HOW FAR IS A WRITTEN WORD WE ARE TRYING TO IGNORE PROCESSED?

A Thesis Submitted to the Committee of Graduate Studies

in Partial Fulfillment of the Requirements of the Degree of Master of Science

in the Faculty of Arts and Science

TRENT UNIVERSITY

Peterborough, Ontario, Canada

Psychology M.Sc. Graduate Program

May 2024

Abstract

How far is a written word we are trying to ignore processed?

Mariana Soutter

It is widely believed that basic mental processes involved in skilled reading are automatic in the sense that they occur without intention. Evidence that reading occurs without intention comes from the observation that the meaning of a colour word (e.g., “red”) affects the time to name the ink-colour of the word in the Stroop task. Evidence also suggests that non-colour words (e.g., house) interfere even though they are irrelevant to the colour naming task. The present study examined which reading processes are triggered without intention in the non-colour word Stroop task. One hundred and twenty skilled English readers completed both a reading aloud task and a colour naming task. In order to identify the reading processes triggered without intention, three psycholinguistic variables were examined, lexicality, word frequency, and neighbourhood density. The findings suggest that processing up to and including the activation of orthographic lexical representations occurs without intention and that intention is required to activate all subsequent reading processes.

Keywords: Stroop effect, Reading, Attention, Visual Word Recognition

Acknowledgements

I would like to thank Dr. Chan-Reynolds for your support and encouragement throughout my thesis. I would also like to thank Yuke, for your friendship throughout grad school and for keeping me grounded. Lastly, I would like to thank my family. I would not be where I am today without their influence and their constant cheering from the sidelines

Table Of Contents

List of Figures

List of Tables

Abbreviations

CB: Counterbalance

CWI: Colour-word interference

DRC Model: Dual-Routed Cascaded Model

IAC Model: Interactive Activation and Competition Model

N: Neighborhood Density

RT: Reaction Time

WF: Word Frequency

How Far is a Written Word we are Trying to Ignore Processed?

Overview

The ability to read is crucial for academic, economic, and social success (Jamieson, 2006; Norton & Wolf, 2012). Reading fluency is a significant predictor of school outcomes for all literacy-based subjects and individuals with greater literacy skills are more likely to pursue post-secondary education, such as university (Bigozzi et al., 2017). Higher literacy skills are also associated with greater employment levels, health, social status, and financial success (National Forum on Literacy and Poverty, n.d; Hanushek et al., 2013; Ivanova, 2011; Jamieson, 2006; Herbers et al., 2012; Almack, 2013). Reading fluency is positively related with stronger relationships and friendships, higher levels of social support and a better sense of community (Sparapani et al., 2018), and a higher overall standard of living (Underwood & Batt, 2003; National Forum on Literacy and Poverty, n.d.). Reading is thus essential for the success of an individual in modern society. Given this, it is important to understand how it is that people read.

Cognitive psychologists use an information-processing approach to understand reading. This approach conceptualizes reading as a collection of complex human behaviours (e.g., proofreading, studying, reading for pleasure, scanning a list, etc.), each of which is composed of a subset of cognitive processes involved in reading such as letter perception, visual word recognition, and the computation of phonology (e.g., Underwood & Batt, 2003). For instance, silent reading and reading aloud can be thought of as two different reading behaviours (i.e., tasks) that emphasize different subskills. Silent reading emphasizes semantics while deemphasizing phonology whereas reading aloud emphasizes phonology while deemphasizing semantic processes (Underwood & Batt, 2003).

Given the relationship between reading proficiency and quality of life, it should not be a surprise that a substantial amount of research examines how reading changes with practice. A major claim in reading research is that with extensive practice reading changes from a slow, controlled process to a fast, automatic process (for detailed discussions of the automatic view of word reading see Neely & Kahan, 2001; see also Augustinova & Ferrand, 2014; for a discussion of automatic vs. controlled processes see Besner & Boutilier, 1997; Flaudias & Llorca, 2014; Guttentag & Haith, 1978; Kinoshita et al., 2017; LaBerge & Samuels, 1974; Logan, 1980; Moors & De Houwer, 2006; Reynolds & Besner, 2006; Samuels & Flor, 1997; Tzelgov et al., 1997; Underwood & Batt, 1993). In their seminal work, Laberge and Samuels (1974) state that automatic processing is often a hallmark for complex, well-practiced behaviours such as reading: "As visual words are processed through many stages en route to meaningfulness, each stage is processed automatically." (p. 295). When visual word recognition becomes automatic in skilled readers, it frees up attentional resources, facilitating subsequent sentence-level comprehension (Macleod, 1991).

One task used to examine claims about the automaticity of reading in reading in skilled readers is the Stroop task (Stroop, 1935). In the Stroop task, participants are asked to name the ink-colour of a printed colour-word while ignoring the meaning of the word. Three conditions are often presented in an experiment, an incongruent colour condition (e.g., the word "green" printed in /red/ ink), a congruent colour condition (e.g., the word "green" printed in /green/ ink), and a neutral condition in which only the colour is displayed (e.g., a solid square printed in /green/ ink). The standard finding is that responses in the incongruent condition are slower than responses in the congruent and neutral conditions (e.g., see Macleod 1991). The observation that the word affects performance across a wide range of conditions despite instructions to ignore the word is taken as evidence that the word is processed all the way to the level of semantics (Neely & Kahan, 2001; Augustinova & Ferrand, 2014; Kinoshita & Mills, 2019) and phonology (Bakan & Alperson, 1967; Posner & Snyder, 1975; Perfetti & Zhang, 1995; Frost, 1998; Coltheart et al., 1999) automatically.

Consistent with the claim that words are read automatically in the Stroop task, the Stroop effect persists even when there is no advantage to reading the word (Zhang et al., 1999; Ferand & Grainger, 1992). Brown et al. (2002) describes this obligatory nature of reading succinctly, ‘‘Processing to the lexical level tends to occur whenever words are present in a display, regardless of participants' intention to read them, and often despite the intention to avoid reading them'' (p. 236). Though the Stroop effect is taken by the majority of cognitive psychologists to indicate that reading is automatic, studies demonstrating that the Stroop effect can be reduced or even eliminated merit the reconsideration of the nature of processing (i.e., controlled vs. automatic processing) involved in word reading (see Besner & Stolz, 1999; Bauer & Besner, 1997; Raz et al., 2002; Risko et al., 2005; O'Malley & Besner, 2002).

One criticism levelled against claims of automatic word reading based on performance in the standard Stroop task is that when colour-naming, the stimuli may be read because the meanings of the words are colour related and may therefore be seen as task relevant. Only a handful of studies have examined performance in the Stroop task using unrelated, non-colour associated words (e.g., non-color words such as "house"; Klein, 1964). In these studies, psycholinguistic properties of the words are manipulated (e.g., word frequency – how frequently a word has been encountered in print) and the effects on colour naming times are examined. Unfortunately, the data from these studies are inconsistent. Klein (1964) initially reported that the more frequently a word was encountered in print (i.e., the word frequency effect) and the more pronounceable the word was (i.e., the pronounceability effect), the more it interfered with colour naming. This finding was replicated by Fox and colleagues (1971) and Bakan and Alperson (1967). Subsequent studies that used a different stimulus presentation method have reported the opposite effect (Monsell et al., 2001; Burt, 2002). For example, in Monsell et al.'s (2001) conceptual replication of Klein's (1964) study, the pronounceability effect was replicated yet the word frequency and lexicality effects were not. A frequency effect in the opposite direction was observed in which there was greater interference for low-frequency words than high-frequency words, and whether the word was a real word or a pseudoword had little effect on colour-naming reaction times. Monsell et al.'s (2001) findings were later replicated (e.g., Burt, 2002).

Although the Stroop task with non-colour words has yielded mix findings, the task is interesting because the nature of the interference from non-colour-words is unlikely to be semantic or phonological. Consequently, it may provide insight to how reading processes themselves interfere with ongoing task performance (i.e., the colour-name response; Posner & Snyder, 1975; Monsell et al., 2001). Therefore, a Stroop variant that employs non-colour words can help us (1) better understand how irrelevant words are processed and (2) test the claim that word reading is automatic.

The current study examined the claim that word reading is automatic by asking participants to perform two tasks. One task was the non-colour word variant of the Stroop task. The other task, which served as the control, was a reading aloud task. The stimuli consisted of printed letter strings that form words (e.g., wheel) and pronounceable pseudowords (e.g., croul). The psycholinguistic properties of the printed letter strings were manipulated to replicate and extend previous work (e.g., Klein, 1964; Fox, 1971; Monsell et al., 2001). It was anticipated that the pattern of interference from these psycholinguistic properties would identify the locus (loci) of attentional effects in the reading system (i.e., the point at which interference is occurring).

Literature Review

Reading is a well-practiced and overlearned every-day behaviour that occurs with relative ease for skilled readers and that appears automatic as practice increases (LaBerge & Samuels, 1974; Watson, 2021). A classic claim in the reading cognition literature is that for skilled readers, reading up to, and including the activation of semantic representations is automatic in the sense that it is involuntary and unintentional (Posner & Snyder, 1975). In fact, skilled reading has a long history of being defined as one's ability to recognize words in an automatic way (Cattell, 1886; Brown et al., 2002; LaBerge & Samuels, 1974). Consistent with this, most theoretical accounts of visual word recognition either implicitly or explicitly include the assumption of automatic processing (e.g., Coltheart et al., 2001; Forster, 1976; McClelland & Rumelhart, 1981; Jacobs & Grainger, 1996; Morton, 1969; Plaut et al., 1996; Seidenberg & McClelland, 1989; Reynold & Besner, 2006). Much of the evidence for the claim that reading is automatic comes from the Stroop task.

Reading as an Automatic Process

Most cognitive psychologists operate under the view that reading up to, and including the activation of semantic representations, is automatic (Neely & Kahan, 2001). This assumption is widely discussed in cognitive textbooks, psychological books, and research articles (Stolz & Besner, 1999). The idea that skilled reading is automatic has been the received view for at least half a century (MacLeod, 1991). Laberge and Samuels' (1974) provided one of the first detailed accounts of this position. Reading automatically was thought to develop with practice, transforming from a controlled process to an automatic process where the word is read fluently

(Laberge & Samuels, 1974). In fact, one distinctive feature of a skilled reader is that, following years of practice, the associations between "visual symbols and speech sounds" becomes "effortless and automatic" (Joo et al., 2021, p. 1). Ventura and colleagues (2019) also state that the "processing of written words, a trademark of perceptual expertise, seems to involve fast and automatic activation of lexical phonological representations." (p. 5).

The prevalent view among contemporary reading researchers remains that reading aloud is an automatic process. In a recently published handbook on the science of reading, Brysbaert (2022) explained how "evidence is strong that in alphabetic writing systems, phonology is assembled automatically in the early stages of visual word recognition, and that various possible phonological codes are co-assembled which can in turn contribute to visual word recognition and text understanding." (p. 8). Another example comes from Joo and colleagues (2021) who state that "cognitive models of reading have proposed that, for the literate brain, viewing printed words produces widespread and automatic activation of phonological and semantic representations" (p. 2) and that "literacy involves automatizing the connections between orthographic (visual), phonological, and semantic codes in the brain" (p. 2). Similar claims are also made by Seidenberg and colleagues (Harm & Seidenberg, 1999; Harm & Seidenberg, 2004; Seidenberg & McClelland, 1989), among others (e.g., Van Orden & Goldinger, 1994; Grainger & Holcomb, 2010). For example, Grainger and Holcomb (2010) stated that "all three types of codes (orthography, phonology, and semantics) are automatically activated both in silent reading and reading aloud" (p. 3).

Automatic Processing

To investigate the claim that reading is automatic it is necessary to discuss what constitutes an automatic process. Automatic processing has long been a central concept in cognitive psychology (James, 1890; Wundt, 1903). Despite this, consensus has not been reached regarding the definition of automaticity, nor the key characteristics of an automatic process (see Moors & De Houwer, 2006; Reynolds & Besner, 2006). For instance, Posner and Snyder (1975; see also Neely & Kahan, 2001) defined automaticity as an unintentional and unconscious process that cannot be prevented, whereas Brown and colleagues (2002) state that for a process to be considered automatic it must be obligatory, autonomous, and without a limited capacity for processing demands. Bargh (1989, 1992, 1994) argued that the four most frequently occurring properties of automatic processes, which he termed “the four horsemen of automaticity”, are that it is unintentional, uncontrollable, efficient, and unconscious. On the other hand, Logan (1997) iterated that any definition of automaticity must include four similar, yet different properties; an automatic process must be fast, autonomous, effortless, and occur without conscious awareness. In fact, a review of the literature by Reynolds and Besner (2006) revealed that at least thirteen properties have been used to define a process as automatic. See Table 1 for an adapted list of properties that have been ascribed to automatic processes.

Table 1.

Characteristics of Automaticity.

Characteristic	Reference
1. Processing is stimulus driven.	Posner & Snyder, 1975; Hasher & Zacks,1979; Brown et al., 2002
2. Processing proceeds without intention.	Bargh, 1994; Boiché et al., 2016; Posner & Snyder, 1975; Cohen et al.,1992; Hasher & Zacks, 1979; Xu et al., 2014; Greenwald et al., 2009; Miles et al., 2019; Wu, 2013; Gold & Ciorciari, 2020
3. Processing proceeds without a requirement for central capacity.	Cohen et al.,1992; McCann et al., 2000; Pashler, 1994; Hasher & Zacks, 1979; Brown et al., 2002; Miles et al., 2019
4. Processing is autonomous and independent. It is not dependent on other processes. It runs the same way every time.	Posner & Snyder, 1975; Hasher & Zacks,1979; Logan, 1988; Brown et al., 2002; Kuhn et al., 2010; Logan, 1997
5. The process is executed obligatorily. It is ballistic in the way that once triggered by a stimulus it occurs from start to finish and cannot be stopped.	Hasher & Zacks, 1979; Besner, 2001; Brown et al., 2002
6. Processing proceeds without conscious control of processing	Atkinson & Shiffrin, 1968; Laberge & Samuels, 1974; McCann et al., 1992; McCann et al., 2000; Pashler, 1994; Posner, 1978; Schneider & Shiffrin, 1977; Stolz & McCann, 2000; Kuhn et al., 2010; Logan, 1997; Boiché et al., 2016; Greenwald et al., 2009; Dijksterhui, 2010; Miles et al., 2019; Happe, 2001
7. Processing proceeds without attention; it is independent of attention.	Bargh, 1994; Cohen et al.,1992; Shiffrin & Schneider, 1977; Laberge & Samuels, 1974; Logan, 1988; Happe, 2001; Pruessner et al., 2020; Gold & Ciorciari, 2020
8. Attentional demands are reduced in an automatic process.	Cohen et al.,1992; Bargh, 1994; Brown et al., 2002; Miles et al., 2019
9. Processing occurs in parallel.	Bargh, 1994; Treisman et al., 1992; James,1890; Shiffrin & Schneider, 1977; Schneider & Chein, 2003; Miles et al., 2019; Gold & Ciorciari, 2020
10. Processing speed is fast.	Neely, 1977; Logan,1988; Ehri, 2005; Boiché et al., 2016; Xu et al., 2014; Greenwald et al., 2009; Miles et al., 2019
11. The stimulus captures attention entirely.	Treisman et al.,1992

12. Processing cannot be altered by expectations.	Treisman et al.,1992; Neely, 1977; Miles et al., 2019
13. Processing cannot be altered with practice.	Hasher & Zacks, 1979; Toner & Moran, 2021

Note. Adapted/Reprinted from *Reading aloud is not automatic: Processing capacity is required to generate a phonological code from print* (p. 5), by M. G. Reynolds, and D. Besner, 2006, Journal of Experimental Psychology Human Perception & Performance. Adapted with permission.

More recently, theorists have argued that automaticity is not a single unified property of cognitive processing, but instead refers to a set of independent properties (Moors & De Houwer 2006; Reynolds & Besner, 2006; Stanovich, 1990). The complexity of defining automaticity is rooted in the vast number of possible definitions of automaticity as different combinations of features. For example, using Table 1 for reference, Posner and Snyder (1975) use properties 1, 2, and 4 as a definition automaticity, whereas the definition of automaticity used by Hasher and Zacks (1979) is comprised of six properties: 1-5 and 13.

The abundance of possible definitions of automaticity makes refuting claims of automatic processing virtually impossible. One solution would be to assess the presence or absence of all features. A process is automatic if all the features are true and not automatic if any feature is absent. However, there are two problems with this solution. Firstly, the all-or-none conception of the distinction between automatic and non-automatic processes is untenable in the face of studies demonstrating the lack of co-occurrence of features central to automatic processes (Moors & De Houwer, 2006). For example, Stroop studies have demonstrated that though Stroop interference meets some criteria of automaticity (e.g., it is unintentional) it does not meet other criteria (e.g., it is not independent of attention). In fact, it would be an extremely rare instance for a process to meet every criterion of automaticity. Secondly, it is impossible to assess all features simultaneously. A more viable solution would be to abandon the all-or-none distinction between automatic and non-automatic processes and take on a decompositional approach, in which researchers separate the features of automaticity and investigate them individually (Moors & De Houwer, 2006). This approach requires that researchers specify the feature or set of features they are referring to when assessing a process as automatic (e.g., referring to a process as being automatic in the sense that X is present, where X represents one of the above features of automaticity).

Intentionality

Moors and De Houwer (2006)'s conceptual analysis of automaticity identified lack of intention as one of the key facets of automatic processing. According to Moors and De Houwer (2006) a process does not require intention if its occurrence is goal independent. In short, an intentional process is one that is caused by a goal, whereas an unintentional process is not.

Consistent with Moors and De Houwer's claim that unintentionality is a core property of automatic processing, many definitions of automaticity include features that are related to the absence of intention (e.g., features 1, 2, 4, 5, 6, & 12 in Table 1). For example, Shiffrin and Schneider (1977) state that a process can be automatic in the sense that it has no requirements for intention. This unintentional process is said to be purely stimulus driven, unavoidable, and once triggered, difficult to suppress or alter (Moors & de Houwer, 2006). Likewise, Posner and Snyder (1975) and Neely and Kahan (2001) defined automaticity as an unintentional and unpreventable process, and Bargh (1989, 1992, 1994) also included unintentionality as one of the four "horsemen of automaticity". Given that multiple criteria for automaticity exist and the prevalence of intention (or lack thereof) in definitions of automaticity, the present study will test the claim that reading occurs automatically by examining which processes occur without intention.

The Stroop Paradigm

Much of the research supporting the claim that reading occurs automatically, in the sense that it occurs without intention, comes from studies using the Stroop task (Stroop, 1935; MacLeod, 1991; Moors & De Houwer, 2006; Augustinova & Ferrand, 2014). In the standard Stroop task, participants are instructed to name the ink-colour of an object (e.g., a word), while ignoring the meaning of the object. The colour-carrier is almost always a word (see MacLeod, 1991 for an extended discussion). Critically, the meaning of the word is manipulated to be incongruent with the target colour (e.g., the word "blue" in /red/ ink), congruent with the target colour (e.g., the word "red" in /red/ ink), or neutral with regards to the target colour (e.g., "XXXX" in red ink). The standard finding is that it takes longer to identify the ink-colour on incongruent trials than on neutral and congruent trials. Sometimes performance is faster on congruent trials than on neutral trials, though this is not always observed (see MacLeod, 1991 for a review). The slowed responses on incongruent trials relative to the other conditions is often referred to as the Stroop Effect. The Stroop task, and resultant Stroop Effect, is seen as a useful tool for studying unintentional processing because it *requires* engagement with a task (e.g., colour-naming) and processing of the relevant dimension (i.e., the colour), but does *not require* processing of the irrelevant dimension (i.e., the word).

Consistent with researchers believing that the Stroop task can be used to assess whether reading is automatic in the sense that it does not require intention, many researchers have claimed that word reading occurs without intention and cannot be stopped based on the persistence of the Stroop effect (e.g., Ashcraft, 1994; Brown et al., 2002, 1995; Cohen et al., 1990; Kahneman & Chajczyk, 1983; Logan, 1980; Posner & Snyder, 1975, 1985; Neely, 1977; Rayner & Pollatsek, 1989; Regan, 1978; Augustinova & Ferrand, 2014). For instance, Augustinova and Ferrand (2014) argue that Stroop studies have "made a significant contribution to the generally held view that word reading is automatic in skilled readers, in the sense that it cannot be prevented or controlled" (p. 343). It has been argued that the difficulty in naming the colour of a colour-word compared to a patch of colour is indicative that the irrelevant word cannot be voluntarily ignored even when doing so would be advantageous to colour-naming performance (Brown et al., 2002). For example, Brown et al. (2002) state that it is a "likely reality that word processing in the Stroop task is truly involuntary" (p. 537) and that the "reading of words cannot be inhibited via voluntary intention alone" (p. 537). Regan (1978) claims that the reading of the irrelevant word in the Stroop Task "will run to completion whether or not it is compatible with the person's intention" (p. 130). Posner and Snyder (1975) claim that "the Stroop phenomenon is a prime example of automatic processing without intention." (p. 392). Further, the Stroop effect has been taken as evidence that words automatically and unintentionally activate their meaning in a way that cannot be prevented (Ashcraft, 1994; Posner & Snyder, 1975; Neely, 1977). Bar-Anan et al. (2007) argue that in the Stroop task "word reading cannot be suppressed even when irrelevant to the task at hand given its automatic activation upon exposure to the stimulus." (p. 3) and Augustinova and Ferrand (2014) state "interference [in the Stroop and semantic Stroop tasks] is considered as evidence that readers cannot refrain from computing the lexical and semantic representations of the word that is read, even when such computations impair performance" (p. 827). In short, the Stroop task and ensuing research exploring the Stroop phenomenon have significantly contributed to the widely held view that reading is automatic (Augustinova & Ferrand, 2014; Besner & Stoltz, 1999).

