple onset colour words. Similarly, of the 144 complex onset neutral trials, 72 were matched to the complex onset colour words and 72 were matched to the simple onset colour words (see Table 2 for a stimulus matrix for Experiments 1–3). Again, having the two types of neutral words matched to both the complex and simple onset colour words meant that we could compare RTs with matched neutral words that differed in onset complexity but were matched on other psycholinguistic variables.

Results

Only data from correct trials are reported below since there were no effects in the error data other than main effects of word type (7.1% of data were removed as errors—see Table 3 for error percentages as a function of condition). Any RTs greater than 2 SDs either side of the overall mean were removed from analysis resulting in the removal of 2.8% of the correct trials.

The data were entered into a 4 (word type: incongruent vs. complex neutral vs. simple neutral vs. congruent) × 2 (onsets: complex vs. simple) repeated-measures analysis of variance (ANOVA) and its Bayesian equivalent (see Figure 2; and see Table 3 for means, percentage errors, and Stroop effects). These analyses revealed a main effect of word type, $F(1.844, 125.389)$, Greenhouse–Geisser = 79.561, $p < .001$, $\eta^2 = .381$, $BF_{10} = 7.22 \times 10^{30}$, but no main effect of onset, $F(1, 68) = 1.138$, $p = .290$, $BF_{01} = 6.255$. However, the interaction between word type and onsets was significant, $F(2.246, 152.704)$, Greenhouse–Geisser = 6.667, $p = .001$, $\eta^2 = .021$, with $BF_{incl} = 143$ for just the interaction (i.e., without the main effects also included in the model).

The interaction was further decomposed into two one-way repeated-measures ANOVAs and their Bayesian equivalents investigating Stroop effects in the complex onset and simple onset conditions independently. A significant ANOVA for the complex onset stimuli, $F(3, 204) = 62.883$, $p < .001$, $\eta^2 = .480$, $BF_{10} = 3.27 \times 10^{25}$, was underpinned by a significant overall Stroop or congruency effect (i.e., mean difference between congruent and incongruent trials) of 119 ms, $t(68) = 12.682$; $p(\text{Holm}) < .001$,

Table 2. Stimulus matrices for all three experiments.

| | Colour names | | | |
|-------------------------------------|--------------|--------|---------------|------|
| | Simple onset | | Complex onset | |
| | Red | Purple | Green | Blue |
| Experiments 1, 2, and 3 | | | | |
| Simple onset colour words | | | | |
| Red | 12 | 4 | 4 | 4 |
| Purple | 4 | 12 | 4 | 4 |
| Complex onset colour words | | | | |
| Green | 4 | 4 | 12 | 4 |
| Blue | 4 | 4 | 4 | 12 |
| Experiments 1 and 3 | | | | |
| Simple onset neutral words (Set 1) | | | | |
| Dog | | 12 | 12 | 12 |
| Farmer | 12 | | 12 | 12 |
| Taste | 12 | 12 | | 12 |
| Cake | 12 | 12 | 12 | |
| Complex onset neutral words (Set 1) | | | | |
| Dry | | 12 | 12 | 12 |
| Frozen | 12 | | 12 | 12 |
| Train | 12 | 12 | | 12 |
| Crew | 12 | 12 | 12 | |
| Experiment 1 | | | | |
| Simple onset neutral words (Set 2) | | | | |
| Far | | 12 | 12 | 12 |
| Carbon | 12 | | 12 | 12 |
| Force | 12 | 12 | | 12 |
| Tape | 12 | 12 | 12 | |
| Complex onset neutral words (Set 2) | | | | |
| Fly | | 12 | 12 | 12 |
| Clever | 12 | | 12 | 12 |
| Floor | 12 | 12 | | 12 |
| Tree | 12 | 12 | 12 | |
| Experiment 2 | | | | |
| Repeated-letter stimuli | | | | |
| Xxx | | 24 | 24 | 24 |
| Xxxxxx | 24 | | 24 | 24 |
| xxxxx | 24 | 24 | | 24 |
| xxxx | 24 | 24 | 24 | |

$d = 0.930$, $BF_{10} = 3.56 \times 10^{11}$. This latter effect resulted from a significant positive facilitation of 28 ms, $t(68) = 2.985$, $p(\text{Holm}) = .010$, $d = .711$, $BF_{10} = 23.056$, and

Table 3. RTs, errors, and Stroop effects as a function of onset complexity in Experiment 1.

| Onset | Word type | | | | Stroop effects | | | | | | | |
|------------|-----------------|---------|-----------|---------|------------------------------|---------|----------------|---------|-----------------|-------------------|-----------------|----------------------------|
| | Incongruent | | Congruent | | Complex neutral ^a | | Simple neutral | | Facilitation | Congruency effect | Interference | Complex vs. simple neutral |
| | Mean (SE) | % error | Mean (SE) | % error | Mean (SE) | % error | Mean (SE) | % error | | | | |
| Complex | 914 (17) | 9.8 | 795 (16) | 6.2 | 823 (14) | 7.1 | 816 (14) | 6.6 | 28* | 119* | 91* | 7 ^b |
| Simple | 900 (17) | 8.5 | 828 (16) | 6.9 | 813 (13) | 6.8 | 821 (14) | 7.2 | -7 ^b | 72* | 79* | -8 |
| Difference | 14 ^b | | 33* | | | | | | 35* | 47* | 12 ^b | |

RT: reaction time.

^aNote that half of the complex onset neutral trials were matched in length and frequency to the complex onset colour words and half were matched to the simple onset colour words. The same was true for the simple onset neutral stimuli. This meant that we could compare RTs for complex and simple onset neutral words that were matched on other psycholinguistic variables.

^bEvidence for null.

* $p < .01$.

a significant Stroop interference effect (91 ms), $t(68)=9.697$, $p(\text{Holm}) < .001$, $d = .711$, $\text{BF}_{10} = 1.99 \times 10^{10}$, when CNC stimuli were used as a colour-neutral comparison baseline. When SNC stimuli were used to this end, the aforementioned congruency effect resulted from significant positive facilitation of 22 ms, $t(68)=2.301$, $p(\text{Holm}) = .045$, $d = .219$ albeit where the Bayes factor was anecdotal, $\text{BF}_{10} = 1.85$; and a significant Stroop interference effect of 97 ms, $t(68)=10.381$, $p(\text{Holm}) < .001$, $d = .761$, $\text{BF}_{10} = 7.07 \times 10^{11}$. RTs to neutral stimuli with complex versus simple onsets did not differ significantly, 6 ms; $t(68)=0.684$, $p(\text{Holm}) = .495$, where the Bayes factor supported the null, $\text{BF}_{01} = 3.632$.

A significant ANOVA for the simple onset stimuli, $F(3, 204)=37.997$, $p < .001$, $\eta^2 = .358$, $\text{BF}_{10} = 2.91 \times 10^{16}$, was underpinned by a significant Stroop congruency effect of 72 ms, $t(68)=7.820$, $p(\text{Holm}) < .001$, $d = 0.571$, $\text{BF}_{10} = 268.861$. This latter effect resulted from a non-significant negative facilitation, (that is, -7 ms; $t(68)=0.761$, $p(\text{Holm}) = .775$, $\text{BF}_{10} = 0.204$, and a significant Stroop interference effect of 79 ms, $t(68)=8.581$, $p(\text{Holm}) < .001$, $d = .626$, $\text{BF}_{10} = 4.46 \times 10^7$, when CNS stimuli were used as a colour-neutral comparison baseline. When SNC stimuli were used to this end, the aforementioned congruency effect resulted from a non-significant negative facilitation, that is, -15 ms; $t(68)=1.627$, $p(\text{Holm}) = .316$, albeit where the Bayes factor was anecdotal, $\text{BF}_{10} = 0.585$, and a significant Stroop interference effect of 87 ms, $t(68)=9.446$, $p(\text{Holm}) < .001$, $d = .690$, $\text{BF}_{10} = 1.55 \times 10^9$. Again, RTs to neutral stimuli with complex versus simple onsets did not differ significantly, 8 ms; $t(68)=0.866$, $p(\text{Holm}) = .775$, where the Bayes factor favoured the null anecdotally, $\text{BF}_{01} = 2.179$.

