CANADIAN PSYCHOLOGICAL ASSOCIATION
SOCIÉTÉ CANADIENNE DE PSYCHOLOGIE

Canadian Journal of Experimental Psychology /
Revue canadienne de psychologie expérimentale

ISSN: 1196-1961

2024, Vol. 78, No. 2, 114–128
https://doi.org/10.1037/cep0000326

On the Association Between Intention and Visual Word Identification

Derek Besner and Colin M. MacLeod
Department of Psychology, University of Waterloo

One of the most fundamental distinctions in cognitive psychology is between processing that is "controlled" and processing that is "automatic." The widely held automatic processing account of visual word identification asserts that, among other characteristics, the presentation of a well-formed letter string triggers sublexical, lexical, and semantic activation in the absence of any intention to do so. Instead, the role of intention is seen as independent of stimulus identification and as restricted to selection for action using the products of identification (e.g., braking in response to a sign saying "BRIDGE OUT"). We consider four paradigms with respect to the role of an intention—defined here as a "task set" indicating how to perform in the current situation—when identifying single well-formed letter strings. Contrary to the received automaticity view, the literature regarding each of these paradigms demonstrates that the relation between an intention and stimulus identification is constrained in multiple ways, many of which are not well understood at present. One thing is clear: There is no simple relation between an intention, in the form of a task set, and stimulus identification. Automatic processing of words, if this indeed ever occurs, certainly is not a system default.

Public Significance Statement
With the extensive practice that readers have in processing words, many theorists argue that this skill becomes "automatic," in the sense of occurring necessarily and without intention. In this article, we argue that there is considerable evidence against the reading of individual words being "automatic" in that particular sense. In many situations, intention demonstrably matters. We examine several frequently studied tasks that involve single words (and nonwords, like "mantiness") and show that the context in which they occur is important for understanding how such stimuli are processed. The "task set" that people adopt—their approach to the particular situation—and when they adopt that set both play pivotal roles in how words are processed. Consequently, the concept of automaticity as typically defined is too broad to capture the many subtleties involved in the skill of visual word recognition.

Keywords: visual word identification, intention, task set, automaticity

A fundamental issue in cognitive psychology is the automatic-controlled distinction, going back at least to Cattell (1886) and his timing of the naming of words and objects, and to Bryan and Harter (1899) and their studies of the effects of practice on telegraph operators. Modern accounts of this distinction have been strongly driven by Posner and Snyder's (1975) and Shiffrin and Schneider's (1977; Schneider & Shiffrin, 1977) seminal work. In a review of the literature relevant to the distinction, Moors and De Houwer (2006, p. 297) maintained that what is meant by automaticity varies quite considerably, such that it "should be diagnosed by looking at the presence of features such as unintentional, uncontrolled/uncontrollable, goal independent, autonomous, purely stimulus driven, unconscious, efficient, and fast." How exactly to accomplish such a diagnosis continues to be challenging.

Schneider and Shiffrin (1977) demonstrated that visual search can become highly efficient after practice, virtually eliminating the influence of search set size, provided that targets never appear as distractors and distractors never appear as targets (consistent mapping). In contrast, search is considerably slower and is strongly influenced by the size of the search set when targets are sometimes distractors and distractors are sometimes targets (varied mapping). These empirical findings led Shiffrin and Schneider (1977) to postulate that practice with consistent mapping is associated with decreased demands on attention and working memory, and hence with parallel processing: On their account, behaviour has become more automatized. In contrast, the varied mapping situation, despite extensive practice, is associated with greater demands on attention and memory and with serial rather than parallel processing.