One reason why researchers have argued that word reading occurs unintentionally based on evidence from the Stroop task is because of the persistence of the Stroop effect across a wide range of conditions. If a process is unintentional, and therefore goal-independent, the effects of the process should be observed despite substantial changes in methodology. In what follows I provide a small sampling of evidence that the Stroop effect persists across a wide range of methodological changes even when steps are taken to discourage the reading of the word. I begin by discussing the original card version of the Stroop task and then switch to the individual stimulus presentation version of the Stroop task. I then discuss the spatial separation of the colour and the word, and finally the use of congruency manipulations, post-hypnotic suggestion, and single-letter colouring and cueing manipulations in an attempt to eliminate the Stroop effect. In Stroop's (1935) original study, participants named the ink-colours of objects printed on sheets of paper. In the experimental condition (now referred to as the incongruent condition), the stimuli consisted of two sheets of paper, each with a list of 100 colour words (e.g., red, blue, green, brown, and purple) printed in incongruent ink colours (e.g., the word "red" printed in /blue/ ink). In the control condition (now referred to as the neutral condition), the sheets consisted of 100 squares of colours (e.g., ■ printed in /red/ ink), with the ink colours presented in the same sequence as on the word sheets[1]. Participants were instructed to serially name the ink-colours of the items on the sheets. Critically, it took participants 47 seconds longer to name the ink-colours on the sheet of colour words compared to the sheet of patches of colour. Stroop's original demonstration is compelling evidence that the words were read unintentionally because

[1] A common misconception is that the original Stroop experiment included a congruent condition in which colour words were printed in their congruent colours (e.g., the word blue printed in /blue/ ink). For example, De Houwer (2003, see also Bugg & Jacoby, 2008; Roloefs, 2005; Rothermund et al., 2022) claim that the Stroop effect refers to the interference observed in colour-naming performance on incongruent trials compared to congruent trials. However, the use of congruent trials in the Stroop task (e.g., Sichel & Chandler, 1969; Glaser & Dolt, 1977; Glaser & Glaser, 1982; Redding & Gerjets, 1977; Zajano & Gorman, 1986) wasn't introduced until several decades after the publication of Stroop's (1935) original study when individual stimulus presentation was used .

the meaning of the word affected colour naming performance despite only hurting task performance.

One of the largest methodological changes to the Stroop task was brought about by the tachistoscope, which allowed for individually presented stimuli to be displayed in a sequential manner and for timing to be recorded with millisecond accuracy (MacLeod, 1991; Sahinoglu & Dogan, 2016). The Stroop task with individual stimulus presentation was first used by Dalrymple-Alford and Budayr (1966, see also Tecce & Dimartino, 1965). The use of individually presented stimuli made it possible for elements of the task and stimuli to be controlled and manipulated in novel ways (MacLeod, 1991). For instance, individual stimulus presentation allowed for congruent trials to be included in experimental designs[2]. With the introduction of computers, researchers began to study how performance was affected by other factors such as response modality (e.g., vocal vs. key-press responses) and the intervals between stimulus presentation (Sahinoglu & Dogan, 2016). Importantly, a significant interference effect comparable to the original Stroop estimate was observed despite presenting the incongruent colour-word stimuli individually (Dalrymple-Alford & Budayr, 1966). Further, a Stroop effect has been consistently demonstrated with individual stimulus presentation (e.g., MacLeod, 1991; Sichel & Chandler, 1969).

One of the first manipulations to test the claim that the word is read unintentionally, was to physically separate the ink-colour and the word. In most Stroop experiments, the word is the colour-carrier. It is therefore possible that the word is processed because it is (intentionally) selected along with the ink-colour as the two dimensions (i.e., the word and the colour) are

---

[2] If congruent items were included on the paper version of the task, then it would be impossible to determine whether performance was due to people reading the words or naming the colour on the card (both responses would be identical for the entire card).

integrated into a single stimulus. Dyer (1973) was the first to examine whether Stroop interference was reliant on stimulus integration. Dyer (1973) separated the colour and word dimensions of the stimulus by presenting a word and a solid colour square laterally on either side of a fixation point. Despite separating the two dimensions, Dyer (1973) reported slower responses on incongruent trials compared to congruent trials. This suggests that the word is processed unintentionally and not (only) because it is selected as part of the integrated stimulus.

One issue with Dyer's (1973) experiment was that the colour-carrier and colour-word switched locations on a trial-by-trial basis. With the target location uncertain, the colour-word could have been processed on a subset of trials because attention was intentionally allocated to a location of where the colour word could appear. Kahneman and Chajczyk (1983) manipulated stimulus integration by separating the dimensions vertically, rather than laterally as in Dyer (1973). In this study, a coloured square was always presented at fixation and the colour-word was presented either above or below the colour-carrier. If the word was only read in Dyer's (1973) study because the colour-word and colour-carrier could appear at the same locations, then no Stroop effect should be observed when (1) the two dimensions are separated, (2) the location of the colour-carrier is known, and (3) the colour-carrier was never presented at a location where words were presented. Inconsistent with this prediction, a robust Stroop effect was still observed, although it was noticeably reduced in magnitude. Again, this supports the claim that the word is processed without intention.

One problem with the study by Kahneman and Chajczyk (1983) is that they included congruent trials. It was therefore possible that participants were motivated to read the word, because there was a benefit to doing so on a subset of trials. Brown et al. (2002) attempted to remove any incentives to read the colour-word. In their study the colour-carrier (solid rectangles of colour) was once again separated from the colour-word. However, unlike previous work by Dyer (1973) and Kahneman and Chajczyk (1983), they excluded congruent trials so that there would be no benefit to reading the colour-word. They also used an abrupt onset spatial cue to direct attention to the spatial location of the colour-carrier. Consistent with the words being read without intention, a reliable Stroop effect was observed. Brown et al. (2002) argued that this constituted evidence that the reading of words is involuntary and cannot be prevented through intention alone.

Post-hypnotic suggestion has also been used to test whether the word is read automatically in the sense that it occurs without intention. For the most part, these studies have been able to reduce, but not eliminate the Stroop effect, consistent with the idea that word reading is occurring without intention (e.g., Raz et al, 2006; Augustinova & Ferrand, 2012). However, two studies were able to successfully eliminate the Stroop effect when using highly suggestible participants (Raz et al., 2002; 2003). For instance, Raz et al. (2002) investigated whether a posthypnotic suggestion to view word stimuli as meaningless symbols of a foreign language could inhibit lexical processing in the Stroop task. Participants were presented with a word blindness suggestion, which is a posthypnotic suggestion. With this manipulation, Stroop interference was successfully eliminated. This suggests that word reading is not automatic in the sense that it is unintentional because it can be affected by goals instilled by the researcher.

One issue with Raz et al.'s (2002) study is that the word blindness suggestion to view words as meaningless symbols may not affect whether the words are unintentionally read to the point of activating semantics. Instead, it may eliminate the Stroop effect by affecting non-semantic processes such as response competition. Augustinova and Ferrand (2012) examined whether posthypnotic suggestion affected the semantic contribution to the Stroop effect by having highly suggestible participants complete a Stroop task using incongruent trials with colour words (e.g., the word “green” in /blue/ ink) and colour-associated words (e.g., the word “grass” in /blue/ ink). They found that posthypnotic suggestion reduced (but did not eliminate) interference from the colour-words. However, posthypnotic suggestion had no effect on the magnitude of interference from the colour-associated words. Augustinova and Ferrand (2012) therefore argued that the persistence of a semantically based Stroop effect in the face of posthypnotic suggestion indicates (1) that semantic activation, and thus reading, is not inhibited (or influenced) by suggestion, and (2) that suggestion in the standard Stroop task influences non-semantic, task-relevant response competition. Therefore, this study suggest that hypnotic suggestion does not prevent participants from unintentionally reading the carrier word, and that reading is automatic in the sense that it cannot be voluntarily prevented or affected by goals instilled by the researcher.

Other studies have examined whether word reading occurs without intention in the Stroop task by colouring and/or cueing a single letter of the colour word (e.g., the word “blue” where only the /b/ is printed in green ink). There are numerous demonstrations that this manipulation reduces the size of the Stroop effect (e.g., Kahneman & Henik, 1981, Experiment 3; Besner et al., 1997, Experiment 1; Besner & Stoltz, 1999; Monahan, 2001; Reynolds et al., 2010) and, in some instances, has been shown to eliminate the Stroop effect entirely (e.g., Besner et al., 1997, Experiment 2; Besner & Stoltz, 1999; 2001). In their first experiment, Besner et al. (1997) demonstrated that the single letter colouring manipulation significantly decreased the Stroop effect when manual responses were made (e.g., by indicating the target colour using a button keypress). However, these demonstrations included congruent trials, which gave participants some incentive to read the irrelevant word. In their second experiment, the single letter colouring manipulation was able to eliminate the Stroop effect when the congruent trials were replaced with neutral trials (e.g., nonword stimuli that begin with the same letters as the colour words in the response set, such as "ret" or "blat"). The elimination of the Stroop effect suggests that semantic activation can be prevented. This finding is inconsistent with word reading occurring automatically in the sense that it proceeds without intention.

Subsequent work has been critical of Besner et al.'s (1997) demonstration that the Stroop effect can be eliminated. These criticisms, summarized by Augustinova and Ferrand (2012), include the use of a "nonstandard" manual mode of response rather than a vocal colour-naming response, the absence of a direct semantic manipulation, and the use of neutral stimuli beginning with the same letters as the colour words in the response set which may have primed the phonology of the colour words. Besner and colleagues addressed some of the criticisms in a study reported by Manwell et al. (2004). This paper reports a Stroop experiment using a single-letter colouring manipulation using a vocal response mode. It also included neutral words that do not have any phonological overlap with the colour words in the response set (e.g., "jail") and colour-associated words (e.g., "sky") as a semantic manipulation in addition to the standard Stroop colour words (i.e., "blue", "yellow", "green", and "red"). With these conditions the standard Stroop effect (incongruent – neutral) was substantially reduced and the semantically based Stroop effect (semantic associates – neutral) was eliminated. Manwell et al. (2004) argued that the elimination of the semantic Stroop effect is consistent with semantic activation being blocked by the single letter colouring and cueing manipulation.

Augustinova et al. (2010) point out that in Manwell et al.'s (2004) study the sample size (e.g., 16 participants) was too small to have enough statistical power to detect an effect, thus their failure to obtain a semantically based Stroop effect is the result of a Type II error.

Augustinova et al. (2010) replicated Manwell et al.'s (2004) study using a larger sample size. In their first experiment, the colour-carrier letter varied in position on each trial (e.g., first, middle, or final letter) as in Manwell et al.'s (2004) original experiment, and in their second experiment the colour-carrier letter was always the first letter. Replicating previous work (e.g., Manwell et al., 2004; Besner & Stoltz, 1999), colouring a single letter in the word significantly reduced the standard Stroop effect in both experiments. However, inconsistent with Manwell et al. (2004), colouring a single letter of the word failed to eliminate or even reduce the semantically based Stroop effect. The semantic Stroop effect was present regardless of whether the coloured and cued letter was always the first letter or if it was varied each trial (e.g., first letter, middle letter, or final letter). Augustinova et al. (2010) argue that these results indicate that semantic activation in the Stroop task cannot be eliminated, and that reading can therefore be considered automatic in the sense that it occurs unintentionally, where semantic activation cannot be inhibited.

One explanation for the persistence of the Stroop effect is that the colour-word is perceived as relevant. De Houwer (2003) discusses the many ways in which the colour-word is related to the ink-colour in the Stroop task beyond being presented in close spatial and temporal proximity. For instance, the colour-word is semantically and phonologically related to the ink-colour, and the colour-words belong to the set of valid responses for the task. Given the overlap between the properties of the two stimuli, it is not unreasonable to assume that participants will see them as connected. Consistent with this account, De Hower (2003) demonstrated that the Stroop effect can be decomposed into two types of conflict, stimulus-response and stimulus-stimulus conflict (De Houwer, 2003). Stimulus-response conflict results from conflict at the level of the response. On congruent trials, the irrelevant word is congruent with both the ink colour and the response but on incongruent trials, the irrelevant response is incompatible with both the ink colour and the response. The difference in stimulus-response compatibility affects the response stage of selection in which incongruent trials interfere with response selection because the irrelevant word (e.g., “blue” printed in/red/ ink) activates the associated response (e.g., “blue”) which interferes with and competes with the correct response (e.g., “red”). Stimulus-stimulus conflict on the other hand, refers to conflict at the lexico-semantic level in which differences in performance between congruent and incongruent trials results from the semantic representations of the irrelevant word being activated. For example, in incongruent trials the irrelevant word (e.g., “blue” printed in/red/ ink) activates the semantic representation of “blue” which creates conflict in activating the semantic representation of “red”.

An important implication from De Houwer ‘s (2003) study is that the words presented in the Stroop task may be read if they are perceived as related to ink-colour. If so, then the standard colour-word Stroop task may not be appropriate for testing the claim that words are read automatically in the sense that word reading occurs independent of intention. Instead, claims about automaticity may require the use of non-colour words. I therefore turn to the non-colour word variant of the Stroop task next.

The Non-Colour Word Variant of the Stroop Task

In the non-colour-word variant of the Stroop task, the primary task is colour naming (as in the Stroop task), but the distractor words are not colour words or colour related words. The first to use this variant of the Stroop Task was Klein (1964). Klein was also the first to manipulate the characteristics of the word stimuli used to carry the ink-colour. Seven different types (i.e., conditions) of stimuli were used (1) colour words in the response set (e.g., “blue”), (2) colour-words not in the response set (e.g., “purple”), (3) colour-associated words (e.g., “sky”),

(4) common, non-colour associated words (e.g., “take”), (5) rare non-colour associated words (e.g. “sol”), (6) unpronounceable nonword consonant strings (e.g.,hjhg), and (7) ink-colour only.

In Klein’s original work, stimuli were written on cards like in the original Stroop (1935) study. The type of word was therefore blocked (one word type or condition per card) and run between subjects. Participants appeared in only one word-type condition. Items on sheets were printed in red, yellow, blue, and green ink. No ink-colour appeared twice in a row. Response times and accuracy were measured manually for each card. Interference was indexed by subtracting the time to complete an experimental card (e.g., common words) from the time to complete a control card (e.g., ink-colour only).

Klein (1964) found that interference decreased in the following order: colour-words in the response set, colour words not in the response set, colour-associated words, common (hereafter, high-frequency) non-colour associated words, rare (hereafter, low frequency) non-colour associated words, unpronounceable nonwords. Klein argued that the pattern of interference could be explained by a semantic gradient in which colour-word interference was largest for words with the greatest semantic association to the ink colours being named. The finding that high-frequency non-colour associated words interfered more than low-frequency non-colour associated words, which in turn interfered more than unpronounceable letter strings, was thought to be a result of pronounceability. Kleins findings were later replicated by Bakan and Alperson (1976) and Fox and colleagues (1971) who also used the card version of the task.

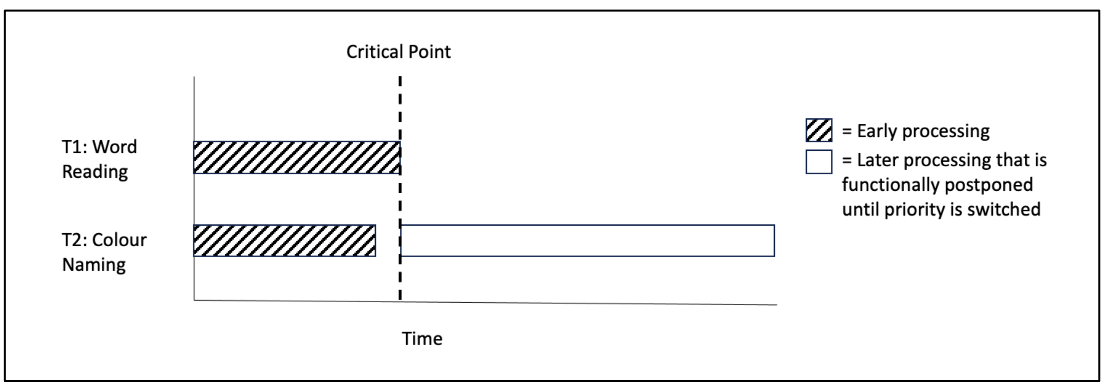
Although Klein’s (1964) findings were replicated by other laboratories using the card version of the task, the pattern of results have been quite different for studies that used individual stimulus presentation (Burt, 1994; 1999; 2001; Kinoshita et al., 2017; Monsell et al., 2001). For instance, Klein (1964; Bakan & Alperson, 1976; Fox et al., 1971) found *slower* colour naming reaction times for high frequency words compared to low frequency words, whereas studies using individual stimulus presentation found *faster* colour naming reaction times for high frequency words than for the low frequency words (Burt, 1994; 1999; 2001; Monsell et al., 2001; see also Navarrete et al., 2015). Similarly, in the card version of the task, responses were *slower* for low-frequency words than for nonwords, whereas the opposite appears to be true in the individual stimulus presentation version of the task (Burt, 2002; Monsell et al., 2001; Kinoshita et al., 2017). It is not known why the two versions of the task produce different patterns, but it is hypothesized that the card version allows participants to alter their response criteria based on the condition, whereas a single criterion is adopted for all conditions in the individual stimulus presentation version.

The observation that non-colour words affect colour-naming performance despite being irrelevant to the task (not part of the response set, semantic set, phonological set, orthographic set and not semantically or associatively related to colour) is consistent with the words being processed in the absence of intention. What is unclear at present is which word reading processes are occurring without intention. For instance, Klein argued that the effects of word frequency were due to phonological interference, suggesting that processing up to phonology occurs without intention. Here, a simple dual-task explanation is proposed to explain the non-colour word effects on colour naming performance (Reynolds et al., n.d). This account will be used to organize the remainder of the thesis.

According to the dual task account, the presentation of the of the stimulus triggers two tasks (1) word reading and (2) colour naming. The processing of the irrelevant non-colour word is *unintentionally* triggered by the presentation of a word-like stimulus; the presentation of a word triggers a word reading task. Processing of the word (the word reading task) is given priority over processing the colour (the colour naming task). Some processing associated with stimuli in both tasks may proceed in parallel, but some portion of the processing associated with the colour naming task is functionally postponed until processing can be disengaged from the word reading task. Word reading processes proceed in a ballistic fashion (once started, they cannot stop) until some critical point in the reading process is reached. At this point, processing associated word reading task is deprioritized and emphasis is given to processing associated with the colour naming task. This account predicts that colour naming performance will be affected by all variables that affect the speed of processing in the reading task prior to the critical point where priority is switched to the colour naming task.

Figure 1.

Dual Task account of the non-colour word Stroop task as a function of task and processing time course.



Note. Figure 1 demonstrates the time course for the processing of the word dimension of the stimulus and for the colour dimension of the stimulus in the non-colour word Stroop task. The presentation of the stimulus triggers two tasks (1) word reading and (2) colour naming, and priority is given to the word reading task. Some early processing occurs in parallel, but a portion of the processing associated with the colour naming task is functionally postponed until priority is switched from the word reading task to the colour naming task.

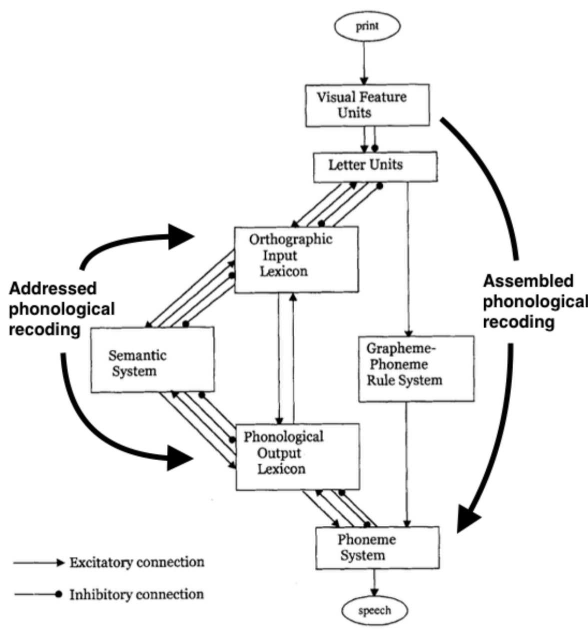
The Dual-Routed Cascaded Model of Reading Aloud.

The Dual-Route Cascaded (DRC) model of reading is one of the most successful computational models of reading aloud (see Figure 1). Critically, it can simulate the effects of all the variables used in non-colour word Stroop studies in the context of reading aloud (Coltheart et al., 2001). It can therefore be used as a framework for discussing how a pronunciation is generated for word-like stimuli and for identifying where in the reading system a variable has its effects. Here I will use the DRC model to discuss the findings from studies using the individual stimulus presentation version of the non-colour word Stroop task (Burt, 1994; 1999; 2002; Kinoshita et al., 2017; Monsell et al., 2001). I will also use the DRC model to illustrate how the dual task account explains performance in the non-colour word Stroop task.

The DRC model is a dual route model because it has two routes or pathways for generating a pronunciation; a lexical route (sometimes called addressed phonology) which uses whole word knowledge and a non-lexical route, which uses spelling-to-sound correspondences (sometimes called assembled phonology or sublexical route). The lexical route consists of the orthographic input lexicon and phonological output lexicon. It receives input from and passes its output to both the letter units and phoneme buffer. It can generate a correct pronunciation for every word known to the model. It is required to read words that have irregular spelling-sound correspondences (e.g., “the”; “pint”). The sublexical route consists of the Grapheme-to-Phoneme Conversion (GPC) system. It takes its input from the letter units and passes its output to the Phoneme Buffer. This system converts orthography (graphemes) into phonology (phonemes) serially, letter by letter, left to right, across the letters using a set of rules based on the most common grapheme to phoneme correspondences. It can generate a correct pronunciation for all words (or word-like stimuli) whose pronunciation follows these rules. The sublexical route is required to read words that are new to the reader and are therefore not stored in memory (e.g., “blart”). The model is cascaded because the model does not wait until processing is complete at one level of representation before passing information onto the next level. Thus, information (or activation in this case) is seen as cascading through the model.

Figure 2.