Effect of onset complexity on the Stroop congruency effect. To assess the effect onset complexity on the Stroop congruency effect further, the data were entered into a 2 (onsets: complex vs. simple) \times 2 (word type: congruent vs.

incongruent) repeated-measures ANOVA and its Bayesian equivalent. These analyses revealed a significant main effect of word type, $F(1, 68)=105.631$, $p < .001$, $\eta^2 = .407$, $\text{BF}_{10} = 3.715 \times 10^{12}$, and as in previous analyses, no main effect of onset, $F(1, 68)=2.401$, $p = .126$, $\text{BF}_{01} = 3.273$. However, the interaction was also significant, $F(1, 68)=9.662$, $p = .003$, $\eta^2 = .025$, $\text{BF}_{\text{incl}} = 47$ for just the interaction, indicating strong support for an effect of onset complexity on the Stroop congruency effect.

Post hoc tests revealed that this interaction was due to a larger significant congruency effect (119 ms) for complex, $t(68)=9.922$, $p(\text{Holm}) < .001$, $d = .864$, $\text{BF}_{10} = 3.559 \times 10^{11}$, than for simple onsets (72 ms), $t(68)=5.984$, $p(\text{Holm}) < .001$, $\text{BF}_{10} = 268.861$. Comparing word types across onset complexity conditions revealed that onset complexity mainly takes its effect on congruent items, $t(68)=3.371$, $p = .002$, $d = .244$, $\text{BF}_{10} = 18.309$, as the effect on Incongruent items remained non-significant, $t(68)=1.385$, $p = .171$, with the Bayes factor providing moderate evidence for the null, $\text{BF}_{01} = 3.042$. Therefore, this latter effect of onset complexity on the Stroop congruency effect is likely to result from changes on Stroop facilitation instead of interference. Two supplementary ANOVAs and their Bayesian equivalents were then run to address this idea directly.

Effect of onset complexity on Stroop interference. A 2 (word type: incongruent vs. neutral) \times 2 (onsets: complex vs. simple) repeated-measures analyses using the neutral trials that matched to the colour word for onset type (CNC and SNS for complex and simple onset Stroop effects, respectively) further revealed the main effect of word type, $F(1, 68)=92.828$, $p < .001$, $\eta^2 = .429$, $\text{BF}_{10} = 3.24 \times 10^{11}$. However, both main effect of onset, $F(1, 68)=1.741$, $p = .191$, $\text{BF}_{10} = 0.318$ and the word type \times onset interaction remained non-significant, $F(1, 68)=1.248$, $p = .268$; with $\text{BF}_{\text{incl}} = 0.302$ for just the interaction, further suggesting that the aforementioned onset complexity effect on the

magnitude of Stroop congruency is unlikely to result from changes on Stroop interference and results from changes in facilitation instead.

Effect of onset complexity on Stroop facilitation. This latter idea is directly suggested by the results of a 2 (word type: congruent vs. neutral) \times 2 (onsets: complex vs. simple) repeated-measures ANOVAs revealing an interaction between word type and onsets, $F(1, 68)=8.581, p=.005, \eta^2=.045$, with $BF_{incl}=57$ indicating strong support for the interaction alone, in addition to main effect of both word type, $F(1, 68)=4.407, p<.040, \eta^2=.016, BF_{10}=106$ and onset, $F(1, 68)=8.819, p=.004, \eta^2=.038, BF_{10}=3.588$, albeit with an anecdotal Bayes Factor favouring the null, $BF_{10}=0.608$. Further post hoc tests additionally revealed that the interaction was due to significant *positive* facilitation of 28 ms for complex onsets, $t(68)=3.594, p(\text{Holm})=.002, d=.224, BF_{10}=23$, and non-significant *negative* facilitation of -7 ms for simple onsets, $t(68)=0.896, p(\text{Holm})>.9, BF_{10}=0.204$.

Discussion

The aim of the present study was to explore how onset complexity affects Stroop task performance. The results from the present experiment convincingly show an onset complexity effect on Stroop task performance, such that, complex onsets increase the magnitude of positive facilitation and the Stroop congruency effect. Both effects were driven by the fact that congruent words with complex onsets were responded to more quickly than congruent words with simple onsets. This result, along with the null effects of onset complexity on both incongruent and neutral words trials suggest the onset complexity effect was driven by the activation of the phonological/phonetic representation of the correct response, facilitating the classification of the colour (this is discussed further in the “General Discussion” section). This facilitating effect of complex onsets is consistent with the notion that subvocal production of the colour name occurs even with a manual response (Parris et al., 2019; but see Kinoshita & Mills, 2020), and is also consistent with a complexity advantage in word reading (Kawamoto & Kello, 1999; Rastle & Davis, 2002).

The finding of robust positive facilitation in a mostly neutral word context is somewhat unusual. Complex onsets appear to have created positive facilitation where otherwise there would not have been. As Parris et al. (2023) have pointed out, when presented in a mostly neutral word trial context, positive facilitation is not observed (Goldfarb & Henik, 2007, Experiment 2; see also Shichel & Tzelgov, 2018); rather there is no positive nor negative facilitation. This is in fact what we observe in the simple onset condition in the present experiment. Thus, the finding of robust positive facilitation in the complex onset

condition is notable and indicates an extra level of information contributing to performance.

Given the effect on congruent trials, it is somewhat surprising that the subvocal production of the incongruent colour name—corresponding to the incorrect response—did not also modify performance. While a larger magnitude of interference was observed for the complex onset incongruent stimuli (i.e., 97 ms as compared with 79 ms), the data are best interpreted as showing no effect of onset complexity on incongruent trial RTs. It is also the case that there was no onset complexity effect for neutral stimuli. An effect of onset complexity on neutral and incongruent trials would be expected if the onset complexity effect in the Stroop task were driven by just the phonological processing of the irrelevant word. However, it is also notable that the size of the effect in the present study (~ 35 ms) was much larger than that reported in studies of intentional word reading (~ 7 ms). Taken together, the present results indicate that the onset complexity effect in the Stroop task is different from that seen in studies of intentional word reading. The lack of an effect of onset complexity on incongruent and neutral trials suggests that the onset complexity effect in the Stroop task requires that the phonological codes of the word and colour name match; only then is the complexity advantage observed. Indeed, the lack of an onset effect on neutral and incongruent trials represents a limitation of the use of onset complexity as a manipulation; If the onset effect were apparent in the neutral and incongruent trials in the mostly xxxx trial context, it would have least enabled us to argue that the disappearance of negative facilitation is indicative of negative facilitation reflecting a different form of conflict to informational conflict.

In intentional reading, the onset complexity effect results from the faster compilation of the phonetic code of the written word but the results from the present study suggest that in the Stroop task the complexity advantage is related to the benefit the phonetic code of the written word has on the subvocal production of the acoustic code of the colour name. Thus, despite the difference in the ease with which neutral words with complex versus simple onsets are read aloud (Rastle & Davis, 2002), there appear to be no implications for colour responding. It is not therefore the pronounceability of the irrelevant word alone that matters; it is the relationship between the pronounceability of the irrelevant word and the correct colour name. Onsets that do not have congruent phonetic codes, whether they refer to incongruent colours or neutral words, do not influence colour categorisation performance.

Experiment 2

The aim of this experiment was to investigate whether onset complexity affects negative facilitation. To do this, the mostly neutral word context employed in Experiment 1 was replaced by a mostly non-pronounceable, repeated-letter

stimuli context; a context shown to repeatedly produce negative facilitation (Entel & Tzelgov, 2018, 2020; Goldfarb & Henik, 2007; Kalanthroff et al., 2013). Given the effect of onset complexity reported above, it was predicted that colour-identification times for colour-congruent words with complex onsets would reduce or eliminate negative facilitation relative to irrelevant words with simple onsets. Such a finding would be consistent with the alternative account of negative facilitation based on differences in pronounceability. In contrast to the alternative account, the task conflict account of negative facilitation indicates that the longer RTs to congruent trials result from competition between task sets. As words with both complex and simple onsets equally fulfil the criteria of being words, they should both equally activate the task set for word reading and thus produce similar amounts of negative facilitation.

As with Experiment 1, in the present experiment, we contrasted the effect of irrelevant colour words with simple onsets (i.e., red and purple) and irrelevant colour words with complex onsets (i.e., blue and green) on Stroop task performance. To ensure orthogonal comparisons for facilitation and interference effects, half the repeated-letter stimuli were employed as the neutral baseline for the complex onset colour word stimuli and half were employed as the baseline for the simple onset colour word stimuli (and the repeated-letter stimuli sets were matched for length with their word counterparts).

Method

Design. A 3 (word type: incongruent vs. congruent vs. repeated-letter neutral) \times 2 (onsets: complex vs. simple) repeated-measures design with RTs as the dependent variable.