Automaticity need not, however, be conceived of as a qualitatively different state—a kind of syndrome. Some theorists (e.g., MacLeod & Dunbar, 1988) have explicitly argued that the automatic-controlled distinction reflects a continuum of learning rather than two discrete modes of processing. Indeed, Shiffrin and Schneider's work with consistent versus varied mapping fits with this characterization in that practice with consistent mapping does not completely eliminate the influence of search set size but instead

This article was published Online First April 11, 2024.
Derek Besner https://orcid.org/0000-0003-0442-2799
Colin M. MacLeod https://orcid.org/0000-0002-8350-7362
This research was supported by Natural Sciences and Engineering Research Council of Canada Discovery Grants to Derek Besner (50503-10030) and Colin M. MacLeod (A7459). The authors are grateful to Rob McCann and Evan Risko for their comments.
Correspondence concerning this article should be addressed to Derek Besner, Department of Psychology, University of Waterloo, Waterloo, ON N2L 3G1, Canada. Email: dbesner@uwaterloo.ca

progressively reduces this influence until it becomes very small. Relatedly, Logan's (1988) distinction between algorithmic and direct retrieval when considering how performance changes over time as a function of practice is explicitly understood in terms of a continuum.

Of the many skills that we develop by dint of extended practice, reading is often characterized as the one that best illustrates automaticity—essentially the claim that Cattell (1886) made. Reading is a uniquely human ability and, for experienced readers, reflects a highly skilled set of processes. Thus, if one wants to investigate "truly" automatic processing of a learned and quintessential human ability, it is no surprise that many of the processes associated with visual word identification (e.g., feature activation; letter activation; orthographic word level activation; application of sublexical spelling-sound correspondences; lexical phonology activation; semantic activation) have been viewed as strong contenders for this status. Yet many of the writings about the putative automaticity of visual word recognition adopt a discrete rather than a continuous perspective, at least as it applies to the case of highly skilled readers (e.g., Augustinova & Ferrand, 2012a, 2012b, 2014; Brown et al., 2002; Burca et al., 2021; LaBerge & Samuels, 1974; Lien et al., 2006; Marcel, 1983; Neely, 1977; Neely & Kahan, 2001; Paap & Herdman, 1998; Posner & Snyder, 1975; among many others). Consequently, it is this skill—visual word identification—that is the focus of this article.

Intention and the Concept of Task Set

In many quarters, the concept of automaticity is cast as requiring no role for intention beyond directing what to do with the products of automatic processing. This is likely why Moors and De Houwer (2006) placed "unintentional" first in their list of features of automaticity cited above. By "intention," we do not mean to suggest something like a conscious process; rather, we view it as generally unconscious, conceiving of intention as the implementation and regulation of a "task set." Whatever task one is performing, some approach to that task must be taken, whether instructed or self-generated. A task set, then, is the set of processes marshalled to make it possible to do what the situation demands without necessarily specifying in great detail how one is able to do that (see Monsell, 2017, for an extended discussion of task set implementation and regulation). Our concern pertains to when such processing is in play in one particular context, that of visual word identification.

The general problem here is that, in the typical experiment, the participant is instructed regarding what they have to do before the experiment starts, and the consequent "task set" that is implemented remains in place throughout the experiment. This context makes it difficult to determine whether a task set is needed beyond the point when the instructions to the participants have been provided (and presumably processed). Fortunately, there are experiments scattered in the visual word recognition literature whose results tell us about whether a task set was needed to allow stimulus identification to take place, or whether the implementation of a task set (e.g., for action) occurred during the time that stimulus identification was taking place (the classic view) or whether it occurred after stimulus identification, as, for example, in tuning parameters for the upcoming trial based on the experience of the preceding trial.

In what follows, we consider evidence from the four paradigms shown in Table 1. The net outcome of all of these findings is to call for a more nuanced understanding of the relation between a task set and stimulus identification. Indeed, the major conclusion reached here is that a task set (a) often precedes stimulus identification (e.g., when it is implemented at the outset and applies to the entire set of trials, or when it is implemented anew as when the task changes across trials, as in task switching), (b) sometimes occurs in parallel with stimulus identification (e.g., when it serves to identify what action is called for on a given trial but does not precede stimulus identification), and (c) sometimes occurs after stimulus identification (e.g., for the purpose of making processing more efficient on the upcoming trial). In other words, there is no fixed relation between a task set and stimulus identification in the context of visual word recognition, contrary to the widely held view that task set plays no role beyond directing what action to take with respect to the products of automatic stimulus identification (e.g., Brown et al., 2002; Neely, 1977; Neely & Kahan, 2001; Posner & Snyder, 1975, among many others).