The structural architecture for Coltheart et al.’s (2001) Dual Routed Cascaded (DRC) model of reading aloud.



Note. Reprinted from *DRC: a dual route cascaded model of visual word recognition and reading aloud* by M. Coltheart, K. Rastle, C. Perry, R. Langdon, and J. Ziegler (2001).

The model is composed of multiple layers of representations that often interact with each other. The visual feature units are used to encode the visual representation of the stimulus. The features of each letter are encoded as the presence and absence of line features based on the Interactive Activation model of Rumelhart and McClelland (1981). Once the features are loaded into the visual feature units activation is passed to the letter units. The letter units consist of a representation of every possible letter in the English language at each letter position. As activation rises in the letter units their activation is passed on, in parallel, to the orthographic Lexicon. The orthographic lexicon contains an entry for each uniquely spelled word. Next is the Phonological Lexicon which contains an entry for each uniquely pronounced word (based on phonemes). Finally, activation is passed onto the phoneme system, which like the letter units, contains a unique representation of each phoneme for each position.

Repetition Priming Effects

Burt (1994; 1999; 2002) examined whether performance in the non-colour word Stroop task is affected by repetition priming. The repetition priming effect refers to the observation that a repeated word is read aloud faster than a non-repeated word in an experimental session. Repetition priming effects are believed to arise, in large part, at visual levels of processing prior to the orthographic lexicon (i.e., at the feature and letter level). In a non-colour word Stroop task, Burt (1994) reported that the colour naming of a target word is facilitated when there is a long interval between prime and target onsets (e.g., a stimulus onset asynchrony (SOA) of 5000 ms or more). Colour-naming of the target word (e.g., the neutral, non-colour associated word "house" printed in /blue/ ink) was faster if the participant silently read the identity prime word (e.g., "house") than if they read an unrelated prime word (e.g., "hat"), before colour-naming the target. Facilitation in colour-naming was also produced when participants were instructed to remember the prime word and recall it after each colour-naming response. Using the same long-SOA repetition priming design, Burt (2002, Experiment 1) also reported facilitation in colour naming when identity primes were used, compared to unrelated prime words. These repetition priming effects suggests that word processing up to and including the orthographic lexicon is occurring prior to priority being switched to the colour-naming task (Burt, 1994, 1999, 2002).

Word Frequency Effects

One of the most common variables that has been examined in the non-colour word Stroop task is word frequency. The word frequency effect refers to the observation that words encountered frequently in text (e.g., cat) are read aloud faster than words that are encountered less frequently (e.g., gnu). The word frequency effect is believed to arise from processing in the orthographic and phonological lexicons (Coltheart et al., 2001; McCann & Besner, 1987; McCann et al., 1988; Morton, 1980; Plourde & Besner, 1997). In the DRC model (Coltheart et al., 2001), this arises because the rate of activation for lexical representations in the orthographic input lexicon and phonological output lexicon is directly related to how often the word is encountered (see McClelland & Rumelhart, 1981; McClelland, 1993; Seidenberg & McClelland, 1989; Plaut et al., 1996; Harm & Seidenberg, 1999 for similar explanations).

The evidence that word frequency affects colour naming performance is complicated. Klein (1964) initially reported greater interference in colour-naming for high-frequency words than for low-frequency words. Monsell et al. (2001) consistently reported faster colour-naming reaction times for high-frequency items than for low-frequency items. Burt reported either no difference in colour-naming performance between high-frequency and low-frequency words (Burt, 1994 Experiment 1 and 2), a trend for faster colour naming for high-frequency words compared to low-frequency words (Burt, 1994 Experiment 3), or significantly faster naming for high-frequency than for low frequency words (Burt, 1994, Experiment 4; 1999, Experiment 1; 2002, Experiment 1 and 2). The presence of word frequency effects in the colour naming task would suggest that lexical processing of the irrelevant word is occurring up to and including the orthographic and phonological lexicons before priority is switched to the colour-naming task.

Lexicality Effects

Another commonly examined variable in the non-colour word Stroop task is lexicality. Lexicality refers to whether the word-like stimulus is a known word (e.g., cat) or a word-like stimulus that is not a known word (e.g., gat). The lexicality effect in reading aloud, which refers to the observation that words are read aloud faster than nonwords, is suggested to arise at a late stage of processing (Coltheart et al., 1977). In the DRC model (Coltheart et al., 2001), the lexicality effect in reading aloud is believed to arise at the level of the phoneme system. Evidence in the phoneme system is accumulated quickly for words because the phonological representation of a word is retrieved in parallel directly from memory using the lexical route. In contrast the phonological representation of a nonword must be computed serially using the non-lexical route; a process that takes more time than the direct memory retrieval process. This account is consistent with other theories (Coltheart et al., 1977; Marshall & Newcombe, 1973; Morton & Patterson, 1980; McClelland & Rumelhart, 1981; McClelland, 1993).

Klein (1964) initially reported greater interference for common words compared to unpronounceable consonant strings, an effect that was replicated by Bakan and Alperson (1967) and Fox et al. (1971). Monsell et al. (2001) however, reported no difference in colour-naming performance for words and pronounceable pseudowords in any of their experiments. Kinoshita et al. (2017) also reported no significant effects of lexicality on colour-naming latencies. In contrast, Burt (2002, Experiment 5) reported greater interference for pseudowords than for words. Burt’s (2002) data is inconsistent with the null effect of lexicality reported by both Monsell et al. (2001) and Kinoshita et al. (2017). However, Burt’s experiment included a repetition manipulation which interacted with familiarity and lexicality. When examining only unrepeated items a null effect of lexicality was reported, likewise to Monsell et al. (2001) and Kinoshita et al. (2017).

Methodological Differences in the Non-Colour Word Stroop Task

It is unclear how to interpret the inconsistencies in the word frequency and lexicality effects in non-colour word variants of the Stroop task due to the methodological inconsistencies across studies. One issue, namely, the use of the blocked card (e.g., Klein, 1964; Bakan & Alperson, 1967; Fox et al., 1971) or individual stimulus presentation (e.g., Burt, 1999; 2002; 2004; Kinoshita et al., 2017; Monsell et al., 2001) could have contributed to the inconsistencies in the word frequency and lexicality effects. In earlier studies (e.g., Klein, 1964; Bakan & Alperson, 1967; Fox et al., 1971), stimuli were presented in a list format in which blocked stimulus presentation was used (e.g., the different stimuli conditions/types were presented on different cards). This form of stimulus presentation allowed for participants to adjust their response criterion as a function of stimulus type, which could have influenced the word frequency and lexicality effects. The list format of stimulus presentation also permitted for the processing of an earlier item to overlap with that of a later item because they were both visible simultaneously. Further, latencies were indexed by measuring the total time it took a participant to read (or colour-name) each cumulative list of items – the latency per item was thus not a precise measure but an estimate based on an average.

Burt (1994;1999; 2002), Monsell et al. (2001), Kinoshita et al. (2017), oppositely to Klein (1964), used trial-by-trial, individual stimulus presentation. This would have increased the likelihood of detecting true effects from different conditions (e.g., word-frequency and lexicality effects). This allowed for response times for individual items to be measured rather than a list of items, increasing accuracy and decreasing the probability of a type II statistical error being made, in which a failure to detect a true effect occurs.

Another difference across studies is whether a reading aloud control condition was included. Whereas some studies (e.g., Monsell et al., 2001) included word reading performance data, other studies (Kinoshita et al., 2017) did not. For example, Kinoshita (2017) only reported colour-naming performance data. Though Kinoshita et al. (2017) addressed several methodological issues in Monsell et al. (2001), they failed to include a reading aloud task as a control. It is therefore difficult to determine if their null lexicality effect in the colour-naming data is truly an absence of difference in response times between words and pseudowords, or if their stimuli are ill-controlled such that no effect would be observed if they had performed a reading aloud task.

Another methodological issue that could have contributed to the inconsistencies in the word frequency and lexicality effects is with regards to stimulus construction. Previous work differed in terms of the type of rare or low-frequency words that were used, whether pronounceability was controlled for, in the number of stimulus items, and in whether any of the items were repeated. Differences between studies were also observed in whether phonological overlap between the stimuli items and the colours were controlled for and whether semantic associations between stimuli words and the colours were accounted for. Inconsistencies in the word frequency and lexicality effects in the literature could also be partially attributed to the matching of items and items sets.

Monsell et al. (2001) argued that many of the rare words used in Klein's (1964) study were so rare they were akin to nonwords. This could have increased the possibility of detecting a familiarity or word frequency effect and decreased the likelihood of detecting a lexicality effect.

Even in later studies (e.g., Monsell et al., 2001), unpronounceable heterogenous letter strings (e.g., lwbq) and homogenous letter strings (e.g., xxxx) were used, making it difficult to decipher between a lexicality effect and a pronounceability effect.

Pronounceability effects were not well controlled in earlier studies, possibly confounding the effect of lexicality. To illustrate, in Klein's (1964) the nonword stimuli used were unpronounceable nonsense syllables making it difficult to decipher between lexicality and pronounceability effects. Whereas later studies (e.g., Monsell et al., 2001; Burt 2002, experiment 5) obtained a lexicality effect in which colour-naming latencies were greater for nonwords than words, Klein (1964) obtained a lexicality effect in the opposite direction in which colour-naming latencies were greater for words than nonwords. Klein used unpronounceable nonword letter-strings, whereas the nonwords in later studies were pronounceable and orthographically acceptable in English.

Inconsistencies in the lexicality effect could also be attributed to poorly constructed nonword stimuli. For example, in Monsell et al. (2001) accuracy rates were only reported for words since they "could not meaningfully compare error rates for words and nonwords: for each word, there was a single, correct pronunciation, but, for each nonword, there was some arbitrariness in what was counted as correct".

The studies also differ in terms of the number of word stimuli that have been used and whether they repeat. For instance, many studies have used small stimuli sets (e.g., 4 or 5 words) in which items were repeated (e.g., Klein, 1964; Burt, 1994, 1999, 2002; Bakan & Alperson, 1967; Fox et al., 1971). Later studies used larger stimulus sets in which no items were repeated (e.g., Monsell et al., 2001; Kinoshita et al., 2017). The repetition of stimuli interacts with word frequency to benefit low-frequency words only (Forster & Davis, 1984). Item repetition therefore attenuates the word-frequency effects (Burt, 2002).

The saliency of low-frequency or rare words is significantly reduced by repeating items decreasing the overall likelihood of obtaining a word-frequency effect (Burt, 2002). Similarly, the repetition of stimuli interacts with the lexicality effect. Repetition increases lexicality effects at long lags (Burt, 2002). As repetition effects are weak for nonwords relative to words, the repetition of items increases the likelihood of obtaining a lexicality effect (Burt 2002; de Heyer et al., 1988). To illustrate, in Burt (2002, experiment 5) a repetition manipulation was used, and a lexicality effect was obtained in which colour-naming latencies were greater for words than pronounceable nonwords. When the lexicality effect was analyzed using unrepeated items only, the lexicality effect was non-significant and colour-naming latencies were no different for words than nonwords.

Though Monsell et al.'s (2001) study addressed some of the methodological issues in previous work (e.g., by using a larger stimulus set, better-matched stimuli, mixed and individual stimulus presentation), several methodological issues are present in their work. For example, some of the word and pseudoword stimuli used in Monsell et al. (2001) began with the same letters of one of the response colour names (e.g., R, G, B, Y, and P), which may have influenced the magnitude of the word frequency effect. As an illustration, consider the word "realm" and the pseudoword, "relk", stimuli used by Monsell et al. (2001). Both begin with the same letter as the response colour "red". Stimuli with significant phonological overlap with one of the response colour words were also included (e.g., the word "tread", which has substantial overlap in phonology with the colour word "red", was used as one of the low-frequency words). This is an issue as colour-naming interference is reduced when the irrelevant word shares the onset letter with one of the response colours (e.g., Coltheart et al., 1999). It is likely that the condition with phonological overlap with the colour names will be slowed – because words in the incongruent condition are slowed more than the congruent condition. This could make it difficult to detect a word frequency effect.

Other issues that could have contributed to the inconsistencies in the word frequency and lexicality effects in the literature are the matching of items and items sets. Many of the studies used stimulus sets that were not well controlled. For instance, the items used by Klein (1964) were not matched on the number of letters, number of phonemes, or number of syllables across lists, and stimuli sets were not matched on neighbourhood density, onset class, number of voiced onsets, onset manner type, onset place of speech, and the proportion of nouns to verbs (for word stimuli). Further, the nonword stimuli used by Klein (1964) were not pronounceable.

Another example can be seen in the studies by Burt (1994;1999; 2002) in which items were not matched on the initial letter, onset phoneme, onset class, number of phonemes, and number of syllables across lists, and stimuli sets were not matched on neighbourhood density, onset class, number of voiced onsets, onset manner type and onset place of speech. Several studies (e.g., Klein, 1964; Fox, 1971; Burt, 2002; Kinoshita et al., 2017) did not take neighbourhood density into account (i.e., neighbourhood density was neither manipulated nor controlled for), variables that have been well demonstrated in the literature to influence lexical processing (see Andrews, 1996; Coltheart et al., 2001). This could have contributed to inconsistencies in the frequency and lexicality effects observed in previous work, since neighbourhood density has been shown to interact with word frequency and lexicality effects.

Another problem with the stimuli in Monsell et al. (2001) was the inclusion of words that were semantically associated with colour-words, which could have contributed to the inconsistencies in the word frequency and lexicality effects. Two such examples were the words "heart" and "blood" which both have a semantic association to the colour red and were included in the high-frequency word stimuli set. These colour-associated words were mixed in with neutral words in their stimuli lists. Following this discussion of methodological differences and inconsistencies in findings across studies using the non-colour word Stroop task, it is evident that a better controlled experiment is required.

The Present Study

The goal of the present study is to examine the dual task account of the non-colour word Stroop task. According to this account, the presentation of a coloured word triggers two tasks (1) word reading and (2) colour naming. The word reading task is triggered unintentionally and proceeds ballistically until processing priority can be redirected to the colour-naming task. The dual-task account predicts that colour-naming performance should be affected by psycholinguistic variables that affect reading processes prior to the point at which priority is redirected to the colour naming task, and not be affected by variables that affect performance after this point. Problematic for this account is the inconsistent evidence that colour-naming performance is affected by word frequency and lexicality.

The inconsistent evidence for the word frequency and lexicality effects in the colour-naming task could be due to confounds in the stimulus sets used in previous work. In order to address these issues, the present study will (1) replicate the general methods used by Monsell et al. (2001), but with (2) better controlled stimuli, and (3) the addition of a new psycholinguistic variable to permit greater precision in estimating the locus of control.

As in Monsell et al. (2001) participants will complete (1) a reading aloud task in which the written stimuli are read aloud to determine whether they produce well established effects and

(2) a colour naming task in which the written words are to be ignored and the ink-colour is to be named aloud. The order of the two tasks will be counterbalanced across participants. The reading aloud task will be used to confirm that the stimuli used in the present study produce the typical patterns of effects observed in the literature. The colour-naming task will be used to assess whether the processes indexed by these effects occur without intention.

Also, like Monsell et al. (2001), the written stimuli will consist of words and nonwords. Three stimulus types will be used, high-frequency words (e.g., “clock”), low-frequency words (e.g., “kilt”), and nonwords (e.g., “chim”). Unlike previous work, this variable will be factorially crossed with neighbourhood density to allow greater precision in assessing the locus of control.

Neighbourhood Density

The concept of a lexical neighbourhood was first formally introduced by Landauer and Streeter (1973), where they defined it as a “set of words in the language from which a given stimulus word is indistinguishable after a specified loss of information about the stimulus word” (p.120). Coltheart and colleagues (1977) were the first psychologists to investigate the effect of neighbourhood density in reading. Their definition of neighbourhood density (often referred to as “Coltheart’s N”), was as the set of words (e.g., neighbours) that can be created from a target word by replacing one single letter and preserving the other letters. Recently, Yarkoni et al. (2008; see Tulkens et al., 2020 for a review) defined neighbourhood density as the mean Levenshtein distance (Levenshtein, 1966) to the 20 closest orthographic neighbours (OLD20) of a given target word among all words between three and eight letters in length in the English Lexicon Project. Levenshtein’s distance (1966) is the minimum number of substitutions, insertions, and deletions needed to generate one letter string from another. Although neighbourhood density is often defined in terms of orthography (orthographic neighbourhood density), the same metrics can be applied to a word's phonology (phonological neighbourhood density).

The neighborhood density effect in reading aloud refers to the observation that words and nonwords with more neighbours are read aloud faster than those with fewer neighbours (Andrews, 1989;1992; Laxon et al., 1992; Mulatti, Reynolds & Besner, 2006; Peereman & Content, 1995; Sears et al., 1995; Tulkens et al., 2020; Reynolds & Besner, 2002; 2004). Two broad accounts of neighbourhood density have been proposed. These are, (1) an early account where the activation of neighbours in the orthographic lexicon facilitates letter processing (e.g., Andrews 1989; Sears et al., 1995; Reynolds & Besner, 2002) and (2) a late account where the activation of neighbours in the orthographic lexicon facilitates phonological processing in the phonological lexicon and phoneme buffer (e.g., Reynolds & Besner, 2002, 2004; Mulatti et al., 2006; Peereman & Content, 1995). Consistent with neighbours affecting lexical processing for words, the joint effects of neighbourhood density and word frequency are interactive in reading aloud such that neighbourhood density influences the time to read low-frequency words and not high-frequency words (e.g., Besner & McCann, 1987; Grainger, 1990). Inconsistent with neighbourhood density affecting early processes, its effects are additive with stimulus contrast in reading aloud (Reynolds & Besner, 2002; 2004). Consistent with the late account, the effect of neighbourhood is affected by the ratio of neighbours with inconsistent phonology for shared orthographic rime units relative to the number of neighbours with phonologically consistent rime (e.g., Peereman, 1995). Also consistent with the late account is evidence from Mulatti et al. (2006) who demonstrated that when orthographic neighbourhood density is controlled for, a facilitatory effect of phonological neighbourhood density on reading aloud is observed, whereas when phonological neighbourhood density was controlled for, a null effect of orthographic neighbourhood density on reading aloud was reported. They provided a computational simulation of these effects by lesioning Coltheart and colleagues' (2001) DRC Model. The model was able to replicate human performance when lesioned so that only feedforward connections from the orthographic lexicon through the phonological lexicon to the phoneme buffer were present and the non-lexical route was removed.

The effects of neighbourhood density therefore likely arise in the connections between the orthographic lexicon and the phonological lexicon. This effect therefore affects performance after the earliest effects of word frequency and before the effects of lexicality. Therefore, including neighbourhood density as a factor permits a more precise identification of the locus of control.

Addressing Methodological Differences in the Non-Colour Word Stroop Task

One concern with previous work (e.g., Klein, 1964; Bakan & Alperson, 1967; Fox et al., 1971) was with regards to the mode of stimulus presentation (i.e., card stimulus presentation vs. individual stimulus presentation). To address this, the present study used trial-by-trial, individual stimulus presentation and measured response times for individual items rather than for a list of items. Another concern with previous work (e.g., Kinoshita et al., 2017) was the failure to include a reading aloud task as a control condition. The present study therefore measures and analyzes performance on both a word reading task and a colour naming task.

There were several issues with stimulus construction in previous work. To not confound lexicality and pronounceability, pronounceability was controlled for by ensuring that all nonwords included in the study were orthographically legal in English, pronounceable, and had a maximum of only two non-arbitrary pronunciations. A large stimulus set in which none of the stimuli were repeated was used to ensure that the effects of word frequency and lexicality were not affected by repetition. Stimuli were better controlled – none of the stimuli had orthographic or phonological overlap in the initial position with any of the four possible ink colours, a more recent corpus was used to obtain stimulus word frequencies, and none of the stimuli had any semantic association to any of the four possible ink colours.

Another concern with previous work was that the stimuli in the different conditions were confounded with other psycholinguistic variables. Therefore, unlike the previous work by Klein (1964), Burt (2002), and Monsell et al. (2001) the stimuli in each condition were matched on the number of letters and the number of phonemes. All the stimuli were monomorphemic and consisted of one syllable. Stimuli sets were also matched on onset phoneme, onset class, syntactic class (words only), onset manner type, and onset place of speech.

The dual task account predicts that the main effects of variables in the colour naming task should be in the same direction as the main effects observed in the reading aloud task. Past research suggests that a portion of the word frequency effect arises from processing in the orthographic lexicon and therefore arise from processes prior to the neighbourhood density effects, which occur due to processing in the connections between the orthographic and phonological lexicons. These, in turn, arise from processes prior to lexicality effects, which occur due to processing in the phoneme buffer. The dual task account predicts that if one effect does not arise in the colour naming task, then no effects arising from subsequent processes should affect colour-naming performance, either.

Method

Participants

One hundred and twenty undergraduate students from Trent University participated for partial course credit toward an eligible class. All participants self-reported as monolingual English speakers with normal colour vision, and normal or corrected to normal visual acuity.

In order to ensure sufficient power in the colour naming task we estimated our effect sizes based on Monsell et al. (2001). An examination of Monsell et al. (2001) suggests that they observed a 50% reduction in effect size in the colour naming task relative to the word reading task. This reduction is comparable to the ones observed by Burt (1994, 1996, 2002). Given the similarity in stimulus construction methods in the present studies and previous work by Reynolds and colleagues (2002; 2004; 2005; 2006; 2008) we used past work from those researchers to estimate the effect sizes in the present work. As our estimate for the effect size for the word frequency effect we used Reynolds and Besner (2008). They reported an effect size of η_p^2=.732 therefore we estimated the effect size in colour naming to be η_p^2=.383. For the lexicality effect we used Reynolds and Besner (2005). They reported a lexicality effect that was also η_p^2=.732. We therefore estimated the effect size to be η_p^2=.383 in colour naming. Our estimate for the size of the N effect in nonwords was based on Reynolds and Besner (2004). They observed an effect size of η_p^2=.500. We therefore estimated the effect size in colour naming to be η_p^2=.259. For words we used Mulatti et al. (2006) who reported an effect size of η_p^2=.248. We therefore estimated the effect size in colour naming to be η_p^2=.130. Based on these calculations we estimated that a sample size where N=120 would ensure power of at least 0.8 to detect our effects. Alpha was set at .05.