Participants. Ninety new participants were recruited through the online testing platform, Testable (testable.org). The inclusion criteria were the same as those in Experiment 1. All participants were paid US\$4.50. Of the 90 participants, nine were removed for reporting a language other than English as their first language, three were removed for having > 70% errors in any one of the experimental conditions, and three were removed because they had previously completed a similar study. Of the remaining 74 participants, 45 were male. The ages ranged from 21 to 57 with an average age of 32.7 ($SD=9.4$).

Stimuli and procedure. The stimuli and procedure were identical to Experiment 1 with the exception of the repeated-letter baseline that replaced the neutral-word baseline. The repeated-letter stimuli used in this experiment were matched in length to the colour words (e.g., xxx for red, xxxxx for green). The stimuli were presented in Testable's default font size and style.

Results

Only data from correct trials were used in this analysis (7.8% of data were removed as errors). None of the error analyses produced anything other than a main effect of word type and thus are not reported below (see Table 4 for error percentages). Any RTs greater than 2 SD s either side of the overall mean were removed from analysis resulting in the removal of 5.3% of correct trials.

Effect of onset complexity on negative facilitation. To assess the effect of onset complexity on negative facilitation, the data were entered into a 2 (word type: congruent vs. repeated-letter) \times 2 (onset: complex vs. simple) repeated-measures ANOVA and its Bayesian equivalent. Analysis revealed no main effect of onset, $F(1, 73)=0.174$, $p=.678$ with $BF_{01}=6.579$ supporting the null, and a non-significant onset \times word type interaction, $F(1, 73)=0.003$, $p=.953$ with $BF_{01}=40.667$, showing strong evidence against the interaction. However, there was also not strong evidence for the presence of negative facilitation. Indeed, the main effect of word type was significant with standard inferential ANOVA, $F(1, 73)=5.445$, $p=.022$, $\eta^2=.021$, but the Bayes factor was anecdotal, $BF_{10}=1.082$ (see Figure 3; see Table 4 for RTs, errors, and Stroop effects).

Effect of onset complexity on the Stroop congruency effect. While the main aim of the present experiment was to explore the effect of onset complexity on negative facilitation, here, we analyse the effect of onset complexity on the Stroop congruency effect to permit a more direct comparison with Experiment 1. The data were entered into a 2 (onset: complex vs. simple) \times 2 (word type: congruent vs. incongruent) repeated-measures ANOVA. This analysis revealed no main effect of onset where $F(1, 73)=0.726$, $p=.397$, $BF_{01}=5.470$. The main effect of word type was significant where $F(1, 73)=161.249$, $p<.001$, $\eta^2=.487$, $BF_{10}=1.416 \times 10^{17}$. The main effect was the result of an overall congruency effect of 118.1 ms where $t(73)=12.698$, $p(\text{Holm})<.001$, $BF_{10}=1.416 \times 10^{17}$. The onset and word type interaction was not significant, $F(1, 73)=1.076$, $p=.303$, and this was supported by the Bayes factor for the contribution of the interaction alone ($BF_{\text{excl}}=4.725$).

In sum, while we observed at least some evidence for negative facilitation, this was not modified by onset complexity. Furthermore, there was evidence against an effect of onset complexity on the Stroop congruency effect.

Discussion

Despite employing an experimental context that has been shown to produce robust negative facilitation effects in previous studies (i.e., 75% repeated-letter stimuli; e.g., Entel & Tzelgov, 2018; Goldfarb & Henik, 2007; Kalanthroff et al., 2013), we were only able to provide

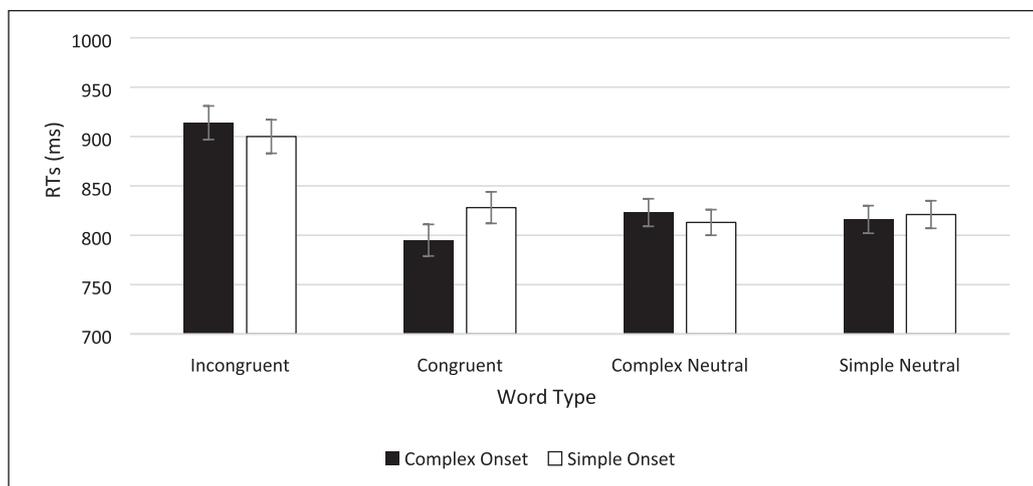


Figure 2. Button press colour classification times in the Stroop task as a function of word type and onset type in Experiment 1.

Table 4. RTs (ms), errors, and Stroop effects as a function of onset complexity in Experiment 2.

| Onset | Word type | | | | | | Stroop effects | | |
|------------|-------------|---------|----------------|---------|-----------------|---------|----------------|-------------------|---------------------|
| | Incongruent | | Congruent | | Neutral (xxxxx) | | Facilitation | Congruency effect | Interference effect |
| | Mean (SE) | % Error | Mean (SE) | % error | Mean (SE) | % error | | | |
| Complex | 955 (16) | 10.8 | 828 (16) | 7.3 | 815 (14) | 7.6 | -13 | 127** | 140** |
| Simple | 940 (16) | 11.7 | 831 (16) | 7.4 | 817 (13) | 7.2 | -14* | 109** | 123** |
| Difference | 15 | | 3 ^a | | | | 1 ^a | 18 | 17 |

RT: reaction time.

^aEvidence for null.

* $p < .05$, ** $p < .001$.

weak evidence for negative facilitation in the present experiment. Parris et al. (2023) noted that for large and robust negative facilitation to be produced via the mostly non-lexical trial context, two other factors need to be present: (1) spare working memory capacity and (2) anticipation of informational conflict in the form of incongruent trials (see Entel & Tzelgov, 2018, 2020). In the present experiment, there was no working memory load imposed on the participants and incongruent trials were included. It is therefore somewhat surprising that the overall main effect of negative facilitation was so small.

In terms of the main aim of the current experiment, the Bayes factor provided evidence for no effect of onset complexity on negative facilitation and, furthermore, there was no effect of onset complexity evident at all in the data—congruent trials were unaffected by onset type. Given that the only difference between the present experiment and Experiment 1 was the presence of the mostly repeated-letter (Experiment 2) context, it is likely that the mostly repeated-letter context was responsible for the lack of an onset complexity effect in the present experiment. However, before interpreting the result from the present experiment, we felt it important to first replicate the onset complexity effect observed in Experiment 1.

Experiment 3

Given the lack of an onset complexity effect in Experiment 2, the aim of the present experiment was to replicate the onset complexity effect observed in Experiment 1. Therefore, the following experiment was identical to Experiment 1 except that only one set of neutral words was used since there was no difference between the two sets in Experiment 1 ($p > .05$).¹

Method

Design. A 4 (word type: incongruent vs. congruent vs. neutral complex vs. neutral simple) \times 2 (onsets: complex vs. simple) repeated-measures design with RTs as the dependent variable.