Table 1
The Four Tasks and the Issues Considered

The tasks	The issues considered
Stroop colour naming	(a) Salience, (b) response control, (c) hypnotic suggestion, (d) instructional set, (e) control of an automatic process, (f) contingency effects, and (g) spatial attention
Letter search through a prime	Words versus nonwords
Task switching	(a) Basic switch cost, (b) stimulus onset asynchrony and task cuing, (c) stimulus set (words vs. nonwords), and (d) Go versus No-Go (words only)
Stimulus switching (without task switching)	(a) Switch cost and (b) the dual-route cascaded model

Stroop Colour Naming

We start with Stroop's (1935; for a review, see MacLeod, 1991) famous phenomenon because many—indeed, perhaps most—cognitive psychologists see the results from this paradigm as so compelling and as speaking very directly to the concept of "automaticity." When the task is to identify the print colour of a word and the reader is instructed to refrain from reading the word (by instruction, the word is irrelevant), the primary result consists of slowed responses to a stimulus consisting of, for example, the word green printed in red (an incongruent trial) as compared to the word green printed in green (a congruent trial). This interference is widely taken to provide the strongest evidence that word identification is indeed automatic in the sense that it is intention-free (among other characteristics): Participants must be reading the words despite their instruction-driven intention not to do so; otherwise, they would not suffer interference.

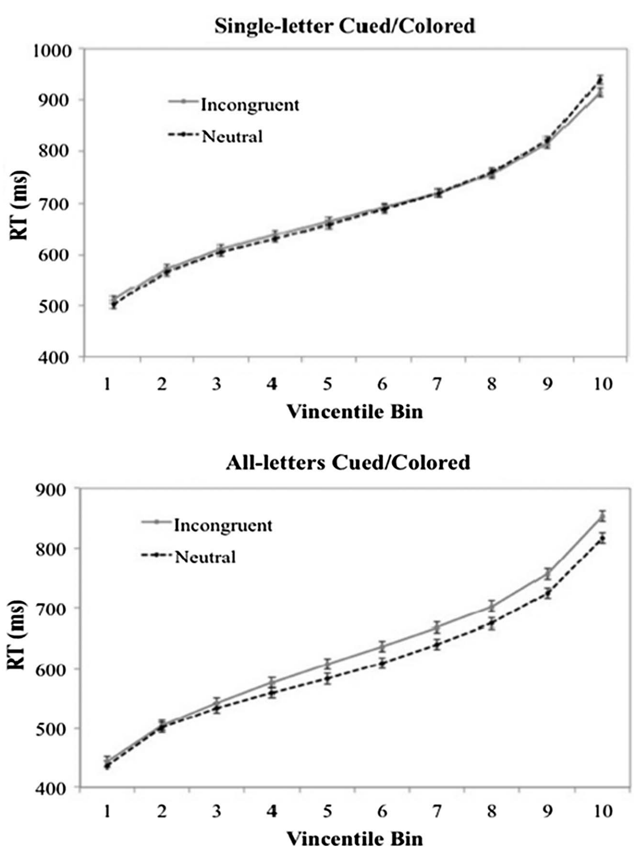
This account is prevalent in introductory psychology textbooks, cognitive psychology textbooks, chapters, and literally hundreds of refereed journal articles (e.g., see Augustinova & Ferrand, 2012a, 2014; Brown et al., 2002; Burca et al., 2021). Indeed, a brief survey of nine cognition textbooks taken more or less at random from our bookshelves revealed that "the" Stroop effect is consistently taken to indicate that word identification occurs in the absence of any intention on the part of the reader; indeed, that it cannot be prevented from occurring. This claim is made repeatedly despite the fact that in some of its variants—such as the semantic Stroop task, where the interfering words are colour-related words (e.g., the word sky printed in red or the word blood printed in yellow) the interference effect (incongruent minus neutral) is absent in approximately 25% of the distribution of response times (RTs)—those at the fast end (Labuschagne & Besner, 2015; see the top panel of Figure 1). This seems an inconvenient fact for the claim that unintentional word identification occurs in all trials. In what follows, we will consider several features of Stroop experiments relevant to the role of intention.