Table 2.

Stimulus properties as a function of condition.

	High -Frequency Words		Low - Frequency Words		Pseudowords	
	High N	Low N	High N	Low N	High N	Low N
Lexicality	word	word	word	word	Non-word	Non-word
Mean number of Syllables	1	1	1	1	1	1
Mean Number of Letters	4.5	4.8	4.4	4.8	4.4	4.8
Mean Number of Phonemes	3.4	3.7	3.5	3.8	3.6	3.9
Written Frequency (CELEX corpus)	1688	5129.3	34.6	16.4	n/a	n/a
Mean number of orthographic Neighbours (CELEX corpus)	9.3	1.5	9.1	0.8	8.5	1.4
Mean number of Phonological Neighbours (CELEX corpus)	20.3	11.8	17.2	9.6	15.7	9.8
Mean OLD20	1.02	1.38	1.02	1.38	n/a	n/a
Mean PLD20	1	1.1	1.02	1.1	n/a	n/a
Frequency count of Voiced Onset						
Voiced	8	7	8	8	7	8
Voiceless	30	31	30	30	31	30
Frequency count of Onset Manner types						
Stop	18	18	17	18	19	18
Affricate	1	1	1	1	0	1
Fricative	14	14	15	14	14	14
Nasal	2	2	2	2	2	2
Glide	3	3	3	3	3	3
Frequency count of different types of places of speech						
Labial	11	12	14	15	15	15
Coronal	15	14	15	14	12	14
Dorsal	12	12	9	9	11	9

Stimuli

The stimuli consisted of 144 words and 72 pronounceable nonwords ranging in length from 4 to 5 letters. The word stimuli were comprised of 36 high-frequency words with many neighbours, 36 high-frequency words with few neighbours, 36 low-frequency words with many neighbours, and 36 low-frequency words with few neighbours. The nonword stimuli consisted of 36 nonwords with many neighbours and 36 nonwords with few neighbours. Characteristics for the word stimuli can be seen in Table 2. Values were calculated using the SUBTLEX lexical database (Brysbaert & New, 2009). Characteristics for the nonword stimuli were calculated using the ARC database (Rastle et al., 2002). For the purposes of the present study, the categorization of neighbourhood density as large and small deliberately confounded orthographic and phonological neighbourhoods calculated using both Coltheart's N methodology (Coltheart et al., 1977) and the Levenstein's distance methodology (e.g., OLD20; Yarkoni et al., 2008).

All stimuli were monosyllabic. None of stimuli began with a first letter or phoneme of the four target colours used in the colour naming task. Items in each condition were matched on the initial letter, onset phoneme, onset class, number of phonemes, number of letters and number of syllables across lists. The high and low frequency stimulus sets were matched on syntactic class. All item sets were matched on onset class, number of voiced onsets, onset manner type and onset place of speech. The distribution of values across different dimensions for the stimuli can be found in Table 2.

Apparatus

A DELL XPS 8930 computer with Windows 10 Pro and an NVIDIA Geforce GTX 1050TI video card and a DELL 24-inch Gaming Monitor (Model S2421HGF) with a native resolution of 1920 × 1080 (running at 120 Hz) were used to conduct the experiment. Stimulus presentation and response collection were controlled using Eprime 3.0 (https://pstnet.com/products/e-prime/). Vocal responses were collected using a Chronos response box and microphone assembly.

Stimulus Displays

The letter strings were presented in a 24-point Courier New font on a single line in the center of the computer screen. The stimuli subtended approximately 0.6° high × 2.0° to 3.2° wide. In the reading aloud task, the letter strings were presented in white. In the colour naming task, the word and nonword stimuli were displayed in four different font colours (e.g., red, yellow, blue, and green). The standard E-Prime colours using were used (Schneider et al., 2002). The assignment of colour to item was randomized for each participant with the condition that each colour occur an equal number of times per stimulus condition and that no colour would occur more than four times in a row. All items were displayed on a black background.

Procedure

All participants were run through the experiment individually in the presence of a research assistant. Each participant completed both the colour-naming and the word-naming tasks. The order of the two tasks was counterbalanced. Participants were assigned to counterbalance pseudo-randomly based on the order in which they arrived in the lab. The apparatus was adjusted for each participant so that they sat approximately 50cm from the computer display and stimulus presentation was at eye level.

For both tasks, participants were instructed to respond vocally, into the microphone, as quickly and as accurately as possible. In the colour naming task, participants were instructed to say aloud the name of the font colour and to ignore the letter string. In the reading aloud task participants were told to read aloud (e.g., pronounce) the letter string. In both tasks they were told that they would be presented with both words (e.g., cat) and word-like stimuli (e.g., blat). Each task consisted of one practice block followed by three experimental blocks. Each practice block contained 24 trials. Each experimental block contained 64 trials. In all, there were 24 practice trials and 192 experimental trials per task. Participants were allowed a brief break between each block. All stimulus types were randomly intermixed in both the practice and experimental blocks with the following conditions: (1) no single condition could occur on more than two consecutive trials and (2) there were an equal number of trials in each condition. Each task took approximately twenty minutes to complete.

In each task a trial unfolded as follows. A trial began with a fixation marker (+) in the center of the screen for 1000 ms, followed by a blank screen for 250 ms. The target letter string was then presented at fixation until a vocal response was made. A blank screen then appeared for 1,500 ms. During this time the experimenter coded the vocal response. In the word-naming task responses were coded as correct, incorrect, or mistrial (e.g., cough or voice key failed to activate). An error was defined as a clear mispronunciation such as an extra or deleted phoneme, a hesitation or pause, or a stutter or lexicalization of the nonwords. The dominant pronunciation of a grapheme was not required for a pronunciation to be deemed correct (e.g., “gnowth” read so as to rhyme with either “growth” or “south” was acceptable). In the colour-naming task, the experimenter coded the response as the colour (yellow, blue, green, red), as the wrong task (e.g., reading the word), or as a mistrial (e.g., cough or voice-key failure).

Results

Table 3.

Mean reaction time (RT) in seconds and percentage error (PE) as a function of trials in the Word Naming and Colour Naming Tasks across conditions with the lexicality, frequency, and neighbourhood (N) effects for each task.

	Word Reading Task						**Colour Naming Task**					
	Low N		**High N**				**Low N**		**High N**			
	RT	PE	RT	PE	*N Effect*	*N Effect*	RT	PE	RT	PE	*N Effect*	*N Effect*
Nonwords	778	4.9	713	3.1	65	1.8	650	0.4	649	0.6	1	-0.2
Low Frequency Words	710	2.7	666	1.7	44	1.0	649	0.5	646	0.3	3	0.2
High Frequency Words	630	0.8	629	0.9	1	-0.1	643	0.6	639	0.4	4	0.2
Lexicality Effect	68	2.2	47	1.4			1	-0.1	3	0.3		
Frequency Effect	80	1.9	37	0.8			6	-0.1	7	-0.1		

Data were analyzed with the R statistical software program (R Core Team, 2022). The data for the word reading task and the colour naming task were analyzed separately. Within each task, the response time (RT) and percentage error (PE) data were analyzed separately with by-subjects (F_s) and by-items (F_i) analyses of variance (ANOVAs) using the EZ package (Lawrence, 2016). In the by-subjects analysis, condition means are calculated for each participant by collapsing across items. In the by-items analysis, condition means are calculated for each item by collapsing across subjects (Clark, 1973). The subject-data were analyzed using a mixed-model ANOVA with stimulus type (high-frequency words, low-frequency words and nonwords) and neighbourhood density (N; high vs. low) as repeated factors and counterbalance (CB) as the between-subjects factor. The item-data were analyzed with stimulus type and neighbourhood density as between-items factors and counterbalance as the repeated factor.

Eight participants were excluded from the formal analysis of the data. Four participants were removed due to a high number of mistrials (> 20%) arising from microphone failures, one was removed because of hardware failure (e.g., a technological error resulting in the subject's data being overridden), and three were removed due to a high number of errors (e.g., wrong-task, stutter, cough, hesitation or a clear mispronunciation such as a phoneme addition or deletion or a lexicalization; >20%). The word-naming and colour-naming performance data can be seen above in Table 3.

Word-Reading Analysis

The word reading data can be seen in Table 3. Prior to analyzing the response time data, trials on which a reading error (1.66%) or mistrial (2.67%) occurred were removed and the remaining trials were analyzed. The remaining participants correct RT data were submitted to a recursive data trimming procedure in which the criterion for outlier removal was calculated separately for each participant in each cell on the basis of the number of observations in each cell (Van Selst & Jolicoeur, 1994). This resulted in the removal of 2.34 % of the correct RT data. See Appendix A for the complete set of ANOVA tables.

Word-Reading RT Analysis

There were no significant effects of counterbalance for the by-subjects analyses, but there was for the by-items analyses. There was a tendency for effects to be slightly smaller in Counterbalance 2. None of the effects changed in a meaningful way across counterbalances and are therefore not discussed further.

There was a significant main effect of stimulus type [$F_s(2, 232) = 143.5$, $p < .001$, $MSE = 5457$, $\eta_p^2 = .553$; $F_i(2, 186) = 100.4$, $p < .001$, $MSE = 4161$, $\eta_p^2 = .519$]. There was an effect of frequency where the high-frequency words (630 ms) were read aloud faster than the low-frequency words (688ms), [$F_s(1,116) = 102.9$, $p < .001$, $MSE = 3673$, $\eta_p^2 = .470$; $F_i(1, 124) = 63.92$, $p < .001$, $MSE = 3216$, $\eta_p^2 = .340$]. There was a lexicality effect where the low-frequency words (688ms) were read faster than non-words (746ms) [$F_s(1, 116) = 146.302$, $p < .001$, $MSE = 2770$, $\eta_p^2 = .559$; $F_i(1,124) = 39.90$, $p < .001$, $MSE = 5364$, $\eta_p^2 = .244$].

There was a significant main effect of neighbourhood density [$F_s(1, 116) = 96.2$, $p < .001$, $MSE = 2410$, $\eta_p^2 = .453$; $F_i(1, 186) = 31.9$, $p < .001$, $MSE = 4161$, $\eta_p^2 = .146$]. Letter strings (i.e., both words and nonwords) with many neighbours (669ms) were read faster than letter strings with few neighbours (706ms).

The two-way interaction between stimulus type and neighbourhood density was significant [$F_s(2, 232) = 59.3$, $p < .001$, $MSE = 1164$, $\eta_p^2 = .338$; $F_i(2, 186) = 8.85$, $p < .001$, $MSE = 4161$, $\eta_p^2 = 0.087$]. Word frequency interacted with neighbourhood density where the effect of neighbourhood density was smaller for high-frequency words (1.3ms) than for low-frequency words (43.6ms) [$F_s(1, 116) = 76.39$, $p < .001$, $MSE = 650.4$, $\eta_p^2 = .397$; $F_i(1, 124) = 9.45$, $p < .001$, $MSE = 3216$, $\eta_p^2 = .071$]. Lexicality interacted with neighbourhood density where the effect of neighbourhood density was smaller for low-frequency words (43.6ms) than for nonwords (65.1ms) by-subjects [$F_s(1, 116) = 19.85$, $p < .001$, $MSE = 1071$, $\eta_p^2 = .146$], but not by-items [$F_i(1,124) = 1.599$, $p = .208$, $MSE = 5411$, $\eta_p^2 = .013$].

Word-Reading PE Analysis

In the percentage error (PE) analysis, there were no effects of counterbalance in the by-subjects analyses, but there were for the by-items analyses. Similar to the reaction time analysis there was a tendency for effects to be slightly smaller in counterbalance 2, when reading was the second task. None of the effects changed in a meaningful way across counterbalances and are therefore not discussed further.

There was a significant main effect of stimulus type [$F_s(2, 232) = 32.00$, $p < .001$, *MSE* =17.8, η_p^2=.216; $F_i(2, 186) = 4.50$, $p = 0.01$, *MSE* =17.0, η_p^2 =.046]. There was an effect of frequency where the high-frequency words (0.9%) were less error prone than the low-frequency words (2.2%), [$F_s(1, 116) = 102.9$, $p <.001$, *MSE* =3672.7, η_p^2= .470; $F_i(1, 124) = 63.9$, $p < .001$, *MSE* =3216.1, η_p^2= .340]. There was an effect of lexicality where low-frequency words (2.2 %) were less error prone than the non-words (4.0%), [$F_s(1, 116) = 146.3$, $p< .001$, *MSE* =2770.3, η_p^2=.558; $F_i(1, 124) = 39.2$, $p < .001$, *MSE* =5411.2, η_p^2=.240].

There was a significant main effect of neighbourhood density [$F_s(1, 116) = 21.85$, $p <$ .001, *MSE* =15.14, η_p^2= .159; $F_i(1, 186) = 31.98$, $p < .001$, *MSE* =4120.4, η_p^2=.147]. Letters strings (i.e., both words and nonwords) with many neighbours (1.9%) were less error prone than those with few neighbours (2.8%).

The two-way interaction between stimulus type and neighbourhood density was significant [$F_s(2, 232) = 22.52$, $p < .001$, *MSE* = 14.76, η_p^2= .163; $F_i(2, 186) = 13.47$, $p < .001$, *MSE* =17.13, η_p^2=0.127]. There was a significant interaction between the word frequency and the neighbourhood density in which the effect of neighbourhood density was smaller for high-frequency words (0.1%) than for low-frequency words (1.0%), [F_s $(1, 116) = 76.39$, $p < .001$, *MSE*=650.45, η_p^2=.397; $F_i(1, 124) = 9.445$, $p = .003$, *MSE* =3216.1, η_p^2=.071]. There was also an interaction between lexicality and the neighbourhood density where the effect of neighbourhood density was smaller for low-frequency words (1.0%) than for nonwords (1.8%), [$F_s(1, 116) =$ 19.85, $p < .001$, *MSE* =1070.9, η_p^2=.146; $F_i(1, 124) = 1.599$, $p = .208$, *MSE* =5411.2, η_p^2=.013).

Word Reading Summary

The typical effects of word frequency (Burt, 2002; Monsell et al., 2001; McCann & Besner, 1987; McCann et al., 1988; Reynolds & Besner, 2005), lexicality (Burt 2002; Monsell et al., 2001; Reynolds & Besner, 2005), and neighbourhood density (Andrews, 1989;1992; Laxon et al., 1992; Peereman & Content, 1995; Reynolds & Besner, 2004; 2006; Sears et al., 1995; Tulkens et al., 2020) on reading aloud performance were observed in the present experiment. Further, the standard interactions effects between word-frequency and neighbourhood density (Besner & McCann, 1987; Grainger, 1990), and between lexicality and neighbourhood density (Besner & McCann, 1987; Grainger, 1990) were observed. If processes indexed by these effects occur without intention, then these effects should also be observed in the colour naming task.

Colour-Naming Analysis

The colour naming data were analyzed identically to the word reading data. The colour naming performance data can be seen in Table 3. Prior to analyzing the response time data, trials on which there was a reading error (0.156%), a colour error (0.312 %), or mistrial (e.g., voice key failure; 2.248 %) were removed. Outlier trimming resulted in the removal of 1.84 % of the remaining correct RT data. See Appendix A for the complete set of ANOVA tables.

Colour-Naming RT Analysis

There were no effects of counterbalance for the by-items analyses, but there were for the by-subjects analyses. There was a tendency for effects to be slightly smaller in Counterbalance 1 when the colour-naming task was performed second. In the subject data, reaction times tended to be faster for participants who completed the reading task first compared to participants who completed the colour-naming task first. None of the effects changed in a meaningful way across counterbalances and are therefore not discussed further.

There was a significant main effect of stimulus type [$F_s(2, 234) = 7.151$, $p = .001$, $MSE = 683.43$, $\eta_p^2 = .058$; $F_i(2, 186) = 8.073$, $p < .001$, $MSE = 316.49$, $\eta_p^2 = .080$]. There was an effect of word frequency where the high-frequency words (641 ms) were colour-named faster than the low-frequency words (648ms), [$F_s(1, 117) = 9.295$, $p = .003$, $MSE = 602.06$, $\eta_p^2 = .074$; $F_i(1, 124) = 9.038$, $p = .003$, $MSE = 346.38$, $\eta_p^2 = .068$]. The effect of lexicality was not significant, colour-naming performance did not differ significantly between low-frequency words (648 ms) and non-words (649 ms), [$F_s(1,117) = 0.554$, $p = .458$, $MSE = 622.63$, $\eta_p^2 = .005$; $F_i(1,124) = 0.369$, $p = .544$, $MSE = 301.78$, $\eta_p^2 = .003$].

There was no main effect of neighbourhood density [$F_s(1, 117) = 2.023$, $p = .158$, $MSE = 458.63$, $\eta_p^2 = .017$; $F_i(2,186) = 2.34$, $p = .128$, $MSE = 316.49$, $\eta_p^2 = .012$]. Colour-naming performance did not differ significantly between letter strings (i.e., both words and nonwords) with many neighbours (645 ms) and letter strings with few neighbours (647 ms).

The two-way interaction between stimulus type and neighbourhood density was not significant [$F_s(2,234) = 0.268$, $p = .765$, $MSE = 432.47$, $\eta_p^2 = .002$; $F_i(2,186) = 0.161$, $p = .851$, $MSE = 316.49$, $\eta_p^2 = .002$]. The effect of neighbourhood density was not affected by word frequency. The effect of neighbourhood density for the high-frequency words (4 ms) did not differ from the effect observed for the low-frequency words (3 ms), [$F_s(1, 117) = 0.061$, $p = .805$, $MSE = 491.72$, $\eta_p^2 < .001$; $F_i(1,124) = 0.044$, $p = .834$, $MSE = 346.38$, $\eta_p^2 < .001$]. The effect of neighbourhood density was not affected by lexicality. There was no difference between the effect of neighbourhood density observed for the low-frequency words (3 ms) and the non-words (1 ms), [$F_s(1,117) = 0.257$, $p = .613$, $MSE = 357.13$, $\eta_p^2 = .002$; $F_i(1,124) = 0.124$, $p = .725$, $MSE = 301.78$, $\eta_p^2 = .001$].

Colour-Naming PE Analysis

In the percentage error (PE) analysis, similar to the reaction time analysis, there were no effects of counterbalance in the by-items analyses, but there were for the by-subjects analyses. There was a tendency for effects to be slightly smaller in Counterbalance 1. None of the effects changed in a meaningful way across counterbalances and are therefore not discussed further.

There was a significant main effect of stimulus type in by-items but not by-subjects [$F_s(2,234) = 0.460$, $p = .632$, $MSE = 1.459$, $\eta_p^2 = 0.004$; $F_i(2, 186) = 8.073$, $p < .001$, $MSE = 316.49$, $\eta_p^2 = .080$]. There was no effect of word-frequency. Colour-naming error rates did not differ significantly for high-frequency words (0.5%) and low-frequency words (0.4%), [$F_s(1,117) = 0.249$, $p = .619$, $MSE = 1.332$, $\eta_p^2 = .002$; $F_i(1,124) = 0.559$, $p = .456$, $MSE = 0.316$, $\eta_p^2 = .005$]. Similarly, there was no effect of lexicality. Colour-naming error rates did not differ significantly for low-frequency words (0.4%) and non-words (0.5%), [$F_s(1,117) = 0.054$, $p = .816$, $MSE = 1.457$, $\eta_p^2 < .001$; $F_i(1, 124) = 0.134$, $p = .715$, $MSE = 0.329$, $\eta_p^2 = .001$].

There was no significant main effect of neighbourhood density [$F_s(1,117) = 0.287$, $p = .593$, $MSE = 1.781$, $\eta_p^2 = .003$; $F_i(1,186) = 2.34$, $p = .128$, MSE = 316.49, $\eta_p^2 = .012$]. Colour-naming error rates did not differ significantly between letter strings (i.e., both words and nonwords) with many neighbours (0.4%) and letter strings with few neighbours (0.5%).

The two-way interaction effect between stimulus type and neighbourhood density was not significant [$F_s(2, 234) = 2.462$, $p = .087$, $MSE = 1.273$, $\eta_p^2 = .021$; $F_i(2,186) = 0.161$, $p = .851$, $MSE = 316.49$, $\eta_p^2 = .002$]. The effect of neighbourhood density did not differ significantly for high-frequency words (0.5%) and low-frequency words (0.4%), [$F_s(1,117) = 2.316$, $p = .131$, $MSE = 472.46$, $\eta_p^2 = .019$; $F_i(1,124) = 0.559$, $p = .456$, $MSE = 0.316$, $\eta_p^2 = .005$]. The effect of neighbourhood density did not differ significantly between low-frequency words (0.4%) and non-words (0.5%), [$F_s(1,117) = 0.882$, $p = .350$, $MSE = 368.78$, $\eta_p^2 = .008$; $F_i(1, 124) = 0.134$, $p = .715$, $MSE = .329$, $\eta_p^2 = .001$].

Colour Naming Summary

Though the typical effects of word frequency, lexicality, and neighbourhood density were found on word reading latencies, this did not translate to the colour naming task. The only significant effect on colour-naming latencies was word frequency. The effects of lexicality, and neighbourhood density were not significant.

Discussion

The goal of the present study was to examine the dual task account of the non-colour word Stroop task. According to this account, the presentation of the non-colour word stimulus triggers two tasks (1) a word reading task and (2) a colour naming task. Word reading is initially triggered unintentionally by the word-like stimulus and proceeds ballistically until processing priority can be redirected to the colour-naming task. In this account, priority is assigned to processing the word (the word reading task) over processing the colour (the colour naming task). Though some processing associated with the colour component and the word component of the stimulus may proceed in parallel, a portion of the processing associated with the colour naming task is functionally postponed. Word reading processes proceed unintentionally and ballistically until some critical point in the reading process is reached where processing associated with the word reading task is deprioritized. Priority is then given to the processing associated with the colour naming task. It is at this point when the portion of the processing associated with the colour naming task that was functionally postponed resumes.