Participants. In total, 124 participants were recruited through the online testing platform, Testable (testable.org). Participants had to fulfil the same criteria as the previous two experiments. None of them participated in Experiments 1 and 2. All participants were paid US\$5. Of the 124 participants, eight were removed for reporting a language other than English as their first language and six were

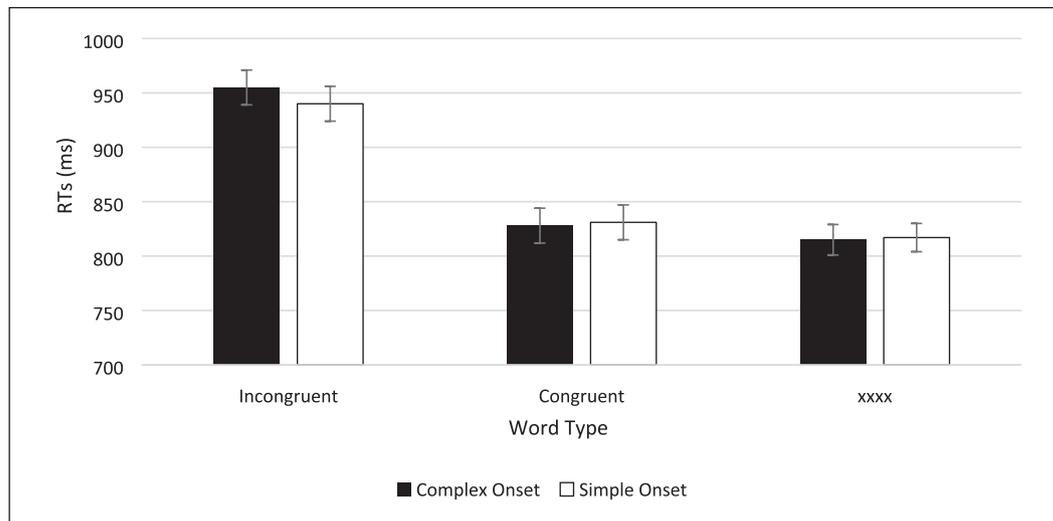


Figure 3. Button press colour classification times in the Stroop task as a function of word type and onset type in Experiment 2.

Table 5. RTs, errors, and Stroop effects as a function of onset complexity in Experiment 3.

| Onset | Word type | | | | | | | | Stroop effects | | | |
|------------|----------------|---------|-----------|---------|-----------------|---------|----------------|---------|---------------------|------------|----------------|----------------------------|
| | Incongruent | | Congruent | | Complex neutral | | Simple neutral | | Facilitation effect | Congruency | Interference | Complex vs. simple neutral |
| | Mean (SE) | % error | Mean (SE) | % error | Mean (SE) | % error | Mean (SE) | % error | | | | |
| Complex | 945 (12) | 12.1 | 829 (12) | 7.9 | 861 (11) | 8.1 | 859 (11) | 8.5 | 32* | 116* | 84* | 2 ^a |
| Simple | 942 (12) | 11.3 | 858 (13) | 8.4 | 860 (11) | 8.5 | 859 (11) | 8.1 | 1 ^a | 84* | 83* | 1 ^a |
| Difference | 3 ^a | | 29* | | | | | | 31* | 32* | 1 ^a | |

RT: reaction time.

^aEvidence for null.

* $p < .01$.

removed for having $> 70\%$ errors in any one of the experimental conditions. Of the remaining 110 participants, 52 were male. The ages ranged from 19 to 64 with an average age of 32.9 years ($SD = 10.1$).

Stimuli and procedure. Identical to Experiment 1 with the exception that: (1) only one set (Set 1) of neutral words were used given no differences were observed between the two sets in Experiment 1 ($p > .8$) and (2) an additional 48 repeated-letter trials were included to enable the calculation of facilitation effects relative to this new baseline. However, as noted in footnote 1, a programming error meant that the data from this condition were unusable and are therefore not reported below.

Results

Only data from correct trials are reported below since there were no effects in the error data other than main effects of word type (8.7% of data were removed as errors—see Table 5 for error percentages as a function of condition).

Any RTs greater than 2 SD s either side of the overall mean were removed from analysis resulting in the removal of 5.1% of the correct trials.

The data were entered into a 4 (word type: incongruent vs. complex neutral vs. simple neutral vs. congruent) \times 2 (onsets: complex vs. simple) repeated-measures ANOVA and its Bayesian equivalent (see Figure 4; see Table 5 for means, percentage errors, and Stroop effects). These analyses revealed a main effect of word type, $F(1.924, 209.7)$, Greenhouse–Geisser = 147.917, $p < .001$, $\eta^2 = .398$, $BF_{10} = 6.250 \times 10^{56}$, and a main effect of onset, $F(1, 109) = 4.776$, $p = .031$, although this was not supported by the Bayes factor that favoured the null, $BF_{10} = 0.314$. The interaction between word type and onsets was also significant, $F(2.103, 229.246)$, Greenhouse–Geisser = 5.087, $p = .006$, $\eta^2 = .011$ with $BF_{incl} = 24.137$ for just the interaction (i.e., without the main effects also included in the model).

The interaction was further decomposed into two one-way repeated-measures ANOVAs and their Bayesian equivalents investigating Stroop effects in the complex

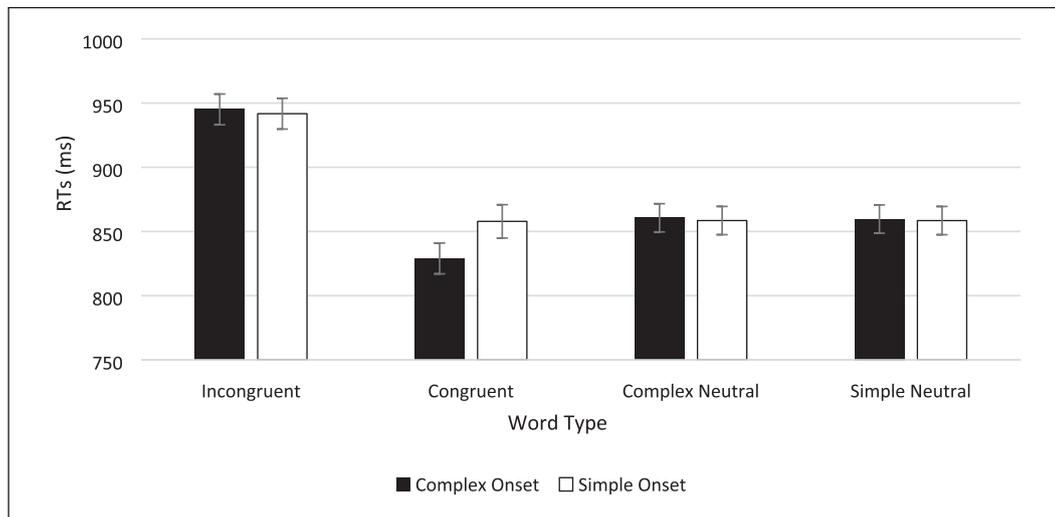


Figure 4. Button press colour classification times in the Stroop task as a function of word type and onset type in Experiment 3.

onset and simple onset conditions independently. A significant ANOVA for the complex onset stimuli, $F(3, 327)=108.593$, $p < .001$, $\eta^2 = .499$, $BF_{10} = 1.722 \times 10^{45}$, was underpinned by a significant overall Stroop congruency effect (i.e., mean difference between congruent and incongruent trials) of 116 ms, $t(109)=17.115$; $p(\text{Holm}) < .001$, $d = 0.968$, $BF_{10} = 8.796 \times 10^{21}$. The congruency effect resulted from significant positive facilitation of 32 ms, $t(109)=4.668$, $p(\text{Holm}) < .001$, $d = .264$, $BF_{10} = 5.088$, and a significant Stroop interference effect (84 ms), $t(109)=12.447$, $p(\text{Holm}) < .001$, $d = .704$, $BF_{10} = 1.939 \times 10^{18}$ when CNC stimuli were used as a colour-neutral comparison baseline. When SNC stimuli were used, the congruency effect resulted from significant positive facilitation of 30 ms, $t(109)=4.353$, $p(\text{Holm}) < .001$, $d = .246$, $BF_{10} = 1.219$, and a significant Stroop interference effect of 87 ms, $t(109)=12.762$, $p(\text{Holm}) < .001$, $d = .721$, $BF_{10} = 5.728 \times 10^{17}$. And consistent with Experiment 1, RTs to neutral stimuli with complex versus simple onsets did not differ significantly, 2 ms; $t(109)=0.315$, $p(\text{Holm}) = .753$, where the Bayes factor supported the null, $BF_{10} = 0.124$.