Salience

For starters, the word has to be more salient than the colour for interference to occur. In the classic Stroop task, reading is considerably faster than colour naming because the word is more salient than the colour. However, when the word and the colour are made to be equally discriminable, Melara and Mounts (1993) and Sabri et al. (2001; see Algom & Chajut, 2019 for further discussion) report no Stroop effect, in line with a Garner-type analysis of discriminability (see Algom & Fitousi, 2016). It is remarkable how little attention their findings have attracted in the literature, both in terms of the design of experiments and the development of theory.

Figure 1

Response Times for the Incongruent and Neutral Conditions When a Single Letter Versus the Entire Word Is Cued/Coloured in a Semantic Stroop Task



Note. These data appear as Figure 2 in "Automaticity Revisited: When Print Doesn't Activate Semantics," by E. M. Labuschagne and D. Besner, 2015, *Frontiers in Psychology*, *6*, Article 117 (https://doi.org/10.3389/fpsyg.2015.00117). CC BY-NC. RT = response time.

Response Control

On the vast majority of incongruent trials in any Stroop experiment, participants rarely read the irrelevant word aloud: Instead, in line with instructions, they correctly name its print colour aloud. On the rare occasions when they do slip and read the word, such responses are readily identified as errors because the trial is incongruent, so their response times are discarded from analysis. Thus, participants are routinely engaged in an act of control (an intention) that makes it possible to avoid executing the phonological code corresponding to the word. The point here is that the intention not to read the word on each trial—the task set—prevents responding aloud with the irrelevant word on all but a few occasions.

Hypnotic Suggestion

Raz et al. (2002, 2003) reported an intriguing finding—that the Stroop effect was eliminated among highly hypnotizable participants by hypnotizing them and instructing them that they would be unable to read the words because they were in an unfamiliar foreign language (although, of course, they were not). This result is inconsistent with the received view that the words must be read despite the intention not to do so. Follow-up work by Augustinova and Ferrand (2012a) paints a more nuanced picture. They reported that the standard Stroop effect was not abolished but was instead reduced in magnitude by a hypnotic suggestion. Rather than being concerned about whether their participants were, for example, as deeply hypnotized (or hypnotizable) as Raz and colleagues' participants were, they reported within the same experiment that the semantic Stroop effect was completely unaffected by hypnosis.

For the present, it remains an open question whether highly hypnotizable participants treated as in the Raz et al. (2002, 2003) studies would also show a reduced semantic Stroop effect. Nonetheless, the bottom line is that hypnosis, for highly hypnotizable participants, at least reduces the magnitude of the Stroop effect when colour words form the incongruent condition. This fact is hard to square with the claim that it is impossible to prevent the irrelevant word from being read. Any reliable reduction in the effect stemming from the application of another instructional manipulation seems inconsistent with the idea that reading of the nominally irrelevant word is always unintended in the sense of being obligatory and unavoidable.