The dual-task account predicted that colour-naming performance should be affected by psycholinguistic variables that affect reading processes prior to the point at which priority is redirected, and not be affected by variables that affect performance after this point. This account was examined in a non-colour word Stroop experiment based on the general methods used by Monsell et al. (2001), but with the addition of a new psycholinguistic variable (e.g., neighbourhood density) to permit greater precision in estimating the locus of control.

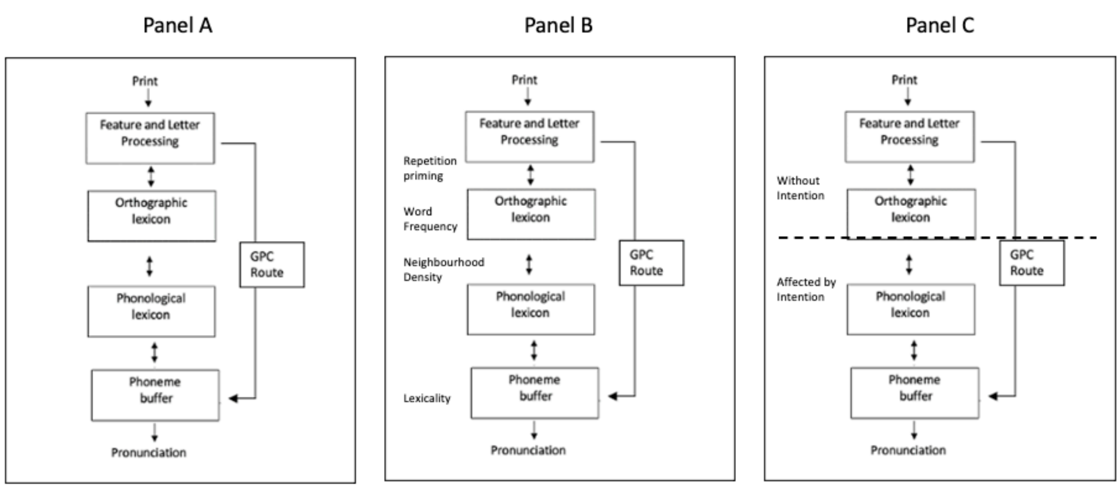
It was argued that the pattern of interference obtained would identify the locus (loci) of control in the reading system using the DRC model's architecture (See Figure 3, Panel A). This locus represents the critical point at which processing ceases to be unintentional in the reading system. Here we argue that if a psycholinguistic variable known to affect a process involved in reading aloud is present when colour-naming the ink colour that a word or nonword is printed in, this indicates that the corresponding process in the reading system indexed by that variable is automatic in the sense that it is occurring unintentionally. Alternatively, if the effect of a psycholinguistic variable known to affect a process involved in reading aloud is absent when colour-naming, this indicates that the processing in the reading system indexed by this variable does not occur unintentionally and that processing associated with the word dimension of the stimulus is halted at this point or before this point in the reading system, so that priority can be reassigned to the processing of the colour dimension of the stimulus.

Word Frequency

There was significant effect of word-frequency in the colour naming task, where colour-naming latencies were significantly shorter for high-frequency words than for low-frequency words. This same pattern was observed in the word-reading task, replicating previous research that has examined the effects of word frequency on reading performance (e.g., Andrews, 1989; Burt, 2002; Forster & Chambers, 1973; Monsell et al., 1989; 2001). This finding is also consistent with previous studies that have examined the effect of word frequency in the non-colour word Stroop task with individual stimulus presentation (e.g., Monsell et al., 2001; Burt, 1994; 1999; 2002). Word frequency effects are widely believed to arise from lexical processing in the orthographic and phonological lexicons (e.g., Coltheart et al., 2001, see Figure 3, Panel B). Therefore, the present observation of a word frequency effect in colour-naming supports the notion that word processing up to and including the orthographic lexicon and/or phonological lexicon is occurring without intention (Coltheart et al., 2001; Reynolds & Besner, 2006; Ziegler et al., 2004, see Figure 3, Panel C). This is consistent with the claim that the word-frequency effect affects a stage in lexical processing that is either automatic or partially automatic (Rabovsky et al., 2008; 2019, Reynolds & Besner, 2006).

Figure 3.

The DRC model's architecture (Panel A), the loci of reading effects in the reading system (Panel B), and the critical point at which processing during word reading switches from being unintentional to being affected by intention (Panel C).



Lexicality

There was no effect of lexicality in the colour naming task. In contrast there was a significant effect of lexicality in the word reading task, where low-frequency words were named significantly faster and more accurately than nonwords, replicating previous research that has examined the effects of lexicality on reading performance (e.g., Kinoshita et al., 2004; Monsell et al., 2001; Reynolds & Besner, 2005; Zevin & Balota, 2000). The absence of a lexicality effect in the colour naming task replicates the previous findings of Monsell et al. (2001; see also Kinoshita et al., 2017, but not Burt, 2002). The received view is that the lexicality effect in reading aloud arises late in processing at the level of the phonemic buffer (Coltheart et al., 1977; Coltheart et al., 2001; McCann et al., 2000; Wydell et al., 2003; Zorzi et al., 1998; Rabovsky et al., 2008; 2019, see Figure 3, Panel B). Therefore, the absence of a lexicality effect in the colour-naming task suggests that processing in the reading system at the phoneme buffer is not occurring without intention (see Figure 3, Panel C). This is consistent with the lexicality effect affecting a stage of processing in the reading system that is not automatic in the sense that it does not occur unintentionally and ballistically. This finding indicates that processing associated with the word dimension of the stimulus is halted at or before the point in the reading system indexed by the lexicality effect (e.g., the phoneme buffer) so that priority can be shifted from the word reading task to the colour-naming task (e.g., the processing of the colour dimension of the stimulus).

Neighbourhood density

There was no effect of neighbourhood density in the colour-naming task. In contrast a significant effect of neighbourhood density in the word-reading task was observed, where words and non-words with large neighbourhoods were read aloud significantly faster and more accurately than those with smaller neighbourhoods replicating previous research (e.g., Andrews, 1989, 1992; Balota et al., 2004; Carreiras et al., 1997; Grainger, 1990; Huntsman & Lima, 2002; Mulatti et al., 2006; Peereman & Content, 1995, 1997; Reynolds & Besner, 2004; Sears et al., 1995; for reviews see Andrews, 1997; Perea, 2015).

The effect of neighbourhood density is widely believed to arise through the connections between the orthographic and phonological lexicons (Mulatti et al., 2006; Peereman & Content, 1995; Reynolds & Besner, 2002; 2004, See Figure 3, Panel B). The absence of a neighbourhood effect in the colour-naming task suggests that lexical processing between the orthographic lexicon and the phonological lexicon in the reading system is not occurring unintentionally (See Figure 3, Panel C). This is consistent with neighbourhood density effect affecting a stage of processing in the reading system that is not automatic in the sense that it does not occur unintentionally and ballistically. This finding indicates that processing associated with the word dimension of the stimulus is halted at or before the point in the reading system indexed by the neighbourhood density effect so that priority can be shifted from the word reading task to the colour-naming task (e.g., the processing of the colour dimension of the stimulus).

Locus of Control

The dual task account of the colour naming task made two predictions (1) if an effect of a variable arises in the colour naming task the effect should be in the same direction as the effect observed in the reading aloud task, and (2) if one effect does not arise in the colour naming task, then no effects arising from subsequent processes should affect colour-naming performance, either. In the word reading task, a significant word-frequency effect, lexicality effect, and neighbourhood density effect was observed. In the colour naming task, only a word frequency effect was observed. Consistent with the first prediction, the effect of word frequency obtained in the colour naming task was in the same direction as the effect of word frequency in the word reading task. Consistent with the second prediction, no effects arose from processes after the word frequency effect.

As noted earlier, previous research indicates that part of the word frequency effect occurs early in the reading process at the level of the orthographic lexicon; that the neighbourhood effect occurs after word frequency effects but before lexicality effects, during processing between the orthographic and phonological lexicons; and that the lexicality effect occurs late in the reading process at the level of the phoneme buffer. The finding that only a word frequency effect was present in the colour-naming task indicates that the word is being unintentionally processed to the level of the orthographic lexicon but that it is not being unintentionally processed at the level of the phonological lexicon or phoneme buffer (See Figure 3, Panel C).

The present findings indicate that during the colour-naming task processing in the orthographic lexicon, as indexed by the presence of the word-frequency effect, is triggered by the word-like dimension of the stimulus and occurs unintentionally. The findings also indicate that in the colour-naming task, processing of the word-dimension of the stimulus is not occurring at the phoneme buffer, as indicated by the absence of a lexicality effect, and is not occurring between the orthographic lexicon and phonological lexicon, as indicated by the absence of a neighbourhood effect. The pattern of interference observed in the colour-naming task suggests that unintentional processing of the word dimension of the stimulus is occurring during, or immediately after, processing in the orthographic lexicon. Intention is then inhibiting the processing of the word, at some point between the orthographic and phonological lexicons, and redirects priority to the task of colour naming the stimulus (i.e., to the processing of the colour dimension of the stimulus).

It is possible to map the results of the present experiments onto the dual-route architecture (see Figure 3, Panel C). Processing up to and including the orthographic input lexicon comprises early processing that is unintentional. Processes that occur after the activation of representations in the orthographic input lexicon, particularly the activation of phonological representations, comprise later processing that is not unintentional. Further the results suggest that the feedforward activation from the orthographic lexicon to the phonological lexicon are either stopped completely or functionally postponed and are therefore not ballistic.

Based on the pattern of interference observed in the present study, I argue that processing up to and including the activation of representations in the orthographic lexicon is automatic in the sense that it is unintentional but that processing beyond the orthographic lexicon is not, in the sense that it either uses intention or occurs after unspecified processes that use intention. The present findings are therefore inconsistent with the widespread claim that reading occurs automatically in the sense that it is unintentional and ballistic.

The Present Study's Findings in the Context of Dual Task Research

The results of the present experiment suggest that processing up to and including the orthographic input lexicon comprise early processing that is unintentional and ballistic. Processes that occur after the activation of representations in the orthographic input lexicon, particularly the activation of phonological representations, comprise later processing that is not unintentional and is not ballistic (i.e., it can either be stopped or postponed). These conclusions are remarkably consistent with studies that have examined reading performance in dual task scenarios, in particular studies using the Psychological Refractory Period (PRP) paradigm (e.g., McCann et al., 2000; Reynolds & Besner, 2006; Ruthruff et al., 2008). In the PRP paradigm, participants are presented with two stimuli in succession with the time between presentation of the first and second stimulus (i.e., the stimulus onset asynchrony; SOA) being varied. Participants are asked to complete two tasks, the first task is usually a choice reaction time task (e.g., tone identification high vs. low) and the second task uses word recognition (e.g., reading aloud task). The standard finding in the PRP task is that as the stimulus onset asynchrony decreases, the response time for task 2 increases (e.g., Welford, 1952; Telford, 1931) suggesting that there is a limited capacity to information processing (Tombu & Joliceur, 2003; 2004; see also Pashler, 1994, for a review). This limitation is theorized to be due to a central bottleneck in information processing where both tasks require access to the same processing system, when the system can only act on one task at a time, resulting in the processing of one task being postponed until processing of the other is completed (e.g., Pashler, 1994).

The PRP paradigm has been used to assess whether skilled reading occurs automatically, in the sense that it occurs independently of central attention and other processing limitations (e.g., McCann & Johnston, 1992; McCann et al., 2000; O'Malley et al., 2008; Reynolds & Besner, 2006). These examinations focus on the assumption that automatic processes should not have a limited processing capacity (Cohen et al.,1992; McCann et al., 2000; Pashler, 1994; Hasher & Zacks, 1979; Brown et al., 2002; Miles et al., 2019; Reynolds & Besner, 2006).

There is a very strong correspondence between the presence / absence of word processing effects in the colour naming task in the current study, and processes that do / do not require the processing bottleneck in studies using the PRP paradigm.

1. Evidence from colour naming studies suggests that the processes underlying the repetition priming effect occur without intention (e.g., Burt, 2002) and evidence from the PRP paradigm suggest that the processes responsible for repetition priming do not require the processing bottleneck (Reynolds & Besner, 2006).
2. Evidence from the present study and other non-colour word Stroop studies suggests that at least some of the processes underlying the word frequency effect occur without intention (e.g., Monsell et al., 2001) and evidence from the PRP paradigm suggest that at least some of the processes responsible for the word frequency do not require

the processing bottleneck (Rabovstky et al., 2008; 2019; Reynolds & Besner, 2006, but see McCann et al., 2000).

3. Evidence from the present non-colour word Stroop study suggests that the processes that give rise to neighbourhood density effects require intention and evidence from the PRP paradigm suggests that the processes responsible for neighbourhood density effects either use or occur after the processing bottleneck (Reynolds & Besner, 2006, Experiments 5-7).
4. Finally, evidence from the present non-colour word Stroop study suggests that lexicality effects require intention and evidence from the PRP paradigm suggests that these processes require the processing bottleneck (McCann et al., 2000; Rabovsky et al., 2008; 2019; Reynolds & Besner, 2006).

The fact that the findings from the non-colour word Stroop task and from studies using the PRP paradigm converge is of theoretical interest. Based on the current findings, processing up to and including the orthographic lexicon appears to be automatic in the sense that it occurs unintentionally and ballistically. Similarly, findings from PRP studies suggest that early processing occurring in the orthographic lexicon is automatic in the sense that it does not require central attention (Johnston et al., 1995), which is often described as a response selection bottleneck (Pashler, 1994). Further, the present findings suggest that processing between the orthographic and phonological lexicons and at the level of the phoneme buffer is not automatic in the sense that it is not unintentional and can be functionally postponed (i.e., is not ballistic). Findings from PRP studies have drawn similar conclusions to the present study, indicating that processing beyond the orthographic lexicon, at the level of the phonological lexicon and phoneme buffer, are not automatic in the sense that attention is required, is limited capacity and can be functionally postponed (Reynolds & Besner, 2006).

This correspondence between the present study and PRP studies suggests that the two qualities of automaticity, (1) intentionality, and (2) attention/ capacity, co-occur in the reading system during processing up to and including the orthographic lexicon before being functionally postponed. This also indicates that the conclusions drawn in the present study regarding the processes involved in reading are generalizable to studies using a different paradigm to investigate the nature of the underlying processes involved in reading.

The Present Study's Findings in the Context of Previous Picture Naming Research

If the dual-task account of colour naming performance is true, then similar patterns of effects should be observed in other Stroop-like tasks. One task that is considered by many to be very similar to the non-colour word Stroop task is the picture-word interference (PWI) task (Hentschel, 1973; Rosinski et al., 1975; Ehri, 1977; Starreveld & La Heij, 2017). In the picture-word interference task, participants are presented with a written distractor word superimposed on a picture of an object and asked to name the object depicted in the picture while ignoring the distractor word. Similar to the Stroop task, a congruency effect is observed in the picture-word interference task. Incongruent words (e.g., the word "cat" superimposed on a line drawing of a pig) printed inside pictures generate significant interference in picture naming relative to words with a congruent meaning (e.g., the word "cow" superimposed on a line drawing of a cow, Rosinski et al., 1975).

The similarity between picture-word interference tasks and the Stroop colour-word interference tasks has been emphasized in the literature with some researchers arguing that they are the same effect. For example, Starreveld and La Heij (2017) argued that "picture–word interference is a Stroop effect" (p. 721). In reading research, the results from these two tasks are sometimes used interchangeably (e.g., Mulatti & Coltheart, 2014; Roelofs, 2003; Roelofs & Piai, 2013). However, some researchers have argued that the two tasks differ in the role of semantics because colours do not have privileged access to semantic information, unlike pictures (De Houwer et al., 1994). Despite this, parallels can still be drawn between the two tasks with regards to the nature of the cognitive processes involved in the tasks and the locus of interference. The PWI task is a promising avenue for examining how the presence of an irrelevant word can interfere with ongoing task performance.

Despite the structural similarities of the picture-word interference and Stroop colour-word interference tasks, some differences do exist. A major difference between tasks is the number of possible target stimuli. The classic Stroop colour word interference task has only a few target colour stimuli whereas the picture word interference task can use many picture stimuli. This is in fact a major advantage of the PWI task. For example, in the PWI task, the relation between the picture and word can be manipulated with greater flexibility compared to in the Stroop colour-word interference task (De Houwer et al., 1994).

The number of targets can affect at least four relevant processes (1) response-set, (2) perceived relevance, (3) response-selection, and (4) the source of interference. First, the small number of targets and responses in the Stroop colour-word interference task creates a clear response set often consisting of between 2 and 4 repeating items (e.g., “blue”, “yellow”, “red”, etc.), whereas the large number of targets and responses in the PWI task does not create a clear response set (De Houwer et al., 1994). Second, the small number of targets and large repetition can make it easier to perceive a relationship between the colours and the words, thereby making the words seem more relevant. Third, the repetition of target stimuli may also shift response selection earlier (or the lack of repetition may shift response selection later in the system) – if a limited number of stimuli repeat, then response selection can be based on physical features instead of a detailed semantic analysis. Finally, a potential difference between the Stroop task and the picture-word interference task is the definition of congruency. In the Stroop task incongruent items are in the same semantic category – colour related words, whereas in the PWI incongruent items are not necessarily in the same category.

Despite the differences between the picture-word interference task and the non-colour word Stroop task similar effects have been observed. Numerous studies have examined picture naming with low- or high-frequency superimposed distractor words. These studies typically report word frequency effects similar those obtained in the present study (Bates et al., 2001; Dhooge & Hartsuiker, 2010, 2011; Miozzo & Caramazza; 2003; Mousikou & Rastle, 2015). This outcome is consistent with words unintentionally activating representations in the orthographic input lexicon.

Unlike the non-colour word Stroop task, Lupker (1982) reported a lexicality effect in which pictures with a non-word superimposed on to it were responded to faster than pictures with a word superimposed on it. This lexicality effect is in the opposite direction to what is observed in word reading, and inconsistent with the present study's colour-naming results in which a null effect of lexicality was observed. However, Lupker (1982) had a limited stimulus set consisting of 12 items that repeated and a semantic relationship between the words and the pictures. More recent studies investigating the effect of lexicality on picture naming performance in the PWI task have also reported that pictures with a pseudoword distractor were named faster than those with a word distractor (Dhooge & Hartsuiker, 2012, Experiment 1). More research is needed to better understand the nature of the lexicality effect in the PWI task. Further, the neighbourhood effect has not yet been explored in the PWI task and provides an avenue for extending the present research.

Future Directions

Two ways that the present research could be extended is by (1) examining other psycholinguistic variables in the context of the colour-word interference task, and (2) searching for parallels between the colour-word interference task and the picture-word interference task.

Future research should examine the locus of control in the non-colour word Stroop task. This can be done by examining whether colour naming is affected by other psycholinguistic variables. The absence of neighbourhood density and lexicality effects suggests that phonological representations are not being activated. Additional converging evidence for this claim would come from looking at the regularity effect (Reynolds & Besner, 2004) which is widely believed to arise at the phoneme system from conflict between the output of the lexical and sublexical routes (Coltheart et al., 2001). Based on the present study's finding of a null effect of lexicality in the colour naming data, it is not anticipated that an effect of regularity would emerge. This prediction could be examined in a replication of the present experiment with the novel addition of regularity as one of the independent variables. If an effect of regularity were to emerge, this would be largely inconsistent with the present findings, which suggest that unintentional processing is not occurring at the level of the phoneme buffer and would be at odds with our finding of a null lexicality effect.

Another study should examine the sublexical computation of phonology. The present study suggests that activation from the orthographic lexicon is not passed on to the phonological lexicon. However, the sublexical computation of phonology starts prior to the orthographic lexicon (see Figure 3, Panel A). A well-established finding is that the time to read a nonword (but not a word) increases linearly with letter length (Coltheart et al., 2001). Therefore, future research should examine whether non-word letter length affects colour-naming performance. These variables could potentially identify the locus of interference in colour naming more precisely and provide further insight into nature of the processing involved in skilled reading.

The picture-word interference task provides another useful way to study whether unintentional word processing (processing up to the orthographic lexicon) and intentional word processing (processing after the orthographic lexicon) is affected by (1) response-set and (2) perceived relevance, (3) response-selection, and (4) the source of interference. Future research should examine how the effects of word frequency, lexicality, and neighbourhood density are affected by the number of target-stimuli. Unintentional processes should affect performance irrespective of the number of unique targets, whereas processes affected by intention should be affected by the number of unique targets.

Some researchers believe that pictures activate semantics differently than colours do, and that in this respect, the PWI task is primarily a semantic task (e.g., De Houwer et al., 1994). Consequently, this means that the distractor word is more likely to be read in the PWI task than in a colour-word interference task because the perceived relevance of the distractor word is greater. To the present study's current knowledge, no study has yet been conducted in which a PWI experiment is designed to make the distractor word as irrelevant as possible to reduce the likelihood of it being read. If an experiment were constructed in such a way, in which the perceived relevance of the distractor word is greatly reduced, based on the present study's findings, it would be expected that the same pattern of effects would emerge.

A final interesting avenue for future investigation would be to examine the effect of homogenous stimuli on performance in the non-colour word Stroop task, to better understand how the homogeneity of stimuli can alter or affect the pattern of effects produced. One way this could be done would be to replicate the present experiment with the addition of another stimulus set, one containing heterogenous stimuli. Stimuli that vary in letter length (e.g., heterogenous stimuli such as words and non-words ranging in letter length from 4 to 10 letters long) could create a signal to pay attention to them, whereas stimuli of the same or close to the same letter length (e.g., as in the present study) do not. Therefore, the inclusion of less homogenous stimuli could test the prediction that homogeneity of stimuli drives the perceived relevance of the distractor word down and makes it less likely they will be read, whereas heterogeneity of stimuli drives this perceived relevance up and increases the probability that the distractor word will be read and that effects (e.g., word frequency, lexicality) will be produced.