A significant ANOVA for the simple onset stimuli, $F(3, 327)=62.590$, $p < .001$, $\eta^2 = .365$, $BF_{10} = 4.933 \times 10^{28}$, was underpinned by a significant Stroop congruency effect of 84 ms, $t(109)=11.294$, $p(\text{Holm}) < .001$, $d = 0.688$, $BF_{10} = 9.508 \times 10^9$. This latter effect resulted from non-significant negative facilitation, that is, -2 ms; $t(109)=0.239$, $p(\text{Holm}) > .9$, $BF_{10} = 0.109$, and a significant Stroop interference effect of 82 ms, $t(109)=11.054$, $p(\text{Holm}) < .001$, $d = .674$, $BF_{10} = 1.813 \times 10^{15}$, when CNS stimuli were used as a colour-neutral comparison baseline. When SNC stimuli were used to this end, the aforementioned congruency effect resulted from non-significant negative facilitation, that is, -1 ms; $t(109)=0.82$, $p(\text{Holm}) > .9$, $BF_{10} = 0.106$, and a significant Stroop interference effect of 83 ms, $t(109)=11.212$, $p(\text{Holm}) < .001$, $d = .683$, $BF_{10} = 1.532 \times 10^{16}$. Again, RTs to

neutral stimuli with complex versus simple onsets did not differ significantly, 1 ms; $t(109)=0.157$, $p(\text{Holm}) > .9$, where the Bayes factor favoured the null, $BF_{10} = 0.112$.

Effect of onset complexity on the Stroop congruency effect. To assess the effect onset complexity on the Stroop congruency effect further, the data were entered into a 2 (onsets: complex vs. simple) \times 2 (word type: congruent vs. incongruent) repeated-measures ANOVA and its Bayesian equivalent. These analyses revealed a significant main effect of word type, $F(1, 109)=201.653$, $p < .001$, $\eta^2 = .426$, $BF_{10} = 1.936 \times 10^{23}$, and a main effect of onset, $F(1, 109)=5.587$, $p = .020$, but where the Bayes factor was anecdotal $BF_{10} = 0.732$. The interaction was also significant, $F(1, 109)=6.382$, $p = .013$, $\eta^2 = .011$, $BF_{\text{incl}} = 7.077$ for just the interaction, indicating strong support for an effect of onset complexity on the Stroop congruency effect.

Post hoc tests revealed that this interaction was due to a larger significant congruency effect (116 ms) for complex, $t(109)=12.220$, $p(\text{Holm}) < .001$, $d = .910$, $BF_{10} = 8.796 \times 10^{21}$, than for simple onsets (84 ms), $t(109)=8.829$, $p(\text{Holm}) < .001$, $BF_{10} = 9.508 \times 10^9$. Comparing word types across onset complexity conditions revealed that onset complexity takes its effect on congruent items, $t(109)=3.370$, $p = .001$, $d = .321$, $BF_{10} = 20.646$, and not on incongruent items remained non-significant, $t(109)=0.410$, $p = .683$, with the Bayes Factor providing moderate evidence for the null, $BF_{10} = 0.115$. Therefore, as shown in Experiment 1, the effect of onset complexity on the Stroop congruency effect results from changes to congruent trials.

Effect of onset complexity on Stroop interference. A 2 (word type: incongruent vs. neutral) \times 2 (onsets: complex vs. simple) repeated-measures analyses using the neutral trials that matched to the colour word for onset type (CNC and

SNS for complex and simple onset Stroop effects, respectively) revealed no word type \times onset interaction, $F(1, 109)=0.018$, $p=.893$; with $BF_{incl}=0.123$ for just the interaction.

Effect of onset complexity on Stroop facilitation. Consistent with Experiment 1, a 2 (word type: congruent vs. neutral) \times 2 (onsets: complex vs. simple) repeated-measures ANOVAs revealed an interaction between word type and onsets, $F(1, 109)=9.410$, $p=.003$, $\eta^2=.033$, with $BF_{incl}=259$, indicating strong support for the interaction alone. Further post hoc tests additionally revealed that the interaction was due to significant *positive* facilitation of 32 ms for complex onsets, $t(109)=4.840$, $p(\text{Holm}) < .001$, $d=.257$, $BF_{10}=5.089$, and non-significant *positive* facilitation of 1 ms for simple onsets, $t(109)=0.093$, $p(\text{Holm}) > .9$, $BF_{10}=0.106$.

Discussion

Given the lack of an onset effect in Experiment 2, the aim of the present experiment was to replicate the onset effect observed in Experiment 1. The results from the present experiment provide a clear replication of Experiment 1 showing faster colour responses to congruent words with complex onsets than to congruent words with simple onsets. This resulted in an onset complexity effect on the Stroop congruency and positive facilitation effects, but not on interference effects. Moreover, the results from the present experiment confirm that colour categorisation responses to neutral words with complex onsets do not differ from colour categorisation responses to neutral words with simple onsets. The results from both Experiments 1 and 3 therefore indicate that colour congruency is required to observe an effect of onset complexity on Stroop task performance.

General discussion

The aim of the present set of experiments was twofold: (1) To investigate whether onset complexity modifies the magnitude of the common Stroop effects of Stroop interference, Stroop (positive) facilitation and the Stroop congruency effect (Experiments 1 and 3) and (2) Investigate whether onset complexity modifies the magnitude of negative facilitation (Experiment 2). Regarding the first aim, the results from both Experiments 1 and 3 provide strong evidence for an effect of onset complexity on Stroop effects. Specifically, both positive facilitation and the Stroop congruency effect (Experiments 1 and 3) were increased when irrelevant colour words had complex onsets (e.g., *blue* and *green* vs. *red* and *purple*) due to the effect of onset complexity on congruent trials. There was no evidence for an effect of onset complexity on Stroop interference nor when the irrelevant word was a neutral,

non-colour-related word indicating that the effect of onset complexity was at the level of encoding the phonology of the correct response. Indeed, as noted above, the lack of an effect on incongruent and neutral trials represents somewhat of a limitation for the use of onset complexity for present purposes: if the onset effect was apparent in the neutral and incongruent trials in the mostly xxxx trial context, it would have least enabled us to argue that the disappearance of negative facilitation is indicative of the phenomenon reflecting a different form of conflict to informational conflict.

The finding of an effect of onset complexity in the Stroop task is novel and is consistent with the complexity advantage observed in word reading studies (Kawamoto & Kello, 1999; Rastle & Davis, 2002). The results are consistent with the notion that complex onsets are beneficial to word processing (O'Seaghdha et al., 1992; Taraban & McClelland, 1987). Furthermore, this result was obtained in a task in which the words are not read aloud, and moreover, responses were manual, supporting the notion that phonological processing of the irrelevant word happens with manual responses (Parris et al., 2019; cf. Kinoshita et al., 2018; see below for a discussion of this issue). However, it is clear that the onset complexity effect reported here is not the same as that reported in studies of intentional reading. The magnitude of the effect was much larger and the effect was limited to when the complex onset of the irrelevant word matched the complex onset of the colour name; colour name congruency is the *sine qua non* of the onset complexity effect in Stroop task performance. In the Stroop task, it is the match between the phonological code of the irrelevant word and the phonetic code of the colour name that results in the onset complexity effect.

In terms of the second aim of the present study, there was no effect of onset complexity on negative facilitation. Indeed, there was no evidence of an onset effect in Experiment 2 indicating that the mostly repeated-letter context prevented the onset effect observed in Experiments 1 and 3. Onset complexity was manipulated because we thought that any factor that makes phonetic encoding more difficult (simple onset stimuli in the present study) would delay colour naming, enhancing negative facilitation. This is in contrast to the task conflict account of negative facilitation, which states that congruent trials RTs are longer because they involve task conflict while non-lexical trials do not. And while simple onset stimuli did delay colour naming of congruent items relative to complex onset stimuli in Experiments 1 and 3, it did not modify negative facilitation in Experiment 2. It is notable, however, that we did not observe large negative facilitation effects despite including experimental conditions that produce it, such as a mostly non-lexical trial context, spare working memory capacity, and exposure to incongruent trials (Parris et al., 2023). Nevertheless, the small main effect of negative

facilitation would not in itself preclude variation in the two onset conditions. However, the evidence for a null interaction effect was clear and therefore the best interpretation of the data is in line with the task conflict account of negative facilitation: Given that both complex and simple onset words equally fulfil the criteria of being words meaning the task conflict account predicts they would both produce similar amounts of negative facilitation relative to a non-lexical baseline.

Since the onset complexity effect observed in Experiments 1 and 3 is one based on positive facilitation, the absence of an onset effect in Experiment 2 could be explained by the general effect of mostly neutral word or mostly non-lexical trial contexts on positive facilitation (Parris et al., 2023; see below for a fuller discussion of this issue). Parris et al. (2023) noted that positive facilitation seems to be prevented when the experimental context is either mostly neutral words or mostly non-lexical stimuli. This would mean that an effect that finds its expression only in positive facilitation would be unlikely to be observed in such contexts. This means that the onset effect is not therefore an ideal test of the alternative account. If the onset complexity effect were also observable in other measures of informational conflict, such as the Stroop interference effect, it could have potentially produced an effect of onset complexity in Experiment 2 in all conditions except the congruent condition thereby indicating that negative facilitation is qualitatively different from informational conflict (the alternative account argues that negative facilitation is just another form of informational (phonological) conflict). Without this more pervasive effect on indices of Stroop task performance, it is unclear whether a strong test of the alternative account has been provided.