Instructional Set

In the context of Stroop experiments, there are various other demonstrations that instructions can alter how a participant processes the irrelevant words (e.g., see the brief review by Bauer & Besner, 1997; see also the extension by Risko et al., 2005). The demonstration by Bauer and Besner (1997) is sufficiently unusual but also sufficiently simple as to merit comment. In separate blocks of trials, they had participants either press one of two buttons corresponding to the two different colours in the experiment (the CLASSIFY condition) or press one button if a target colour appeared but another button if the target colour was absent (the DETECT condition). There was a significant congruency effect in the CLASSIFY condition—the usual Stroop interference—but there was no congruency effect in the DETECT condition. Clearly, how the identical stimuli were processed in these two instructional conditions differed. Following the instructions yielded a qualitative difference in these two conditions, even if it is not evident exactly how the different task sets caused the different patterns. Possibly, the CLASSIFY condition encouraged prototypical "reading," whereas the DETECT condition promoted a top-down matching process that does not invoke reading (look for *x*, turning the task into a YES/NO one). Whatever the answer, these results are not readily accommodated by the classic interpretation of the Stroop effect.

Testing Hebrew-Arabic bilingual participants, Tzelgov et al. (1990) varied how frequently words appeared in the first language versus in the second language across blocks of trials. In so doing, they demonstrated that Stroop interference was controllable, as evidenced by reduced response time in the incongruent condition, but only when the words in the first language were more frequent. Apparently, participants successfully developed an expectation—a kind of task set—only in the language with which they were more familiar. Surely the opposite result would be expected under an automaticity account because the first language should be the more automatic one and therefore harder to control. Shortly thereafter, Tzelgov et al. (1992) reported that Stroop interference decreased as the proportion of incongruent trials increased, again providing participants with an expectation upon which they evidently could act to control the likelihood of word reading.

There is even evidence that having another person in the room reduces the magnitude of the Stroop effect (Augustinova & Ferrand, 2012b; Sharma et al., 2010), a manipulation that obviously is not exerting its effect via a change in instructions. We expect that there are other factors—beyond the obvious factor of training (MacLeod & Dunbar, 1988)—that also modulate the size of the effect in the standard version of the task. Such findings undermine the notion that the word must always be read because reading is intention-free and must always occur.

Task Set Control of a Putatively Automatic Process

How might processing the irrelevant word in the Stroop context depend on a task set? Cohen et al. (1990) provided an existence proof in the form of a parallel distributed processing model with separate input units (one for colours and one for words) and a common set of output units, mediated by two sets of task units—one for colours and one for words. Making the task set units for the words weak relative to those for the colours allowed the model to produce the standard Stroop effect on response time and to do so without producing an unacceptably high proportion of errors for incongruent trials. The instructions to the reader (in this case, to the model) were implemented as settings in the task units, clearly constituting a task set, one that was necessary to produce a Stroop effect.

Contingency Effects

If such a dampening process exists for skilled readers (analogous to weakening the word settings in the task demand units in the Cohen et al., 1990, computational model), and if doing so typically is sufficient to prevent participants from uttering the nominally irrelevant word aloud, why do participants not just make the dampening process strong enough to entirely prevent any processing of the irrelevant word? One answer to this question is that we are always learning—most often incidentally—about events in our environment. Much of this learning is associative and is susceptible to contingencies, the processing of which need not be conscious (e.g., see the reviews of contingency learning by MacLeod, 2019; Schmidt, 2021). The Stroop paradigm is no exception to this general claim about learning because processing the irrelevant word sometimes can help to identify its colour (in the congruent condition), even if at other times there is a cost to doing so (in the incongruent condition; Schmidt & Besner, 2008). MacLeod's and Schmidt's reviews unpack what the elements of contingency learning are, and Lin and MacLeod (2018) have explored how remarkably quickly these contingencies are learned, unlearned, and then relearned (see also Schmidt et al., 2007, 2010).