Conclusion

The findings from the present experiment are consistent with two conclusions. First, some, but not all, word reading processes are automatic in the sense that they proceed without intention. The present study suggests that processing up to the activation of representations in the orthographic lexicon proceed without intention. Second, the non-colour word Stroop task can be conceptualized as a dual-task scenario in which the stimuli associated with the irrelevant task (word reading) are given priority processing and that priority cannot switch to the relevant task (colour naming) until a certain point of processing is reached. Future research needs to test the boundaries of these two claims by examining other psycholinguistic variables in the non-colour word Stroop task and by assessing whether the dual task account applies in other similar tasks (e.g., picture word interference).

References

Almack B. (2013). Research Shows Reading Levels Predict Future Income. *Dyslexia, General, Reading Fluency.* https://onlinereadingtutor.com/research-shows-reading-levels-predict-future-income/

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(5), 802–814. https://doi.org/10.1037/0278-7393.15.5.802

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy?. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(2), 23. https://doi.org/10.1037/0278-7393.18.2.234

Andrews, S. (1996). Lexical retrieval and selection processes: Effects of transposed-letter confusability. *Journal of Memory and Language*, *35*(6), 775-800. https://doi.org/10.1006/jmla.1996.0040

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review, 4*(4), 439–461. https://doi.org/10.3758/BF03214334

Ashcraft. M. H. (1994). Human memory and cognition. Harper Collins.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *Psychology of learning and motivation*, *2*, 89-195. https://doi.org/10.1016/S0079-7421(08)60422-3

Augustinova, M., Flaudias, V., & Ferrand, L. (2010). Single-letter coloring and spatial cuing do not eliminate or reduce a semantic contribution to the Stroop effect. *Psychonomic Bulletin & Review*, *17*, 827-833. https://doi.org/10.3758/PBR.17.6.827

Augustinova, M., & Ferrand, L. (2012). Suggestion does not de-automatize word reading: Evidence from the semantically based Stroop task. *Psychonomic Bulletin & Review*, *19*, 521-527. https://doi.org/10.3758/s13423-012-0217-y

Augustinova, M., & Ferrand, L. (2014). Automaticity of word reading: Evidence from the semantic Stroop paradigm. *Current Directions in Psychological Science*, *23*(5), 343-348. https://doi.org/10.1177/0963721414540

Bakan, P., & Alperson, B.(1967). Pronounceability, intensity, and interference in the color-word test. *American Journal of Psychology, 80*,416-420. doi: 10.2307/1420375.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual Word Recognition of Single-Syllable Words. *Journal of Experimental Psychology: General, 133*(2), 283–316. https://doi.org/10.1037/0096-3445.133.2.283

Bar-Anan, Y., Liberman, N., Trope, Y., & Algom, D. (2007). Automatic processing of psychological distance: Evidence from a Stroop task. *Journal of Experimental Psychology: General, 136*(4), 610–622. https://doi.org/10.1037/0096-3445.136.4.610

Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 3–51). New York: Guilford Press.

Bargh, J. A. (1992). The ecology of automaticity: Toward establishing the conditions needed to produce automatic processing effects. *American Journal of Psychology, 105*, 181–199. https://doi.org/10.2307/1423027

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition, 1* (pp. 1–40). Hillsdale, NJ: Erlbaum.

Bates, E., Burani, C., D'Amico, S., & Barca, L. (2001). Word reading and picture naming in Italian. *Memory & Cognition, 29*(7), 986–999. https://doi.org/10.3758/BF03195761

Bauer, B., & Besner, D. (1997). Processing in the Stroop task: Mental set as a determinant of performance. *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale, 51*(1), 61–68. https://doi.org/10.1037/1196-1961.51.1.61

Besner, D. (2001) The myth of ballistic processing: Evidence from Stroop's paradigm. *Psychonomic Bulletin & Review, 8*, 324–330. https://doi.org/10.3758/BF03196168

Besner D., Stolz J. A.,& Boutilier C. (1997). The stroop effect and the myth of automaticity. *Psychonomic Bulletin and Review, 4* (2), 221-5. https://doi: 10.3758/BF03209396.

Besner, D., & McCann, R. S. (1987). Word frequency and pattern distortion in visual word identification and production: An examination of four classes of models. In M. Coltheart (Ed.), *Attention and performance 12: The psychology of reading* (pp. 201–219). Lawrence Erlbaum Associates, Inc.

Besner, D., & Stolz, J. A. (1999). Unconsciously controlled processing: The Stroop effect reconsidered. *Psychonomic Bulletin & Review, 6*(3), 449–455. https://doi.org/10.3758/BF03210834

Besner, D., & Stolz, J. A. (2001). The Stroop effect and single letter coloring: What replicates and what doesn't. *Psychonomic Bulletin & Review, 8*(4), 858. https://doi.org/10.3758/BF03196229

Bigozzi, L., Tarchi, C., Vagnoli, L., Valente, E., & Pinto, G. (2017). Reading fluency as a predictor of school outcomes across grades 4–9. *Frontiers in psychology, 8,* 200. https://doi.org/10.3389/fpsyg.2017.00200

Boiché, J., Marchant, G., Nicaise, V., & Bison, A. (2016). Development of the generic multifaceted automaticity scale (GMAS) and preliminary validation for physical activity. *Psychology of Sport and Exercise, 25*, 60–67. https://doi.org/10.1016/j.psychsport.2016.03.003

Brysbaert, M. (2022). Chapter "Word recognition II: Orthography-phonology" Phonological Coding. *Visual Language Processing*.

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*(4), 977-990. https://doi.org/10.3758/BRM.41.4.977

Brown, T. L., Joneleit, K., Robinson, C. S., & Brown, C. R. (2002). Automaticity in reading and the Stroop task: testing the limits of involuntary word processing. *The American journal of psychology, 115*(4), 515–543. https://doi.org/10.2307/1423526

Bugg, J. M., Jacoby, L. L., & Toth, J. P. (2008). Multiple levels of control in the Stroop task. *Memory & cognition, 36*(8), 1484–1494. https://doi.org/10.3758/MC.36.8.1484

Burt, J. S. (1994). Identity primes produce facilitation in a colour naming task. *The Quarterly Journal of Experimental Psychology*, *47*(4), 957-1000. https://doi.org/10.1080/14640749408401103

Burt, J.S. (1999). Associative priming in color naming: interference and facilitation. *Memory & Cognition, 27*, 454–464. https://doi.org/10.3758/BF03211540

Burt, J.S., (2002). Why do non-color words interfere with color naming? *Journal of*

Experimental Psychology: Human Perception and Performance, *28*, 1019–1038. https://doi.org/10.1037/0096-1523.28.5.1019

Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of orthographic neighborhood in visual word recognition: cross-task comparisons. *Journal of experimental psychology. Learning, memory, and cognition*, *23*(4), 857–871. https://doi.org/10.1037//0278-7393.23.4.857

Cattell, J. M. (1886). The time it takes to see and name objects. *Mind, 11*, 63-65. http://www.jstor.org/stable/2247157.

Clark, H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychology research. *Journal of Verbal Learning and Verbal Behavior, 12*, 335-359. https://doi.org/10.1016/S0022-5371(73)80014-3

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

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159. doi:10.1037/0033-2909.112.1.155

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. *Attention and performance, 6,* 535–555. https://doi.org/10.1037/0096-1523.25.2.461

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

Coltheart M, Rastle K, Perry C, Langdon R, and Ziegler J (2001). The DRC model: A model of visual word recognition and reading aloud. *Psychological Review.108*, 204–258. https://doi.org/10.1037/0033-295X.108.1.204

Dalrymple-Alford, E. C., & Budayr, B. (1966). Examination of some aspects of the Stroop Color-Word Test. *Perceptual and Motor Skills, 23* (3.2), 1211–1214. https://doi.org/10.2466/pms.1966.23.3f.1211

den Heyer, K., Goring, A., Gorgichuk, S., Richards, L., & Landry, M. (1988). Are lexical decisions a good measure of lexical access? Repetition blocking suggests the affirmative. *Canadian Journal of Psychology / Revue canadienne de psychologie, 42*(3), 274–296. https://doi.org/10.1037/h0084192

De Houwer, J., Fias, W., & d'Ydewalle, G. (1994). Comparing color-word and picture-word Stroop-like effects: A test of the Glaser and Glaser (1989) model. *Psychological Research, 56*(4), 293–300. https://doi.org/10.1007/BF00419659

De Houwer, J. (2003). On the role of stimulus-response and stimulus-stimulus compatibility in the Stroop effect. *Memory & Cognition, 31*(3), 353–359. https://doi.org/10.3758/BF03194393

Dhooge, E., & Hartsuiker, R. J. (2010). The distractor frequency effect in picture–word interference: Evidence for response exclusion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(4), 878–891. https://doi.org/10.1037/a0019128

Dhooge, E., & Hartsuiker, R. J. (2011). How do speakers resist distraction?: Evidence from a taboo picture-word interference task. *Psychological Science, 22*(7), 855–859. https://doi.org/10.1177/0956797611410984

Dhooge, E., & Hartsuiker, R. J. (2012). Lexical selection and verbal self-monitoring: Effects of lexicality, context, and time pressure in picture-word interference. *Journal of Memory and Language, 66*(1), 163–176. https://doi.org/10.1016/j.jml.2011.08.004

Dijksterhuis, A. (2010). Automaticity. In D. E. Carlston (Ed.), *The Oxford handbook of social Cognition* (pp. 239–256). Oxford University Press.

Dyer, F. N. (1973). Interference and facilitation for color naming with separate bilateral presentations of the word and color. *Journal of Experimental Psychology, 99*(3), 314–317. https://doi.org/10.1037/h0035245

Ehri, L. C. (1977). Do adjectives and functors interfere as much as nouns in naming pictures? *Child Development, 48*(2), 697–701. https://doi.org/10.2307/1128678

Ehri, L. C. (2005). Development of Sight Word Reading: Phases and Findings. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 135–154). Blackwell Publishing. https://doi.org/10.1002/9780470757642.

Ferrand, L., & Grainger, J. (1992). Phonology and orthography in visual word recognition: Evidence from masked non-word priming. The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 45A(3), 353–372. https://doi.org/10.1080/02724989208250619

Flaudias, V., & Llorca, P.-M. (2014). A brief review of three manipulations of the Stroop task focusing on the automaticity of semantic access. *Psychologica Belgica, 54*(2), 199–221. https://doi.org/10.5334/pb.am

Frost R. (1998). Toward a strong phonological theory of visual word recognition: true issues and false trails. *Psychological bulletin, 123*(1), 71–99. https://doi.org/10.1037/0033-2909.123.1.71

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal*

Learning & Verbal Behavior, 12(6), 627–635. https://doi.org/10.1016/S0022-5371(73)80042-8

Forster, K. I. (1976). Accessing the mental lexicon. *New approaches to language mechanisms*, 257-287.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(4), 680–698. https://doi.org/10.1037/0278-7393.10.4.680

Fox, L. A., Shor, R. E., & Steinman, R. J. (1971). Semantic gradients and interference in naming color, spatial direction, and numerosity. *Journal of Experimental Psychology, 91*(1), 59–65. https://doi.org/10.1037/h0031850

Glaser, W. R., & Dolt, M. O. (1977). A functional model to localize the conflict underlying the Stroop phenomenon. *Psychological Research, 39*(4), 287–310. https://doi.org/10.1007/BF00308930

Glaser, M. O., & Glaser, W. R. (1982). Time course analysis of the Stroop phenomenon. *Journal of Experimental Psychology: Human Perception and Performance, 8*(6), 875–894. https://doi.org/10.1037/0096-1523.8.6.875

Gold, J., & Ciorciari, J. (2020). A Review on the Role of the Neuroscience of Flow States in the Modern World. *Behavioral sciences (Basel, Switzerland), 10*(9), 137. https://doi.org/10.3390/bs10090137

Grainger, J., & Holcomb, P. J. (2010). Neural constraints on a functional architecture for word recognition. *The neural basis of reading* (pp. 3-32). https://doi.org/10.4074/S0003503307004058

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language, 29*(2), 228-244. https://doi.org/10.1016/0749-596X(90)90074-A

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103*(3), 518–565. https://doi.org/10.1037/0033-295X.103.3.518

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology, 97*(1), 17–41. https://doi.org/10.1037/a0015575

Guttentag, R. E., & Haith, M. M. (1978). Automatic processing as a function of age and reading ability. *Child Development* (pp. 707-716).

Hanushek, E. A., Schwerdt G., Wiederhold S., & Woessmann, L. (2013). Returns to skills around the world: Evidence from PIAAC. *National Bureau of Economic Research (NBER)*, working paper,19762:3. https://EconPapers.repec.org/RePEc:nbr:nberwo:19762

Happe, F. (2001). Automaticity of Action, *Psychology of. International Encyclopedia of the Social & Behavioral Sciences*; Smelser, NJ, Baltes, PB, Eds.

Harm, M. W., & Seidenberg, M. S. (1996). Computational bases of two types of developmental dyslexia. *Proceedings of the eighteenth annual conference of the cognitive science society* (pp. 364-369). Routledge.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review, 106*(3), 491–528. https://doi.org/10.1037/0033-295X.106.3.491

Harm, M. W., & Seidenberg, M. S. (2004). Computing the Meanings of Words in Reading:

Cooperative Division of Labor Between Visual and Phonological Processes. *Psychological Review, 111*(3), 662–720. https://doi.org/10.1037/0033-295X.111.3.662

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108*(3), 356–388. https://doi.org/10.1037/0096-3445.108.3.356

Hentschel, U. (1973). *Two new interference tests compared to the Stroop Color-Word Test*. Lund University.

Herbers J. E., Cutuli J. J., Supkoff L. M., Heistad D., Chan C., Hinz E., & Masten A. S. (2012). Early Reading Skills and Academic Achievement Trajectories of Students Facing Poverty, Homelessness, and High Residential Mobility. Sage Journals. https://doi.org/10.3102/0013189X12445320

Huntsman, L. A., & Lima, S. D. (2002). Orthographic neighbors and visual word recognition. *Journal of Psycholinguistic Research, 31*(3), 289–306. https://doi.org/10.1023/A:1015544213366

Ivanova, I. (2011). The Cost of Poverty in BC. *Centre for Policy Alternatives, 14.* https://www.policyalternatives.ca/sites/default/files/uploads/publications/BC%20Office/2011/07/CCPA_BC_cost_of_poverty_full_report.pdf

James, W. (1890). *The principles of psychology, Vol. 1*. Henry Holt and Co. https://doi.org/10.1037/10538-000

Jamieson, D. G. (2006). Literacy in Canada. *Paediatrics & child health*, 11(9), 573-574. https://doi.org/10.1093/pch/11.9.573

Johnston, J. C., McCann, R. S., & Remington, R. W. (1995). Chronometric evidence for two types of attention. *Psychological Science, 6*(6), 365–369. https://doi.org/10.1111/j.1467-9280.1995.tb00527.x

Joo, S. J., Tavabi, K., Caffarra, S., & Yeatman, J. D. (2021). Automaticity in the reading circuitry. *Brain and Language, 214*, 104906. https://doi.org/10.1016/j.bandl.2020.104906

Kahneman, D., & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy & J. R. Pomerantz (eds.),*Perceptual Organization* (pp. 181–211). Hillsdale, nj: erlbaum.

Kahneman, D., & Chajczyk, D. (1983). Tests of the automaticity of reading: dilution of Stroop effects by color-irrelevant stimuli. *Journal of experimental psychology. Human perception and performance, 9*(4), 497–509. https://doi.org/10.1037//0096-1523.9.4.497

Kinoshita, S., Lupker, S. J., & Rastle, K. (2004). Modulation of regularity and lexicality effects in reading aloud. *Memory & Cognition, 32*(8), 1255–1264. https://doi.org/10.3758/BF03206316

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. http://dx.doi.org/10.1037/xlm0000311

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://www.jstor.org/stable/1420768

Kuhn, M. R., Schwanenflugel, P. J., & Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Reading research quarterly, 45*(2), 230-251. https://.doi.org/10.1598/RRQ.45.2.4

Lawrence, M. A. (2016). Ez: Easy analysis and visualization of factorial experiments.

https://CRAN.R-project.org/ package=ez

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*(2), 293–323. https://doi.org/10.1016/0010-0285(74)90015-2

Landauer, T. K., & Streeter, L. A. (1973). Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Verbal Learning & Verbal Behavior, 12*(2), 119–131. https://doi.org/10.1016/S0022-5371(73)80001-5

Laxon, V. J., Masterson, J., Pool, M., & Keating, C. (1992). Nonword naming: Further exploration of the pseudohomophone effect in terms of orthographic neighborhood size, graphemic changes, spelling-sound consistency, and reader accuracy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(4), 730–748. https://doi.org/10.1037/0278-7393.18.4.730

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet physics doklady, 10* (8), 707-710.

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

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*(4), 492–527. https://doi.org/10.1037/0033-295X.95.4.492

Logan, G. D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading & writing quarterly: Overcoming learning difficulties, 13*(2), 123-146.

Lupker, S. J. (1982). The role of phonetic and orthographic similarity in picture–word interference. *Canadian Journal of Psychology / Revue canadienne de psychologie, 36*(3), 349–367. https://doi.org/10.1037/h0080652

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

Manwell, L.A., Roberts, M.A., & Besner, D. (2004). Single letter coloring and spatial cuing eliminates a semantic contribution to the Stroop effect. *Psychonomic Bulletin & Review, 11*, 458-462. https://doi.org/10.3758/BF03196595

Marshall, J. C. & Newcombe, F. (1973). Patterns of paralexia: A psycholinguistic approach. *Journal of Psycholinguistic Research, 2*, 175–199.

McCann, R., & Besner, D. (1987). Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word frequency effects in naming. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 14–24. https://doi.org/10.1037/0096-1523.13.1.14

McCann, R. S., Besner, D., & Davelaar, E. (1988). Word recognition and identification: Do word-frequency effects reflect lexical access? *Journal of Experimental Psychology: Human Perception and Performance, 14*(4), 693–706. https://doi.org/10.1037/0096-1523.14.4.693

McCann, R. S., Folk, C. L., & Johnston, J. C. (1992). The role of spatial attention in visual word processing. *Journal of Experimental Psychology: Human Perception and Performance, 18*(4), 1015–1029. https://doi.org/10.1037/0096-1523.18.4.1015

McCann, R. S., Remington, R. W., & Van Selst, M. (2000). A dual-task investigation of automaticity in visual word processing. *Journal of Experimental Psychology: Human Perception and Performance, 26*(4), 1352–1370. https://doi.org/10.1037/0096-1523.26.4.1352

McClelland, J. L. (1993). Toward a theory of information processing in graded, random, and interactive networks. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance 14: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 655–688). The MIT Press.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88*(5), 375–407. https://doi.org/10.1037/0033-295X.88.5.375

Miles, A., Charron-Chénier, R., & Schleifer, C. (2019). Measuring automatic cognition: Advancing dual-process research in sociology. *American Sociological Review, 84*(2), 308-333. https://doi.org/10.1177/0003122419832497

Mills, L., Kinoshita, S., & Norris, D. (2019). No Negative Priming Effect in the Manual Stroop Task. *Frontiers in psychology, 10*, 1764. https://doi.org/10.3389/fpsyg.2019.01764

Miozzo, M., & Caramazza, A. (2003). When more is less: a counterintuitive effect of distractor frequency in the picture-word interference paradigm. *Journal of experimental psychology. General*, *132*(2), 228–252. https://doi.org/10.1037/0096-3445.132.2.228

Monahan, J. S. (2001). Coloring single Stroop elements: Reducing automaticity or slowing color processing?. *The Journal of general psychology, 128*(1), 98-112. https://doi.org/10.1080/00221300109598901

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General, 118*(1), 43–71. https://doi.org/10.1037/0096-3445.118.1.43

Monsell, S., Taylor, T.J. & Murphy, K. (2001). Naming the color of a word: Is it responses or task sets that compete?. *Memory & Cognition, 29*, 137–151. https://doi.org/10.3758/BF03195748

Moors, A., & De Houwer, J. (2006). Automaticity: a theoretical and conceptual analysis. *Psychological bulletin, 132*(2), 297. https://doi.org/10.1037/0033-2909.132.2.297

Morey, R. D., Rouder, J. N., & Jamil, T. (2018). BayesFactor: Computation of Bayes factors for common designs. R package version 0.9. 12-4.2. https://cran.r-project.org/web/packages/BayesFactor/BayesFactor.pdf

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76*(2), 165–178. https://doi.org/10.1037/h0027366

Morton, J. & Patterson, K. E. (1980) A new attempt at an interpretation or an attempt at a new interpretation. In M. Coltheart, K. E. Patterson & J. C. Marshall (Eds) *Deep dyslexia*, London: Routledge and Kegan Paul.