The effect of a mostly non-lexical trial context

The finding of a lack of an onset complexity effect in Experiment 2 stands in stark contrast to the findings of Experiments 1 and 3. The only difference between the experiments is the mostly repeated-letter context in Experiment 2 (which replaced the mostly neutral word context in Experiment 1). It therefore seems clear that the mostly repeated-letter context is responsible for the lack of an onset effect. What is unclear is why the mostly repeated-letter context would have this effect. Indeed, Parris et al. (2023) pointed out that the effect of the mostly repeated-letter context needs further explanation given two opposing accounts from task conflict theorists: according to Kalanthroff et al.'s (2018) model the mostly non-lexical context induces a low task conflict control state because task conflict is not encountered often enough, reducing proactive control (see also Spinelli and Lupker [2021, 2022] who argue that the mostly non-lexical trial context reduces proactive control—although not in a way that is

related to task conflict per se). Given that proactive control is low, task conflict is high, which leads to an inhibition mechanism modifying the response threshold to all lexical stimuli. This raising of the response threshold would not happen for repeated-letter string trials (e.g., xxxx) because the task unit for word reading would not be activated for these trials. As responses to congruent trials would be slowed relative to non-lexical trials under these conditions, negative facilitation would result. And since this raising of the response threshold would occur equally for all lexical stimuli it would not differentiate words with simple and complex onsets, which would therefore not differ in terms of negative facilitation—as supported by the results from the present study.

In contrast to this account, Entel and Tzelgov (2020) reasoned that a larger number of non-lexical trials means that participants are less likely to inadvertently read the congruent word, thereby increasing RTs to congruent words, and revealing negative facilitation and thus task conflict relative to an xxxx baseline. This account would also presumably predict reduced phonological processing of congruent trials in this context. If phonological processing of words is reduced, one would expect mitigation of phonological-based effects, especially in a manual response Stroop task (Parris et al., 2019). A future experiment investigating onset complexity with a vocal response Stroop task that promotes phonological processing of irrelevant stimuli might provide an interestingly contrasting set of results.

In contrast to both of these accounts, the alternative account of negative facilitation, tested in the present study, argued that the mostly repeated-letter, non-pronounceable trial context reduces control over word processing and therefore predicts larger informational conflict and facilitation effects. This position was motivated by the finding of Kinoshita et al. (2018) who reported larger semantic conflict effects relative to a neutral-word baseline in the mostly repeated-letter trial context (see also Shichel & Tzelgov, 2018, for evidence of greater semantic conflict in a mostly repeated-letter trial context). According to this position the onset complexity effect should be larger in the mostly repeated-letter trial context of Experiment 2 than in the mostly neutral word context of Experiment 1 because control over phonological (as well as semantic) processing should be reduced and onset effects therefore, larger, in the former. This position finds no support in the current set of results (although the position might still hold true for semantic conflict effects). In sum, the present results are consistent with the notion that the mostly repeated-letter trial context discourages phonological processing of presented letter strings and perhaps reduces inadvertent reading of congruent items. However, since there is evidence that proactive control is reduced in this mostly non-lexical trial context (Spinelli & Lupker, 2021, 2022) it is possible that

positive facilitation results from proactive control and thus intentional and not inadvertent reading. If this is the case, then the effect of the mostly non-lexical context is not via lessening (Entel & Tzelgov, 2019) or increasing (Parris et al., 2023) word processing, but on the amount of proactive control that operates during task performance (Kalanthoff et al., 2018; Spinelli & Lupker, 2021, 2022).

The lack of an effect on onset complexity on neutral trials

A further notable finding from the present study was the lack of an effect of onset complexity on neutral word stimuli in Experiments 1 and 3. This finding has implications for another account of performance in selective attention paradigms. Parris et al. (2023) acknowledged the similarities between their alternative account of negative facilitation and the response exclusion hypothesis (REH; Mahon et al., 2007) in the picture–word interference literature. Under the REH, irrelevant words get obligatorily processed right up to the point of a representation entering an articulatory buffer; no selection occurs before this very late point in processing and selection does not involve selection by competition. Under this account, words have privileged access to the articulators. Thus, as with Roelofs (2003) and Glaser and Glaser (1989) models that are based on architectural differences between word reading/naming and colour naming, the REH is based on architectural differences between word reading/naming and picture naming. Specifically, the REH describes this privileged access as being based on the “quasi rule-like relationship between orthography and phonology” (p. 524; Mahon et al., 2007) and as such leads to a “production-ready” representation for the articulators to produce. The REH uniquely predicts that low-frequency words should interfere more than high-frequency words, a finding—mirrored in the colour-word Stroop literature (e.g., Burt, 2002)—that theories based on connectionist architecture (e.g., Cohen et al., 1990), find difficult to explain. Under the REH, distractor frequency effects arise because the earlier the response to the distractor enters the articulatory buffer, the earlier it can be removed from the buffer. As low-frequency words would take longer to reach the buffer, it takes longer to remove them from the buffer and thus colour naming times would be slowed.

The aim of the present work was to test and further develop the alternative account of negative facilitation presented in Parris et al. (2023). The alternative account, like the REH, assumes that the phonological representation of a pronounceable distractor delays the response to the target colour. As with the REH, this means that phonological representations whose phonetic codes are produced sooner can be excluded sooner. Hence high-frequency words can be excluded faster than low-frequency words and thus colour naming high-frequency distractors is faster

(Burt, 1994, 1999, 2002). Given this, the REH and the alternative account would predict that neutral words with complex onsets would have a useable production-ready representation before neutral words with simple onsets. This means that neutral words with complex onsets should be colour-named faster than neutral words with simple onsets. We did not observe this effect in the present study, which means that our results contrast with this aspect of the REH and the alternative account (assuming that the onset complexity effect observed here is the result of phonological encoding). Nevertheless, future research should aim to further investigate this effect perhaps in a study that uses many neutral words to avoid repetition of a small stimuli set.

Onset segment effects in manual Stroop tasks

The present results have implications for recent debates on the type of processing that happens with manual response Stroop tasks. It has been argued that the mode of response used to classify the colour of the printed word (e.g., responding manually via a keypress vs. responding vocally by saying the colour name aloud) determines the magnitude and type of facilitation and interference that results (Augustinova et al., 2019; Fennell & Ratcliff, 2019; Glaser & Glaser, 1989; Kinoshita et al., 2017; McClain, 1983; Redding & Gerjets, 1977; Sharma & McKenna, 1998; Turken & Swick, 1999; Zahedi et al., 2019; see also MacLeod, 1991; Parris et al., 2022). Accounts of differences between these two response modes propose differential access to the systems (i.e., phonological, lexico-semantic, or response-level processing) that are assumed to produce interference and facilitation (Glaser & Glaser, 1989; Kinoshita et al., 2017; Sharma & McKenna, 1998; Sugg & McDonald, 1994; Turken & Swick, 1999; Virzi & Egeth, 1985; see also Fennell & Ratcliff, 2019; Zahedi et al., 2019). Consistent with this, it has recently been argued that vocal and manual responding involve different tasks (naming vs. classification, respectively), and as such, the type of evidence that is accumulated during Stroop task performance is different, leading to qualitatively different Stroop effects and, importantly for present purposes, that phonological encoding does not occur with manual responses (Kinoshita et al., 2017; Kinoshita & Mills, 2020).

Onset segment effects on Stroop task performance were originally reported with a vocal response Stroop task. Coltheart et al. (1999) developed a novel Stroop paradigm by creating stimuli that either shared the initial phoneme or end phoneme of a colour name. For example, if the colour to be named was red, the to-be-ignored word would be *rat* (sharing initial phoneme) or *pod* (sharing the end phoneme) or a word that shares no phoneme at all (e.g., *fit*). Words that share a phoneme with a colour name have been shown to produce a naming latency advantage when naming the

colour not the printed word (Regan, 1978). Coltheart et al. (2001) reasoned that if, as assumed by the dual-route cascaded (DRC) model, there is a serial order component in processing of print-to-speech, there will be an advantage for colour names that share an *initial* phoneme with the to-be-ignored written word compared with items that share an *end* phoneme because the computation of phonology from print is left to. Consistent with their predictions, Coltheart and colleagues reported that phonemic overlap was significant at both positions. However, facilitation was greater for items with overlap in the initial phoneme position that they argued is incompatible with models that do not assume a serial grapheme-to-phoneme processing route in reading (see also Marmurek et al., 2006; Mousikou et al., 2015).² In contrast to the computation-of-phonology-from-orthography account of this *position-sensitive Stroop effect*, Kinoshita et al. (2017, 2020) have argued that it reflects phonological encoding (and hence that any left-to-right serial processing effects reflect phonological encoding and not phonological computation).