Put simply, the default task set on an ongoing basis, largely unconscious, is to explore the environment for contingencies and to modify processing in response to those contingencies. In this regard, there is now an extensive literature on how manipulating the likelihood of congruent and incongruent trials (and indeed of "neutral" control trials such as *table* or *xxxx*) across the set of trials in an experiment, whether experiment-wide (e.g., Tzelgov et al., 1992) or at the level of individual items (e.g., Jacoby et al., 2003), can dramatically alter the magnitude of the Stroop effect. Again, if the processing of the word on each trial was inevitable because it was intention-free, it is not at all obvious why such manipulations should alter the effect.

Spatial Attention

Spatial attention, despite being widely ignored in the visual word recognition literature as well as in the leading computational models of reading aloud (e.g., see Coltheart et al., 2001; Plaut et al., 1996), is nonetheless intimately intertwined with those processes. Intention can be seen as operating by directing spatial attention. Indeed, Lachter et al. (2004) confirmed the major claim made by Broadbent (1958) to the effect that there is no identification without attention.

Spatial attention can be deployed broadly or more narrowly in accord with the elements of an experiment (including the instructions). When set broadly, as when the spatial cue is only 50% valid, attention is distributed across both cued and uncued locations but is biased to the cued location. This results in a cueing effect such that a related word that appears outside the spatially cued location can nonetheless "prime" the target word within the cued location (Besner et al., 2005; Waechter et al., 2011). But when attention is set more narrowly, as when the spatial cue is a 100% valid cue for the location of the target, a word that appears in the uncued location no longer primes a word in the cued location (Besner et al., 2005; Waechter et al., 2011).

Just as it can within an object, the narrowing of spatial attention can also apply within a word. Consider the case of colouring only a single letter in a word (e.g., red where only the e is in green with the other letters in white). Now imagine spatially cuing that letter via a small arrow pointing to the e. Compared to a fully coloured word (e.g., red entirely in green), this narrowing of attention reduces the interference caused by the word (Besner, 2001; Besner & Stolz, 1999; Besner & Young, 2024; Labuschagne & Besner, 2015). Here, the task set controls the distribution of spatial attention, which in turn is best understood as a necessary preliminary to visual word recognition (e.g., Besner et al., 2016; McCann et al., 1992; see the bottom panel of Figure 1).[1]

In this section, we have shown that the Stroop effect is sensitive to a variety of manipulations that should not affect the amount of interference—if the processing of the irrelevant word was purely automatic in the sense of not requiring a task set to process the word and not being differently affected by changes in task set. The strong implication is that, while it remains possible that some subset of processes in this task could be "automatic" in various ways, at least some clearly are not: These can be influenced by the design of the experiment and even by the instructions (the "task set").

Letter Search Through a Prime

All models of visual word recognition include a mechanism for the translation of spelling into sound sublexically (i.e., without having to access the word in the reader's lexicon). This mechanism can be applied both to familiar words and to new words (which, of course, start out life as nonwords). Not surprisingly, there is considerable debate regarding how to characterize this mechanism (e.g., Coltheart et al., 2001; Frost, 1998; Perry et al., 2007; Plaut et al., 1996; Rastle & Coltheart, 2006). Yet, despite our ongoing case for the importance of task set, all accounts assume, typically implicitly, that this conversion process unfolds in the complete absence of a task set. Indeed, the terms "task set" and "automatic" appear nowhere in descriptions of the leading computational models as advanced, for example, by Plaut et al. (1996) and Coltheart et al. (2001). Instead, it is implicitly assumed that the mere presentation of a word or nonword suffices to trigger sublexical spelling-to-sound conversion. In other formulations, this is explicitly seen as an "automatic" and obligatory process (e.g., see Frost, 1998; Van Orden et al., 1992, among many others). Indeed, the "regularity effect"—the fact that exception words like *pint* are slower to read aloud than are regular words like tint, hint, splint, and mint is often construed as evidence that such sublexical phonological recoding is automatic because if one could prevent such a computation, then one would do so—and hence there would be no regularity effect because word-level processing (the lexical route) itself would suffice. This claim has been challenged by Robidoux and Besner (2011) with the aid of simulation evidence from the leading computational models.