Mousikou, P., & Rastle, K. (2015). Lexical frequency effects on articulation: A comparison of picture naming and reading aloud. *Frontiers in Psychology, 6,* Article 1571. https://doi.org/10.3389/fpsyg.2015.01571

Mulatti, C., & Coltheart, M. (2014). Color naming of colored non-color words and the response-exclusion hypothesis: A comment on Mahon et al. and on Roelofs and Piai. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 52,* 120–122. https://doi.org/10.1016/j.cortex.2013.08.018

Mulatti, C., Reynolds, M. G., & Besner, D. (2006). Neighborhood effects in reading aloud: New findings and new challenges for computational models. *Journal of Experimental Psychology: Human Perception and Performance*, 32(4), 799-810. https://doi.org/10.1037/0096-1523.32.4.799

National Forum on Literacy and Poverty.(n.d.). https://www.frontiercollegefoundation.ca/getattachment/6f6bd77f-043f-4ce3-96a0-9a2ec2902d73/Discussion-Paper-Literacy-and-Poverty-by-Frontier.aspx

Navarrete, E., Sessa, P., Peressotti, F., & Dell'Acqua, R. (2015). The distractor frequency effect in the colour-naming Stroop task: An overt naming event-related potential study. *Journal of Cognitive Psychology, 27*(3), 277-289. https://doi.org/10.1080/20445911.2014.1002786

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106*(3), 226–254. https://doi.org/10.1037/0096-3445.106.3.226

Neely, J. H., & Kahan, T. A. (2001). Is semantic activation automatic? A critical re-evaluation. In H. L. Roediger III, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 69–93). American Psychological Association. https://doi.org/10.1037/10394-005

Norton, E. S., & Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: implications for understanding and treatment of reading disabilities. *Annual review of psychology, 63*, 427–452. https://doi.org/10.1146/annurev-psych-120710-100431

O'Malley, S., Reynolds, M. G., Stolz, J. A., & Besner, D. (2008). Reading aloud: Spelling-sound translation uses central attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(2), 422–429. https://doi.org/10.1037/0278-7393.34.2.422

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*(2), 220–244. https://doi.org/10.1037/0033-2909.116.2.220

Peereman, R. (1995). Naming regular and exception words: Further examination of the effect of phonological dissension among lexical neighbours. *European Journal of Cognitive Psychology, 7*(3), 307–330. https://doi.org/10.1080/09541449508402451

Peereman, R., & Content, A. (1995). Neighborhood size effect in naming: Lexical activation or sublexical correspondences?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(2), 409. https://doi.org/10.1037/0278-7393.21.2.409

Peereman, R., & Content, A. (1997). Orthographic and phonological neighborhoods in naming: Not all neighbors are equally influential in orthographic space. *Journal of Memory and Language, 37*(3), 382–410. https://doi.org/10.1006/jmla.1997.2516

Perea, M. (2015). Neighborhood effects in visual word recognition and reading. In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 76–87). Oxford University Press.

Perfetti, C. A., & Zhang, S. (1995). Very early phonological activation in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 24–33. https://doi.org/10.1037/0278-7393.21.1.24

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*(1), 56–115. https://doi.org/10.1037/0033-295X.103.1.56

Plourde, C. E., & Besner, D. (1997). On the locus of the word frequency effect in visual word recognition. *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale, 51*(3), 181–194. https://doi.org/10.1037/1196-1961.51.3.181

Posner, M. I. (1978). Chronometric explorations of mind. Hillsdale, NJ: Erlbaum.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium.* Lawrence Erlbaum, 392-409.

Pruessner, L., Barnow, S., Holt, D. V., Joormann, J., & Schulze, K. (2020). A cognitive control framework for understanding emotion regulation flexibility. *Emotion, 20*(1), 21–29. https://doi.org/10.1037/emo0000658

R Core Team (2022). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. https:// www.R-project.org/

Rabovsky, M., Álvarez, C. J., Hohlfeld, A., & Sommer, W. (2008). Is lexical access autonomous? Evidence from combining overlapping tasks with recording event-related brain potentials. *Brain Research, 1222*, 156–165. https://doi.org/10.1016/j.brainres.2008.05.066

Rabovsky, M., Conrad, M., Álvarez, C. J., Paschke-Goldt, J., & Sommer, W. (2019). Attentional modulation of orthographic neighborhood effects during reading: Evidence from event-related brain potentials in a psychological refractory period paradigm. *PLoS ONE, 14*(1), Article e0199084. https://doi.org/10.1371/journal.pone.0199084

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC nonword database. *The Quarterly Journal of Experimental Psychology: Section A, 55*(4), 1339-1362. https://doi.org/10.1080/02724980244000099

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Prentice-Hall, Inc.

Raz, A., Shapiro, T., Fan, J., & Posner, M. I. (2002). Hypnotic suggestion and the modulation of

Stroop interference. *Archives of General Psychiatry, 59*(12), 1155–1161. https://doi.org/10.1001/archpsyc.59.12.1155

Raz, A., Landzberg, K. S., Schweizer, H. R., Zephrani, Z. R., Shapiro, T., Fan, J., & Posner, M. I. (2003). Posthypnotic suggestion and the modulation of Stroop interference under cycloplegia. *Consciousness and Cognition: An International Journal, 12*(3), 332–346. https://doi.org/10.1016/S1053-8100(03)00024-2

Raz, A., Kirsch, I., Pollard, J., & Nitkin-Kaner, Y. (2006). Suggestion Reduces the Stroop Effect. *Psychological Science, 17*(2), 91–95. https://doi.org/10.1111/j.1467-9280.2006.01669.x

Redding, G. M., & Gerjets, D. A. (1977). Stroop effect: Interference and facilitation with verbal and manual responses. *Perceptual and Motor Skills, 45*(1), 11–17. https://doi.org/10.2466/pms.1977.45.1.11

Regan, J. (1978). Involuntary automatic processing in color-naming tasks. Perception & *Psychophysics, 24*(2), 130–136. https://doi.org/10.3758/BF03199539

Reynolds, M., Mulatti, C., Peressotti, F., Job, R., & Dell' Aqua, R. (n.d). A dual task account of non-color word interference on color naming. [Master's Thesis, Trent University, Canada, University of Padova, Italy, University of Trento, Italy] Università di Padova. http://colab.psy.unipd.it/wp-content/uploads/2017/02/jephpp_colnam_15.pdf

Reynolds, M., & Besner, D. (2002). Neighbourhood density effects in reading aloud: New insights from simulations with the DRC model. *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale, 56*(4), 310–318. https://doi.org/10.1037/h0087407

Reynolds, M., & Besner, D. (2004). Neighbourhood density, word frequency and spelling–sound regularity effects in naming: Similarities and differences between skilled readers and the dual route cascaded computational model. *Canadian Journal of Experimental Psychology*, 58, 13–31. https://doi.org/10.1037/h0087437

Reynolds, M., & Besner, D. (2005). Contextual control of lexical and sublexical routines when reading English aloud. *Psychonomic Bulletin & Review*, *12*, 113–118. https://doi.org/10.3758/BF03196355

Reynolds, M., & Besner, D. (2006). Reading aloud is not automatic: Processing capacity is required to generate a phonological code from print. *Journal of Experimental Psychology: Human Perception and Performance, 32*(6), 1303–1323. https://doi.org/10.1037/0096-1523.32.6.1303

Reynolds, M., & Besner, D. (2008). Contextual effects on reading aloud: Evidence for pathway control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(1), 50–64. https://doi.org/10.1037/0278-7393.34.1.50

Reynolds, M., Kwan, D., & Smilek, D. (2010). To group or not to group: An ecological consideration of the Stroop effect. *Experimental Psychology*, *57*(4), 275–291. https://doi.org/10.1027/1618-3169/a000033

Risko, E.F., Schmidt, J.R. & Besner, D. (2006). Filling a gap in the semantic gradient: Color associates and response set effects in the Stroop task. *Psychonomic Bulletin & Review, 13,* 310–315. https://doi.org/10.3758/BF03193849

Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review, 110*(1), 88–125. https://doi.org/10.1037/0033-295X.110.1.88

Roelofs, A., & Piai, V. (2013). Associative facilitation in the Stroop task: Comment on Mahon et al. (2012). *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 49*(6), 1767–1769. https://doi.org/10.1016/j.cortex.2013.03.001

Rosinski, R. R., Golinkoff, R. M., & Kukish, K. S. (1975). Automatic semantic processing in a picture-word interference task. *Child Development, 46*(1), 247–253. https://doi.org/10.2307/1128859

Rothermund, K., Gollnick, N., & Giesen, C. G. (2022). Accounting for Proportion Congruency Effects in the Stroop Task in a Confounded Setup: Retrieval of Stimulus-Response Episodes Explains it All. *Journal of cognition, 5*(1), 39. https://doi.org/10.5334/joc.232

Ruthruff, E., Allen, P. A., Lien, M. C., & Grabbe, J. (2008). Visual word recognition without central attention: evidence for greater automaticity with greater reading ability. *Psychonomic bulletin & review*, *15*(2), 337–343. https://doi.org/10.3758/pbr.15.2.337

Sahinoglu, B., & Dogan, G. (2016). Event-Related Potentials and the Stroop Effect. The *Eurasian journal of medicine, 48*(1), 53–57. https://doi.org/10.5152/eurasianjmed.2016.16012

Samuels J. S. & Flor R. F. (1997) The importance of automaticity for developing expertise in Reading. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 13*(2), 107-121. https://doi.org/10.1080/1057356970130202

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime (Version 2.0). *Computer software and manual*. Pittsburgh, PA: Psychology Software Tools Inc.

Schneider, W., & Chein, J. M. (2003). Controlled & automatic processing: Behavior, theory, and biological mechanisms. *Cognitive Science, 27*(3), 525–559. https://doi.org/10.1016/S0364-0213(03)00011-9

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological review, 84*(2), 127. https://doi.org/ 10.1037/0033-295X.84.2.127

Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21*(4), 876–900. https://doi.org/10.1037/0096-1523.21.4.876

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*(4), 523–568. https://doi.org/10.1037/0033-295X.96.4.523

Shiffrin, R. M., Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual leaming, automatic attending, and a general theory. *Psychological Review, 84*(2): 127-190. https://doi.org/10.1037/0033-295X.84.2.127

Sichel, J. L., & Chandler, K. A. (1969). The color-word interference test: The effects of varied color-word combinations upon verbal response latency. *The Journal of Psychology: Interdisciplinary and Applied, 72*(2), 219–231. https://doi.org/10.1080/00223980.1969.10543502

Sparapani, N., Connor, C. M., McLean, L., Wood, T., Toste, J., & Day, S. (2018). Direct and Reciprocal Effects among Social Skills, Vocabulary, and Reading Comprehension in First Grade. *Contemporary educational psychology, 53,* 159–167. https://doi.org/10.1016/j.cedpsych.2018.03.003

Stanovich, K. E. (1990). Concepts in developmental theories of reading skill: Cognitive resources, automaticity, and modularity. *Developmental Review, 10*(1), 72–100. https://doi.org/10.1016/0273-2297(90)90005-O

Starreveld, P. A., & La Heij, W. (2017). Picture-word interference is a Stroop effect: A theoretical analysis and new empirical findings. *Psychonomic Bulletin & Review, 24*(3), 721–733. https://doi.org/10.3758/s13423-016-1167-6

Stolz, J. A., & Besner, D. (1999). On the myth of automatic semantic activation in reading. *Current Directions in Psychological Science, 8*(2), 61–65. https://doi.org/10.1111/1467-8721.00015

Stolz, J. A., & McCann, R. S. (2000). Visual word recognition: Reattending to the role of spatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 26*(4), 1320–1331. https://doi.org/10.1037/0096-1523.26.4.1320

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

Tecce, J. J., & Dimartino, M. (1965). Effects of heightened drive (shock) on performance in a tachistoscopic color-word interference task. *Psychological Reports, 16*(1), 93–94. https://doi.org/10.2466/pr0.1965.16.1.93

Telford, C. W. (1931). Refractory phase of voluntary and associative response. *Journal of Experimental Psychology*, 14, 1-35.

Tombu, M., & Jolicoeur, P. (2003). A central capacity sharing model of dual-task performance. *Journal of experimental psychology. Human perception and performance*, *29*(1), 3–18. https://doi.org/10.1037//0096-1523.29.1.3

Tombu, M., & Jolicœur, P. (2004). Virtually No Evidence for Virtually Perfect Time-Sharing. *Journal of Experimental Psychology: Human Perception and Performance, 30*(5), 795–810. https://doi.org/10.1037/0096-1523.30.5.795

Toner, J., & Moran, A. (2021). Exploring the orthogonal relationship between controlled and automated processes in skilled action. *Review of Philosophy and Psychology, 12*(3), 577–593. https://doi.org/10.1007/s13164-020-00505-6

Treisman, A., Vieira, A., & Hayes, A. (1992). Automaticity and preattentive processing. *The American Journal of Psychology, 105*(2), 341–362. https://doi.org/10.2307/1423032

Tulkens, S., Sandra, D., & Daelemans, W. (2020). Orthographic Codes and the Neighborhood Effect: Lessons from Information Theory. *Proceedings of the 12th Language Resources and Evaluation Conference* (pp. 172-181).

Tzelgov, J., Porat, Z., & Henik, A. (1997). Automaticity and consciousness: Is perceiving the word necessary for reading it? *The American Journal of Psychology, 110*(3), 429–448. https://doi.org/10.2307/1423567

Underwood, G., & Batt, V. (2003). Reading and understanding an introduction to the psychology of reading. Blackwell.

Van Orden, G. C., & Goldinger, S. D. (1994). Interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception and Performance, 20*(6), 1269–1291. https://doi.org/10.1037/0096-1523.20.6.1269

Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, Section A, 47*(3), 631–650. https://doi.org/10.1080/14640749408401131

Ventura, P., Fernandes, T., Leite, I., Pereira, A., & Wong, A. C. (2019). Is holistic processing of written words modulated by phonology?. *Acta psychologica, 201*, 102944. https://doi.org/10.1016/j.actpsy.2019.102944

Watson, A. (2021). U.S. daily reading time by age 2020. *Statista*.

https://www.statista.com/statistics/412454/average-daily-time-reading-us-by-age/Wu

Welford, A. T. (1952). The psychological refractory period and the timing of high-speed performance-a review and a theory. *British Journal of Psychology*, 43(1), 2.

Wundt, W. (1903). *Naturwissenschaft Und Psychologie*. Kessinger Publishing

Wydell, T. N., Vuorinen, T., Helenius, P., & Salmelin, R. (2003). Neural Correlates of Letter-String Length and Lexicality during Reading in a Regular Orthography. Journal of Cognitive Neuroscience, 15(7), 1052–1062. https://doi.org/10.1162/089892903770007434

Xu, Y., Li, Y., Ding, W., & Lu, F. (2014). Controlled versus automatic processes: which is dominant to safety? The moderating effect of inhibitory control. *PloS one, 9*(2), e87881. https://doi.org/10.1371/journal.pone.0087881

Yarkoni, T., Balota, D. & Yap, M. (2008) Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review, 15*, 971–979. https://doi.org/10.3758/PBR.15.5.971

Zajano, M. J., & Gorman, A. (1986). Stroop interference as a function of percentage of congruent items. *Perceptual and Motor Skills, 63*(3), 1087–1096. https://doi.org/10.2466/pms.1986.63.3.1087

Zevin, J. D., & Balota, D. A. (2000). Priming and attentional control of lexical and sublexical pathways during naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(1), 121–135. https://doi.org/10.1037/0278-7393.26.1.121

Ziegler, J. C., Ferrand, L., & Montant, M. (2004). Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory & Cognition, 32*(5), 732–741. https://doi.org/10.3758/BF03195863

Zhang, S., Perfetti, C. A., & Yang, H. (1999). Whole word, frequency-general phonology in semantic processing of Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(4), 858–875. https://doi.org/10.1037/0278-7393.25.4.858

Zorzi, M., Houghton, G., & Butterworth, B. (1998). Two routes or one in reading aloud? A connectionist dual-process model. *Journal of Experimental Psychology: Human Perception and Performance, 24*(4), 1131–1161. https://doi.org/10.1037/0096-1523.24.4.1131

Appendix A. The complete set of ANOVA tables.

Please note that in the following ANOVA tables "cb" is used to represent the effect of counterbalance and a colon is used to represent an interaction between effects (e.g., a two-way interaction, "cb:wordFrequency" or a three-way interaction, "cb:wordFrequency:neighbourhoodDensity").

Table 1. Analysis of Variance by Subject of Stimulus Types based on Orthographic Neighbourhood Density (ON) in the Word-Naming Task (RT)

effect	DFn	DFd	SSn	SSd	*F*	*p*	*p<.05*	*ηp2*
(Intercept)	1	116	327838501	11484412.4	3311.38109	3.86E-87	*	0.96615492
cb	1	116	1725.356	11484412.4	0.01742721	8.95E-01		0.00015021
wordFrequency	2	232	1566342.02	1266126.2	143.50518	2.73E-41	*	0.55299544
neighbourhoodDensity	1	116	231851.043	279595.6	96.1915136	6.68E-17	*	0.45332401
cb:wordFrequency	2	232	2240.683	1266126.2	0.20528697	8.15E-01		0.00176659
cb:neighbourhoodDensity	1	116	4315.81	279595.6	1.79056452	1.83E-01		0.01520126
wordFrequency:neighbourhoodDensity	2	232	137928.792	270033.4	59.2509672	1.63E-21	*	0.33809209
cb:wordFrequency:neighbourhoodDensity	2	232	4165.068	270033.4	1.78921552	1.69E-01		0.01518998

Table 2. Analysis of Variance by Subject of High- and Low-Frequency Words based on Orthographic Neighbourhood Density (ON) in the Word-Naming Task (RT)

effect	DFn	DFd	SSn	SSd	*F*	*p*	*p<.05*	*ηp2*
(Intercept)	1	116	200346311	5411951.27	4294.23158	1.72E-93	*	0.97369753
cb	1	116	2326.5981	5411951.27	0.04986841	8.24E-01		0.00042972
wordFrequency	1	116	378034.174	426035.37	102.930336	1.07E-17	*	0.47015109
neighbourhoodDensity	1	116	49128.2938	78569.34	72.5331517	6.86E-14	*	0.3847236
cb:wordFrequency	1	116	1625.5234	426035.37	0.44259404	5.07E-01		0.00380096
cb:neighbourhoodDensity	1	116	956.1022	78569.34	1.41159199	2.37E-01		0.0120226
wordFrequency:neighbourhoodDensity	1	116	49690.4712	75451.92	76.3942729	2.09E-14	*	0.39707145
cb:wordFrequency:neighbourhoodDensity	1	116	2616.6512	75451.92	4.02284705	4.72E-02	*	0.03351734

Table 3. Analysis of Variance by Subject of Low-Frequency Words and Nonwords based on Orthographic Neighbourhood Density (ON) in the Word-Naming Task (RT)

Effect	DFn	DFd	SSn	SSd	*F*	*p*	*p<.05*	*ηp2*
(Intercept)	1	116	237340120	10729053.5	2566.06549	5.83E-81	*	0.95674975
cb	1	116	43.512709	10729053.5	0.00047045	9.83E-01		4.0556E-06
wordFrequency	1	116	405295	321350.5	146.301989	2.76E-22	*	0.55776166
neighbourhoodDensity	1	116	348522.71	395707.8	102.167886	1.31E-17	*	0.46829941
cb:wordFrequency	1	116	1.744052	321350.5	0.00062956	9.80E-01		5.4272E-06
cb:neighbourhoodDensity	1	116	8203.93039	395707.8	2.4049458	1.24E-01		0.0203112
wordFrequency:neighbourhoodDensity	1	116	21256.3253	124222.8	19.8492822	1.94E-05	*	0.14611255
cb:wordFrequency:neighbourhoodDensity	1	116	72.274404	124222.8	0.06749027	7.95E-01		0.00058147

Table 4. Analysis of Variance by Subject of Stimulus Types based on Orthographic Neighbourhood Density (ON) in the Word-Naming Task (PE)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	116	13352.798	8772.591	176.564093	4.7184E-25	*	0.60350568
cb	1	116	43.740808	8772.591	0.57838485	0.44848919		0.00496134
wordFrequency	2	232	1138.06025	4123.543	32.0149445	5.2684E-13	*	0.21629534
neighbourhoodDensity	1	116	330.681239	1755.866	21.8462146	8.0297E-06	*	0.15848251
cb:wordFrequency	2	232	18.323754	4123.543	0.51546827	0.59790323		0.00442403
cb:neighbourhoodDensity	1	116	0.967256	1755.866	0.06390106	0.80088054		0.00055057
wordFrequency:neighbourhoodDensity	2	232	664.798996	3424.212	22.5210039	1.1516E-09	*	0.16258186
cb:wordFrequency:neighbourhoodDensity	2	232	6.860702	3424.212	0.23241596	0.79280068		0.00199958

Table 5. Analysis of Variance by Subject of High- and Low-Frequency Words based on Orthographic Neighbourhood Density (ON) in the Word-Naming Task (PE)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	116	200346311	5411951.27	4294.23158	1.715E-93	*	0.97369753
cb	1	116	2326.5981	5411951.27	0.04986841	0.8236852		0.00042972
wordFrequency	1	116	378034.174	426035.37	102.930336	1.0701E-17	*	0.47015109
neighbourhoodDensity	1	116	49128.2938	78569.34	72.5331517	6.8639E-14	*	0.3847236
cb:wordFrequency	1	116	1625.5234	426035.37	0.44259404	0.50719368		0.00380096
cb:neighbourhoodDensity	1	116	956.1022	78569.34	1.41159199	0.23721917		0.0120226
wordFrequency:neighbourhoodDensity	1	116	49690.4712	75451.92	76.3942729	2.0861E-14	*	0.39707145
cb:wordFrequency:neighbourhoodDensity	1	116	2616.6512	75451.92	4.02284705	0.04721521	*	0.03351734

Table 6. Analysis of Variance by Subject of Low-Frequency Words and Nonwords based on Orthographic Neighbourhood Density (ON) in the Word-Naming Task (PE)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	116	237340120	10729053.5	2566.06549	5.8284E-81	*	0.95674975
cb	1	116	43.512709	10729053.5	0.00047045	0.98273263		4.0556E-06
wordFrequency	1	116	405295	321350.5	146.301989	2.759E-22	*	0.55776166
neighbourhoodDensity	1	116	348522.71	395707.8	102.167886	1.3126E-17	*	0.46829941
cb:wordFrequency	1	116	1.744052	321350.5	0.00062956	0.98002543		5.4272E-06
cb:neighbourhoodDensity	1	116	8203.93039	395707.8	2.4049458	0.12367638		0.0203112
wordFrequency:neighbourhoodDensity	1	116	21256.3253	124222.8	19.8492822	1.9422E-05	*	0.14611255
cb:wordFrequency:neighbourhoodDensity	1	116	72.274404	124222.8	0.06749027	0.79548715		0.00058147