Kinoshita et al.'s arguments that phonological encoding does not happen with manual responses and that the left-to-right serial processing effects observed in Coltheart et al. (1999) and subsequent studies reflect phonological encoding, predicts that left-to-right serial processing effects should not be observed with manual responses. However, Parris et al. (2019) replicated Coltheart et al. (1999) with both vocal and manual response versions of the Stroop task, indicating that vocal and manual responses do not differ qualitatively, but only quantitatively (the effect of onset overlap was larger in the vocal response). The results from the present experiments are also therefore notable since they also indicate that phonological encoding occurs with the manual response Stroop task. According to Kinoshita and colleagues' position on the type of processing that can happen with a manual response, the onset complexity effect observed in the present study would be due to phonological computation and not phonological encoding.

In conclusion, the results from the present study show for the first time that onset complexity modifies Stroop task performance whereby congruent words (but not neutral or incongruent words) with complex onsets result in faster response times relative to congruent words with simple onsets, thereby increasing Stroop facilitation and congruency effects. This is consistent with the complexity advantage in intentional reading. Furthermore, in contrast to the alternative account of negative facilitation that proposes that negative facilitation results from difference in the pronounceability of irrelevant letter strings, onset complexity did not modify the magnitude of negative facilitation. However, two findings mean that we cannot conclusively reject the alternative account: (1) the negative facilitation effect was surprisingly small in our study reducing the likelihood of its modification and (2) the onset complexity effect took its effect by modifying

positive facilitation only, the very index of performance that appears to be generally prevented by mostly neutral (repeated-letter or neutral words) trial contexts. Without a concomitant effect on incongruent and neutral trials, it is not possible to determine whether negative facilitation is qualitatively different from informational conflict (the alternative account argues that negative facilitation is just another form of informational (phonological) conflict).

The results from this study have implications for understanding the levels of processing of the irrelevant word with a manual response and indicate the need for a future study comparing onset complexity effects in manual and vocal response Stroop tasks. Moreover, the results from this study are favourable to the task conflict account of negative facilitation, a theoretical approach that has provided insights into the nature of selective attention including the suggestion of a new potential form of impairment that might explain symptoms of clinical disorders, such as OCD (Kalanthoff et al., 2017).

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The writing was supported in part by ANR Grant ANR-19-CE28-0013 and RIN Tremplin Grant 19E00851 awarded to BP, LF and MA and by the British Academy Grant (SRG1819\190912) awarded to BP.

ORCID iDs

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

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

Maria Augustinova  <https://orcid.org/0000-0001-9514-4418>

Notes

1. We had also included a repeated-letter condition (i.e., 48 xxxx trials) in the present experiment to enable calculation of facilitation effects relative to this baseline, but an error in programming meant that the data for this condition were unusable.
2. See Marmurek et al. (2006) for a replication of this effect and for findings showing the same effect with non-words and an (albeit weaker) effect of onset overlap on Stroop interference, and Mousikou et al. (2015) for a replication with both pronounceable and non-pronounceable non-words.

References

- Augustinova, M., Clarys, D., Spatola, N., & Ferrand, L. (2018). Some further clarifications on age-related differences in Stroop interference. *Psychonomic Bulletin & Review*, 25, 767–774. <https://doi.org/10.3758/s13423-017-1427-0>

- 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*, Article 1786. <https://doi.org/10.3389/fpsyg.2019.01786>
- Berent, I., & Marom, M. (2005). Skeletal structure of printed words: Evidence from the Stroop task. *Journal of Experimental Psychology: Human Perception and Performance, 31*(2), 328–338. <https://doi.org/10.1037/0096-1523.31.2.328>
- 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), 85–99. <https://doi.org/10.1037/a0019252>
- Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*, 977–990. <https://doi.org/10.3758/BRM.41.4.977>
- Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition, 1*(1), 9. <https://doi.org/10.5334/joc.10>
- Burt, J. S. (1994). Identity primes produce facilitation in a colour naming task. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 47*(A), 957–1000. <https://doi.org/10.1080/14640749408401103>
- Burt, J. S. (1999). Associative priming in color naming: Interference and facilitation. *Memory and Cognition, 27*(3), 454–464. <https://doi.org/10.3758/BF03211540>
- Burt, J. S. (2002). Why do non-colour words interfere with colour naming? *Journal of Experimental Psychology: Human Perception and Performance, 28*(5), 1019–1038. <https://doi.org/10.1037/0096-1523.28.5.1019>
- 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. <https://doi.org/10.1037/0033-295X.97.3.332>
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204–256. <https://doi.org/10.1037/0033-295X.108.1.204>
- Coltheart, M., Woollams, A., Kinoshita, S., & Perry, C. (1999). A position-sensitive Stroop effect: Further evidence for a left-to-right component in print-to-speech conversion. *Psychonomic Bulletin & Review, 6*(3), 456–463. <https://doi.org/10.3758/BF03210835>
- Dalrymple-Alford, E. C. (1972). Associative facilitation and interference in the Stroop color-word task. *Perception & Psychophysics, 11*(4), 274–276. <https://doi.org/10.3758/BF03210377>
- Dalrymple-Alford, E. C., & Budayr, B. (1966). Examination of some aspects of the Stroop color-word test. *Perceptual and Motor Skills, 23*(3), 1211–1214. <https://doi.org/10.2466/pms.1966.23.3f.1211>
- Entel, O., & Tzelgov, J. (2018). Focusing on task conflict in the Stroop effect. *Psychological Research, 82*(2), 284–295. <https://doi.org/10.1007/s00426-016-0832-8>
- Entel, O., & Tzelgov, J. (2019). When working memory meets control in the Stroop effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(7), 1387–1406. <https://doi.org/10.1037/xlm0000790>
- Entel, O., Tzelgov, J., Bereby-Meyer, Y., & Shahar, N. (2015). Exploring relations between task conflict and informational conflict in the Stroop task. *Psychological Research, 79*, 913–927. <https://doi.org/10.1007/s00426-014-0630-0>
- Fennell, A., & Ratcliff, R. (2019). Does response modality influence conflict? Modelling vocal and manual response Stroop interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*, 2098–2119. <https://doi.org/10.1037/xlm0000689>
- Ferrand, L., Ducrot, S., Chausse, P., Maïonchi-Pino, N., O'Connor, R. J., Parris, B. A., Perret, P., Riggs, K. J., & Augustinova, M. (2020). Stroop interference is a composite phenomenon: Evidence from distinct developmental trajectories of its components. *Developmental Science, 23*(2), e12899. <https://doi.org/10.1111/desc.12899>
- Frederiksen, J. R., & Kroll, J. F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance, 2*, 361–379. <https://doi.org/10.1037/0096-1523.2.3.361>
- Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General, 118*(1), 13–42. <https://doi.org/10.1037/0096-3445.118.1.13>
- 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. <https://doi.org/10.1037/0096-1523.33.5.1170>
- 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*, Article 1157. <https://doi.org/10.3389/fpsyg.2014.01157>
- Hasshim, N., & Parris, B. A. (2021). The role of contingency and correlation in the Stroop task. *Quarterly Journal of Experimental Psychology, 74*(10), 1657–1668. <https://doi.org/10.1177/174702182110325>
- Kalanthroff, E., Avnit, A., Henik, A., Davelaar, E., & Usher, M. (2015). Stroop proactive control and task conflict are modulated by concurrent working memory load. *Psychonomic Bulletin and Review, 22*(3), 869–875. <https://doi.org/10.3758/s13423-014-0735-x>
- Kalanthroff, E., Davelaar, E., Henik, A., Goldfarb, L., & Usher, M. (2018). Task conflict and proactive control: A computational theory of the Stroop task. *Psychological Review, 125*(1), 59–82. <https://doi.org/10.1037/rev0000083>
- Kalanthroff, E., Goldfarb, L., Usher, M., & Henik, A. (2013). Stop interfering: Stroop task conflict independence from informational conflict and interference. *The Quarterly Journal of Experimental Psychology, 66*(7), 1356–1367. <https://doi.org/10.1080/17470218.2012.74160>
- Kalanthroff, E., Henik, A., Simpson, H. B., Todder, D., & Anholt, G. E. (2017). To do or not to do: Task control deficit in obsessive-compulsive disorder. *Behavior Therapy, 48* (5) 603–613. <https://doi.org/10.1016/j.beth.2017.01.004>
- Kawamoto, A. H., & Kello, C. T. (1999). Effect of onset cluster complexity in speeded naming: A test of rule-based approaches. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 361–375. <https://doi.org/10.1037/0096-1523.25.2.361>

- Kinoshita, S. (2000). The left-to-right nature of the masked onset priming effect in naming. *Psychonomic Bulletin & Review*, 7(1), 133–141. <https://doi.org/10.3758/bf03210732>
- Kinoshita, S., De Wit, B., & Norris, D. (2017). The magic of words reconsidered: Investigating the automaticity of reading color-neutral words in the Stroop task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(3), 369–384. <https://doi.org/10.1037/xlm0000311>
- Kinoshita, S., & Mills, L. (2020). Phonological encoding in the oral but not manual Stroop task: Evidence for the role of a speech production process. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(8), 1494–1504. <https://doi.org/10.1037/xlm0000828>
- Kinoshita, S., Mills, L., & Norris, D. (2018). The semantic Stroop effect is controlled by endogenous attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(11), 1730–1742. <https://doi.org/10.1037/xlm0000552>
- Klein, G. S. (1964). Semantic power measured through the interference of words with color-naming. *The American Journal of Psychology*, 77(4), 576–588. <https://doi.org/10.2307/1420768>
- Littman, R., Keha, E., & Kalanthroff, E. (2019). Task conflict and task control: A mini-review. *Frontiers in Psychology*, 10, Article 1598. <https://doi.org/10.3389/fpsyg.2019.01598>
- MacKay, D. G. (1987). *The organization of perception and action: A theory of action and other cognitive skills*. Springer-Verlag.
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109(2), 163–203. <https://doi.org/10.1037/0033-2909.109.2.163>
- 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. [https://doi.org/10.1016/S1364-6613\(00\)01530-8](https://doi.org/10.1016/S1364-6613(00)01530-8)
- Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(3), 503–535. <https://doi.org/10.1037/0278-7393.33.3.503>
- Marmurek, H. H. C., Proctor, C., & Javor, A. (2006). Stroop-like serial position effects in colour naming of words and non-words. *Experimental Psychology*, 53(2), 105–110. <https://doi.org/10.1027/1618-3169.53.2.105>
- McClain, L. (1983). Effects of response type and set size on Stroop color-word performance. *Perceptual & Motor Skills*, 56, 735–743. <https://doi.org/10.2466/pms.1983.56.3.735>
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134–140. [https://doi.org/10.1016/S1364-6613\(03\)00028-7](https://doi.org/10.1016/S1364-6613(03)00028-7)
- Monsell, S., Taylor, T. J., & Murphy, K. (2001). Naming the colour of a word: Is it responses or task sets that compete? *Memory & Cognition*, 29(1), 137–151. <https://doi.org/10.3758/BF03195748>
- Mousikou, P., Rastle, K., Besner, D., & Coltheart, M. (2015). The locus of serial processing in reading aloud: Orthography-to-phonology computation or speech planning? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(4), 1076–1099. <https://doi.org/10.1037/xlm0000090>
- O’Seaghdha, P. G., Dell, G. S., Peterson, R. R., & Juliano, C. (1992). Modeling form-related priming effects in comprehension and production. In R. Reilly & N. E. Sharkey (Eds.), *Connectionist approaches to language processing (Vol. 1)*, pp. 373–408. Lawrence Erlbaum.
- Parris, B. A., Hasshim, N., Ferrand, L., & Augustinova, M. (2023). Do task sets compete in the Stroop task and other selective attention paradigms? *Journal of Cognition*, 6, 23. <https://doi.org/10.5334/joc.272>
- Parris, B. A., Hasshim, N., Wadsley, M., Augustinova, M., & Ferrand, L. (2022). The loci of Stroop effects: A critical review of methods and evidence for levels of processing contributing to color-word Stroop effects and the implications for the loci of attentional selection. *Psychological Research*, 86(4), 1029–1053. <https://doi.org/10.1007/s00426-021-01554-x>
- Parris, B. A., Sharma, D., Weekes, B. S. H., Momenian, M., Augustinova, M., & Ferrand, L. (2019). Response modality and the Stroop task. *Experimental Psychology*, 66(5), 361–367. <https://doi.org/10.1027/1618-3169/a000459>
- Parris, B. A., Wadsley, M. G., Hasshim, N., Benattayallah, A., Augustinova, M., & Ferrand, L. (2019). An fMRI study of response and semantic conflict in the Stroop task. *Frontiers in Psychology*, 10, 2426. <https://doi.org/10.3389/fpsyg.2019.02426>
- Rastle, K., & Davis, M. H. (2002). On the complexities of measuring naming. *Journal of Experimental Psychology: Human Perception and Performance*, 28(2), 307–314. <https://doi.org/10.1037/0096-1523.28.2.307>
- Redding, G. M., & Gerjets, D. A. (1977). Stroop effects: Interference and facilitation with verbal and manual responses. *Perceptual & Motor Skills*, 45, 11–17. <https://doi.org/10.2466/pms.1977.45.1.11>
- Regan, J. (1978). Involuntary automatic processing in colour-naming tasks. *Perception & Psychophysics*, 24(2), 130–136. <https://doi.org/10.3758/BF03199539>
- Roelofs, A. (2003). Goal-referenced selection of verbal action: Modelling attentional control in the Stroop task. *Psychological Review*, 110(1), 88–125. <https://doi.org/10.1037/0033-295X.110.1.88>
- Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124(2), 207–231. <https://doi.org/10.1037/0096-3445.124.2.207>
- Rouder, J. N. (2014). Optional stopping: No problem for Bayesians. *Psychonomic Bulletin & Review*, 21(2), 301–308. <https://doi.org/10.3758/s13423-014-0595-4>
- Schiller, N. O. (2008). The masked onset priming effect in picture naming. *Cognition*, 106(2), 952–962. <https://doi.org/10.1016/j.cognition.2007.03.007>
- Schmidt, J. R., Liefooghe, B., & De Houwer, J. (2020). An episodic model of task switching effects: Erasing the homunculus from memory. *Journal of Cognition*, 3(1), Article 22. <https://doi.org/10.5334/joc.97>
- Sharma, D., & McKenna, F. P. (1998). Differential components of the manual and vocal Stroop tasks. *Memory & Cognition*, 26(5), 1033–1040. <https://doi.org/10.3758/BF03201181>

- Shichel, I., & Tzelgov, J. (2018). Modulation of conflicts in the Stroop effect. *Acta Psychologica, 189*, 93–102. <https://doi.org/10.1016/j.actpsy.2017.10.007>
- Spinelli, G., & Lupker, S. J. (2021). Proactive control in the Stroop task: A conflict-frequency manipulation free of item-specific, contingency-learning, and color-word correlation confounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47*(10), 1550–1562.
- Spinelli, G., & Lupker, S. J. (2022). Robust evidence for proactive conflict adaptation in the proportion-congruent paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 49*(5), 675–700. <https://doi.org/10.1037/xlm0001144>
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643–662. <https://doi.org/10.1037/0096-3445.121.1.15>
- Sugg, M. J., & McDonald, J. E. (1994). Time course of inhibition in color-response and word-response versions of the Stroop task. *Journal of Experimental Psychology: Human Perception and Performance, 20*(3), 647–675.
- Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory and Language, 26*, 608–631. [https://doi.org/10.1016/0749-596X\(87\)90105-7](https://doi.org/10.1016/0749-596X(87)90105-7)
- Turken, A. U., & Swick, D. (1999). Response selection in the human anterior cingulate cortex. *Nature Neuroscience, 2*, 920–924. <https://doi.org/10.1038/13224>
- Virzi, R. A., & Egeth, H. E. (1985). Toward a Translational Model of Stroop Interference. *Memory & Cognition, 13*(4), 304–319.
- Zahedi, A., Rahman, R. A., Stürmer, B., & Sommer, W. (2019). Common and specific loci of Stroop effects in vocal and manual tasks, revealed by event-related brain potentials and post-hypnotic suggestions. *Journal of Experimental Psychology: General, 148*(9), 1575–1594. <https://doi.org/10.1037/xge0000574>