Table 7.Analysis of Variance by Item of Stimulus Types based on Orthographic Neighbourhood Density (ON) in the Word-Naming Task (RT)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	116	13352.798	8772.591	176.564093	4.7184E-25	*	0.60350568
cb	1	116	43.740808	8772.591	0.57838485	0.44848919		0.00496134
wordFrequency	2	232	1138.06025	4123.543	32.0149445	5.2684E-13	*	0.21629534
neighbourhoodDensity	1	116	330.681239	1755.866	21.8462146	8.0297E-06	*	0.15848251
cb:wordFrequency	2	232	18.323754	4123.543	0.51546827	0.59790323		0.00442403
cb:neighbourhoodDensity	1	116	0.967256	1755.866	0.06390106	0.80088054		0.00055057
wordFrequency:neighbourhoodDensity	2	232	664.798996	3424.212	22.5210039	1.1516E-09	*	0.16258186
cb:wordFrequency:neighbourhoodDensity	2	232	6.860702	3424.212	0.23241596	0.79280068		0.00199958

Table 8. Analysis of Variance by Item of High- and Low-Frequency Words based on Orthographic Neighbourhood Density (ON) in the Word-Naming Task (RT)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	124	108371088	398792.32	33696.7745	6.885E-153	*	0.99633361
wordFrequency	1	124	205563.905	398792.32	63.917791	7.7306E-13	*	0.34013699
neighbourhoodDensity	1	124	31420.6009	398792.32	9.769884	0.00220988	*	0.073035
cb	1	124	1104.4814	27205.71	5.03408	0.02662774	*	0.03901356
wordFrequency:neighbourhoodDensity	1	124	30374.1469	398792.32	9.4445	0.00260448	*	0.07077474
wordFrequency:cb	1	124	1695.2617	27205.71	7.726778	0.00628896	*	0.0586576
neighbourhoodDensity:cb	1	124	748.1841	27205.71	3.410124	0.0671818		0.02676493
wordFrequency:neighbourhoodDensity:cb	1	124	1895.9175	27205.71	8.641341	0.00392085	*	0.06514816

Table 9. Analysis of Variance by Item of Low-Frequency Words and Nonwords based on Orthographic Neighbourhood Density (ON) in the Word-Naming Task (RT)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	124	128235161	670985.64	23698.2121	1.882E-143	*	0.99479477
wordFrequency	1	124	212116.804	670985.64	39.1997716	5.7188E-09	*	0.24019501
neighbourhoodDensity	1	124	197636.708	670985.64	36.5238098	1.635E-08	*	0.22752892
cb	1	124	80.30859	46153.51	0.215764	0.64310122		0.00173701
wordFrequency:neighbourhoodDensity	1	124	8653.30246	670985.64	1.5991542	0.20839475		0.01273221
wordFrequency:cb	1	124	285.6547	46153.51	0.7674646	0.38269646		0.00615116
neighbourhoodDensity:cb	1	124	2908.92584	46153.51	7.8153713	0.0060056	*	0.05929029
wordFrequency:neighbourhoodDensity:cb	1	124	287.66295	46153.51	0.7728601	0.38103415		0.00619414

Table 10. Analysis of Variance by Item of Stimulus Types based on Orthographic Neighbourhood Density (ON) in the Word-Naming Task (PE)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	186	7258.33842	3186.921	423.622349	7.9191E-50	*	0.69489307
wordFrequency	2	186	153.874755	3186.921	4.49033796	0.01245913	*	0.04605931
neighbourhoodDensity	1	186	6.1786E-30	3186.921	3.6061E-31	1		1.9387E-33
cb	1	186	85.6509743	3099.931	5.13917306	0.02454254	*	0.02688707
wordFrequency:neighbourhoodDensity	2	186	461.624264	3186.921	13.4710139	3.4387E-06	*	0.12652283
wordFrequency:cb	2	186	2.40516973	3099.931	0.0721567	0.93041112		0.00077528
neighbourhoodDensity:cb	1	186	3.9911E-29	3099.931	2.3947E-30	1		1.2875E-32
wordFrequency:neighbourhoodDensity:cb	2	186	7.21550919	3099.931	0.2164701	0.80555922		0.00232223

Table 11. Analysis of Variance by Item of High- and Low-Frequency Words based on Orthographic Neighbourhood Density (ON) in the Word-Naming Task (PE)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	124	108371088	398792.32	33696.7745	6.885E-153	*	0.99633361
wordFrequency	1	124	205563.905	398792.32	63.917791	7.7306E-13	*	0.34013699
neighbourhoodDensity	1	124	31420.6009	398792.32	9.769884	0.00220988	*	0.073035
cb	1	124	1104.4814	27205.71	5.03408	0.02662774	*	0.03901356
wordFrequency:neighbourhoodDensity	1	124	30374.1469	398792.32	9.4445	0.00260448	*	0.07077474
wordFrequency:cb	1	124	1695.2617	27205.71	7.726778	0.00628896	*	0.0586576
neighbourhoodDensity:cb	1	124	748.1841	27205.71	3.410124	0.0671818		0.02676493
wordFrequency:neighbourhoodDensity:cb	1	124	1895.9175	27205.71	8.641341	0.00392085	*	0.06514816

Table 12. Analysis of Variance by Item of Low-Frequency Words and Nonwords based on Orthographic Neighbourhood Density (ON) in the Word-Naming Task (PE)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	124	128235161	670985.64	23698.2121	1.882E-143	*	0.99479477
wordFrequency	1	124	212116.804	670985.64	39.1997716	5.7188E-09	*	0.24019501
neighbourhoodDensity	1	124	197636.708	670985.64	36.5238098	1.635E-08	*	0.22752892
cb	1	124	80.30859	46153.51	0.215764	0.64310122		0.00173701
wordFrequency:neighbourhoodDensity	1	124	8653.30246	670985.64	1.5991542	0.20839475		0.01273221
wordFrequency:cb	1	124	285.6547	46153.51	0.7674646	0.38269646		0.00615116
neighbourhoodDensity:cb	1	124	2908.92584	46153.51	7.8153713	0.0060056	*	0.05929029
wordFrequency:neighbourhoodDensity:cb	1	124	287.66295	46153.51	0.7728601	0.38103415		0.00619414

Table 13. Analysis of Variance by Subject of Stimulus Types based on Orthographic Neighbourhood Density (ON) in the Colour-Naming Task (RT)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	116	295442206	5683829.18	29.613980977	4.92E-103	*	0.98112475
cb	1	116	1890.02292	5683829.18	0.03857306	8.45E-01		0.00033242
wordFrequency	2	232	11899.1555	139403.85	9.90146295	7.47E-05	*	0.07864454
neighbourhoodDensity	1	116	567.73474	47111.71	1.39789517	2.39E-01		0.01190733
cb:wordFrequency	2	232	259.15864	139403.85	0.21564973	8.06E-01		0.0018556
cb:neighbourhoodDensity	1	116	56.1928	47111.71	0.13835976	7.11E-01		0.00119134
wordFrequency:neighbourhoodDensity	2	232	95.31939	87037.6	0.12703761	8.81E-01		0.00109395
cb:wordFrequency:neighbourhoodDensity	2	232	1141.52567	87037.6	1.52137667	2.21E-01		0.01294553

Table 14. Analysis of Variance by Subject of High- and Low-Frequency Words based on Orthographic Neighbourhood Density (ON) in the Colour-Naming Task (RT)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	116	195820045	3720523.68	6105.35699	3.68E-102	*	0.98135455
cb	1	116	1514.22946	3720523.68	0.04721126	8.28E-01		0.00040683
wordFrequency	1	116	6923.66088	54876.83	14.6354068	2.12E-04	*	0.11203246
neighbourhoodDensity	1	116	625.627835	43541.96	1.66673322	1.99E-01		0.01416486
cb:wordFrequency	1	116	224.143124	54876.83	0.47379932	4.93E-01		0.00406786
cb:neighbourhoodDensity	1	116	436.027329	43541.96	1.16161909	2.83E-01		0.00991467
wordFrequency:neighbourhoodDensity	1	116	2.654017	48727.58	0.00631811	9.37E-01		5.4463E-05
cb:wordFrequency:neighbourhoodDensity	1	116	487.894639	48727.58	1.1614731	2.83E-01		0.00991344

Table 15.Analysis of Variance by Subject of Low-Frequency Words and Nonwords based on Orthographic Neighbourhood Density (ON) in the Colour-Naming Task (RT)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	116	198704626	3986491.71	5781.96024	8.14E-101	*	0.98033218
cb	1	116	1703.5592	3986491.71	0.04957062	8.24E-01		0.00042715
wordFrequency	1	116	379.57659	72908.58	0.60391912	4.39E-01		0.00517924
neighbourhoodDensity	1	116	305.91771	43163.1	0.82214796	3.66E-01		0.0070376
cb:wordFrequency	1	116	159.02008	72908.58	0.25300629	6.16E-01		0.00217634
cb:neighbourhoodDensity	1	116	151.38564	43163.1	0.406846	5.25E-01		0.00349503
wordFrequency:neighbourhoodDensity	1	116	83.74383	41486.95	0.23415276	6.29E-01		0.00201449
cb:wordFrequency:neighbourhoodDensity	1	116	123.13936	41486.95	0.34430502	5.58E-01		0.00295936

Table 16. Analysis of Variance by Subject of Stimulus Types based on Orthographic Neighbourhood Density (ON) in the Colour-Naming Task (PE)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	116	164.781707	191.3829	99.876617	2.44E-17	*	0.46265604
cb	1	116	7.673717	191.3829	4.6511529	3.31E-02	*	0.03855042
wordFrequency	2	232	1.545387	350.7852	0.5110389	6.01E-01		0.00438618
neighbourhoodDensity	1	116	1.272526	216.6049	0.6814851	4.11E-01		0.00584056
cb:wordFrequency	2	232	1.809498	350.7852	0.5983771	5.51E-01		0.00513195
cb:neighbourhoodDensity	1	116	2.493615	216.6049	1.3354238	2.50E-01		0.01138125
wordFrequency:neighbourhoodDensity	2	232	6.906933	310.9496	2.576637	7.82E-02		0.02172972
cb:wordFrequency:neighbourhoodDensity	2	232	1.41615	310.9496	0.528296	5.90E-01		0.00453363

Table 17. Analysis of Variance by Subject of High- and Low-Frequency Words based on Orthographic Neighbourhood Density (ON) in the Colour-Naming Task (PE)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	116	195820045	3720523.68	6105.35699	3.68E-102	*	0.98135455
cb	1	116	1514.22946	3720523.68	0.04721126	8.28E-01		0.00040683
wordFrequency	1	116	6923.66088	54876.83	14.6354068	2.12E-04	*	0.11203246
neighbourhoodDensity	1	116	625.627835	43541.96	1.66673322	1.99E-01		0.01416486
cb:wordFrequency	1	116	224.143124	54876.83	0.47379932	4.93E-01		0.00406786
cb:neighbourhoodDensity	1	116	436.027329	43541.96	1.16161909	2.83E-01		0.00991467
wordFrequency:neighbourhoodDensity	1	116	2.654017	48727.58	0.00631811	9.37E-01		5.4463E-05
cb:wordFrequency:neighbourhoodDensity	1	116	487.894639	48727.58	1.1614731	2.83E-01		0.00991344

Table 18. Analysis of Variance by Subject of Low-Frequency Words and Nonwords based on Orthographic Neighbourhood Density (ON) in the Colour-Naming Task (PE)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	116	198704626	3986491.71	5781.96024	8.14E-101	*	0.98033218
cb	1	116	1703.5592	3986491.71	0.04957062	8.24E-01		0.00042715
wordFrequency	1	116	379.57659	72908.58	0.60391912	4.39E-01		0.00517924
neighbourhoodDensity	1	116	305.91771	43163.1	0.82214796	3.66E-01		0.0070376
cb:wordFrequency	1	116	159.02008	72908.58	0.25300629	6.16E-01		0.00217634
cb:neighbourhoodDensity	1	116	151.38564	43163.1	0.406846	5.25E-01		0.00349503
wordFrequency:neighbourhoodDensity	1	116	83.74383	41486.95	0.23415276	6.29E-01		0.00201449
cb:wordFrequency:neighbourhoodDensity	1	116	123.13936	41486.95	0.34430502	5.58E-01		0.00295936

Table 19. Analysis of Variance by Item of Stimulus Types based on Orthographic Neighbourhood Density (ON) in the Colour-Naming Task (RT)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	186	159959534	59140.33	503082.677	3.665967E- 321	*	0.99963042
wordFrequency	2	186	6317.80766	59140.33	9.93494889	7.95E-05	*	0.09651676
neighbourhoodDensity	1	186	485.154619	59140.33	1.52584143	2.18E-01		0.0081367
cb	1	186	765.785761	38417.37	3.70759794	5.57E-02		0.01954375
wordFrequency:neighbourhoodDensity	2	186	43.189693	59140.33	0.06791714	9.34E-01		0.00072976
wordFrequency:cb	2	186	227.873097	38417.37	0.55163067	5.77E-01		0.00589654
neighbourhoodDensity:cb	1	186	6.254405	38417.37	0.03028108	8.62E-01		0.00016277
wordFrequency:neighbourhoodDensity:cb	2	186	366.090213	38417.37	0.88622392	4.14E-01		0.00943934

Table 20. Analysis of Variance by Item of High- and Low-Frequency Words based on Orthographic Neighbourhood Density (ON) in the Colour-Naming Task (RT)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	124	106054167	42770.25	307473.458	2.46E-212	*	0.99959688
wordFrequency	1	124	3899.98489	42770.25	11.306881	1.03E-03	*	0.08356472
neighbourhoodDensity	1	124	465.6623	42770.25	1.3500535	2.48E-01		0.01077027
cb	1	124	704.30388	27705.06	3.1522642	7.83E-02		0.02479126
wordFrequency:neighbourhoodDensity	1	124	4.42038	42770.25	0.01281562	9.10E-01		0.00010334
wordFrequency:cb	1	124	181.20888	27705.06	0.8110395	3.70E-01		0.00649814
neighbourhoodDensity:cb	1	124	106.5081	27705.06	0.47670002	4.91E-01		0.00382963
wordFrequency:neighbourhoodDensity:cb	1	124	160.49869	27705.06	0.71834657	3.98E-01		0.00575975

Table 21. Analysis of Variance by Item of Low-Frequency Words and Nonwords based on Orthographic Neighbourhood Density (ON) in the Colour-Naming Task (RT)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	124	107579814	38167.4	349510.195	8.76E-216	*	0.99964534
wordFrequency	1	124	129.02049	38167.4	0.4191676	5.19E-01		0.003369
neighbourhoodDensity	1	124	297.15337	38167.4	0.9654054	3.28E-01		0.00772538
cb	1	124	748.21354	26142.68	3.5489277	6.19E-02		0.02782405
wordFrequency:neighbourhoodDensity	1	124	41.51925	38167.4	0.1348896	7.14E-01		0.00108664
wordFrequency:cb	1	124	159.93692	26142.68	0.7586131	3.85E-01		0.00608065
neighbourhoodDensity:cb	1	124	71.09189	26142.68	0.3372032	5.63E-01		0.00271201
wordFrequency:neighbourhoodDensity:cb	1	124	37.00373	26142.68	0.1755162	6.76E-01		0.00141345

Table 22. Analysis of Variance by Item of Stimulus Types based on Orthographic Neighbourhood Density (ON) in the Colour-Naming Task (PE)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	186	89.5827438	65.01978	256.266493	7.9346E-37	*	0.57943908
wordFrequency	2	186	0.20439628	65.01978	0.29235496	0.74684591		0.00313375
neighbourhoodDensity	1	186	2.7229E-31	65.01978	7.7892E-31	1		4.1877E-33
cb	1	186	0.00594428	67.83407	0.01629912	0.89854962		8.7622E-05
wordFrequency:neighbourhoodDensity	2	186	0.61318885	65.01978	0.87706487	0.41771538		0.0093427
wordFrequency:cb	2	186	0.05817002	67.83407	0.07975066	0.9233781		0.0008568
neighbourhoodDensity:cb	1	186	1.4603E-31	67.83407	4.0042E-31	1		2.1528E-33
wordFrequency:neighbourhoodDensity:cb	2	186	0.17451006	67.83407	0.23925197	0.78745839		0.002566

Table 23. Analysis of Variance by Item of High- and Low-Frequency Words based on Orthographic Neighbourhood Density (ON) in the Colour-Naming Task (PE)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	124	106054167	42770.25	307473.458	2.464E-212	*	0.99959688
wordFrequency	1	124	3899.98489	42770.25	11.306881	0.00102742	*	0.08356472
neighbourhoodDensity	1	124	465.6623	42770.25	1.3500535	0.24750033		0.01077027
cb	1	124	704.30388	27705.06	3.1522642	0.07827485		0.02479126
wordFrequency:neighbourhoodDensity	1	124	4.42038	42770.25	0.01281562	0.91005025		0.00010334
wordFrequency:cb	1	124	181.20888	27705.06	0.8110395	0.36955884		0.00649814
neighbourhoodDensity:cb	1	124	106.5081	27705.06	0.47670002	0.49121121		0.00382963
wordFrequency:neighbourhoodDensity:cb	1	124	160.49869	27705.06	0.71834657	0.39831976		0.00575975

Table 24. Analysis of Variance by Item of Low-Frequency Words and Nonwords based on Orthographic Neighbourhood Density (ON) in the Colour-Naming Task (PE)

Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ηp2
(Intercept)	1	124	107579814	38167.4	349510.195	8.76E-216	*	0.99964534
wordFrequency	1	124	129.02049	38167.4	0.4191676	0.51854951		0.003369
neighbourhoodDensity	1	124	297.15337	38167.4	0.9654054	0.32774195		0.00772538
cb	1	124	748.21354	26142.68	3.5489277	0.06192575		0.02782405
wordFrequency:neighbourhoodDensity	1	124	41.51925	38167.4	0.1348896	0.71404129		0.00108664
wordFrequency:cb	1	124	159.93692	26142.68	0.7586131	0.38544604		0.00608065
neighbourhoodDensity:cb	1	124	71.09189	26142.68	0.3372032	0.56250193		0.00271201
wordFrequency:neighbourhoodDensity:cb	1	124	37.00373	26142.68	0.1755162	0.67598004		0.00141345

Appendix B. Complete Set of Stimulus Items by Condition.

High-Frequency Words with High Orthographic Neighbourhood Density
Chin, clock, catch, came, camp, coal, crown, damp, deck, dust, face, farm, file, flat, found, felt, form, house, horse, keep, might, night, pipe, pound, prime, sound, sweet, spare, state, store, team, trace, trick, take, tight, wage, wing, wish

High-Frequency Words with Low Orthographic Neighbourhood Density
Chain, clear, cloth, cloud, club, court, crowd, dance, depth, desk, faith, fault, first, floor, force, fresh, from, hence, huge, kept, myth, north, piece, plus, proud, smart, smile, spend, stand, storm, teeth, trade, trend, twice, type, waist, wheel, which

Low-Frequency Words with High Orthographic Neighbourhood Density
Chop, clack, crape, clank, crock, corse, creed, douse, dent, dime, fade, fang, flan, flap, flog, foal, foil, hound, hitch, kilt, mime, nave, peach, pout, prow, snare, snout, spate, stoop, stoke, tame, tram, tamp, toil, tome, wipe, weft, wound

Low-Frequency Words with Low Orthographic Neighbourhood Density
Chirp, cache, crib, crumb, crypt, copse, cusp, doff, delve, dirge, farce, fern, flirt, fluff, fret, frizz, froze, hewn, hoax, kiln, midge, neigh, peeve, poove, preen, snail, smirk, spume, skimp, sprig, tempt, troop, tryst, twang, tuft, waive, weigh, whoop

Non-Words with High Orthographic Neighbourhood Density
Chim, clonk, carge, cabe, cass, cofe, crowl, dant, dess, dunt, fage, fank, fime, flas, fouth, fent, fote, hoise, hotch, keem, minch, nirth, pite, pouse, prine, sount, sweak, spale, stabe, stome, teap, trave, trink, tate, titch, wame, wilk, wime

Non-Words with Low Orthographic Neighbourhood Density
Chail, clebe, cloph, clort, cluf, couce, croul, danch, delch, desp, faich, fauch, firnt, floog, forve, freph, frob, hente, huce, kelk, mysk, norsh, pieph, plyt, proun, smarg, smife, speng, stasp, stoil, teesh, trafe, trech, twife, tyse, waish, wheef, whilb

Appendix C. List of Effects Simulated in the DRC model of Reading Aloud.

The complete list of effects successfully simulated in Coltheart et al.'s (2001) DRC model of Reading Aloud.

1. Frequency effect
2. Lexicality effect
3. Regularity effect
4. Interaction of regularity with frequency
5. Interaction of regularity with position of irregularity
6. Consistency effect
7. Pseudohomophone effect
8. Base word frequency effect on pseudohomophone reading
9. Absence of N effect on pseudohomophone reading
10. Presence of N effect on nonword reading
11. Whammy effect
12. Strategy effects
13. Homophone and pseudohomophone priming
14. Repetition priming
15. Onset effect in masked form of priming
16. Triple interaction between regularity, frequency, and repetition
17. Length effect
18. Interaction between lexicality and letter length

Appendix D. Characteristics often ascribed to automatic processes.

Note. Adapted/Reprinted from *Reading aloud is not automatic: Processing capacity is required to generate a phonological code from print* (p. 5), by M. G. Reynolds, and D. Besner, 2006, Journal of Experimental Psychology Human Perception & Performance. Adapted with permission.

1. Processing is stimulus driven.
2. Processing proceeds without intention.
3. Processing proceeds without a requirement for central capacity.
4. Processing is autonomous and independent. It is not dependent on other processes. It runs the same way every time.
5. The process is executed obligatorily. It is ballistic in the way that once triggered by a stimulus it occurs from start to finish and cannot be stopped.
6. Processing proceeds without conscious control of processing.
7. Processing proceeds without attention; it is independent of attention.
8. Attentional demands are reduced in an automatic process.
9. Processing occurs in parallel.
10. Processing speed is fast.
11. The stimulus captures attention entirely.
12. Processing cannot be altered by expectations.
13. Processing cannot be altered with practice.