Given the claim made in the computational models that such processing is free of any task set, it follows that forcing the reader to adopt a particular set—here, explicitly attending to the letter level—should not influence the sublexical spelling-to-sound conversion to print. Ferguson and Besner (2006) reported that prime items consisting of pseudohomophones like BRANE speed the processing of sound-identical target words like BRAIN, and that prime items like MARKED speed the processing of morphologically related words like MARK. In the former case, sublexical spelling-to-sound conversion is required to produce priming of BRAIN by a letter string like BRANE. Sublexical spelling-to-sound conversion is not required, however, for prime words like MARKED; lexical activation for both MARK and MARKED can be based on orthography alone.

The central question here is: What happens when the reader is given a task set that requires explicit identification of a letter in the prime before making a lexical decision about the target? In a second experiment, Ferguson and Besner (2006) demonstrated that such a letter search through a prime like BRANE eliminated priming of the target BRAIN (see also Kahan et al., 2006), whereas letter search through a prime like MARKED did not prevent morphological priming for MARK.

The BRANE/BRAIN result is consistent with the idea that sublexical spelling-to-sound correspondences cannot operate because this requires a task set, and a different task set is being used to support letter search. These two different task sets cannot operate at the same time. In contrast, priming driven by orthography persists despite letter search on the prime, consistent with the idea that orthographic processing in aid of lexical activation does not require a task set as a preliminary to operate. Nevertheless, it may be that a more demanding operation might disrupt orthographic operations that support lexical access. For example, a prime like JUGDE (which transposes the letters G and D) would surely prime a target like JUDGE when letter search is not required on the prime, but such priming might be eliminated were letter search through JUGDE to be required.

Task Switching

Experiments from another distinct literature are easier to understand as manifestations of "task set" because they explicitly use that terminology and appeal to that idea theoretically. Allport et al. (1994) and Rogers and Monsell (1995) were the first modern-day researchers to report that, for consecutive trials within an experiment, a response time cost was evident when the task switched relative to when it did not. This was true even when the task for the upcoming trial was perfectly predictable. An illustration would be an alternating runs paradigm where an Arabic numeral is presented on each trial. On two consecutive trials, the task would be to make a parity judgment ("is it odd or even?"), whereas on the next two trials, a magnitude judgment would be called for ("is it larger/smaller than 5?"); this sequence would of course keep repeating. The cost is evident in the slower response times for the switch trials than for the nonswitch trials.

Basic Switch Cost

"Switch" costs have been well documented in numerous articles and are associated with a complicated literature (see reviews by Monsell, 2017; Pashler, 2000, among others). The switch cost indexes, at least in part, whatever it is that participants do that enables

[1] Of course, the reduction or elimination of the semantic Stroop effect does not demand the interpretation that semantic processing itself has been derailed. It is sufficient that a prior process has been stopped; Neely and Kahan (2001) make this point explicitly. It should also be noted that the Augustinova laboratory (see Augustinova & Ferrand, 2014; Burca et al., 2021) takes the position that there is no convincing evidence that the semantic Stroop effect is influenced by colouring only a single letter and spatial cuing. Their experiments do not, however, implement procedural details that Besner and Young (2024) see as important, in particular spacing between the letters when crossed with the spatial attention manipulation.

them to follow the instructions. Despite the robustness of the switch cost effect, it is not clear whether it reflects a delay before target processing can begin, a slowing at the level of the response (e.g., proactive interference from the other task—so-called attentional inertia), or other possibilities (see Monsell, 2017, for discussion). Indeed, multiple factors could well be involved. That reading-based processing cannot begin until the task set is in place would constitute clear evidence that intention matters.

To be more concrete, imagine a Stroop-like task-switching experiment. It should not be surprising that switching tasks between colour naming and word reading across trials, even when the switch is perfectly predictable, yields a switch cost in both directions. Of particular interest, though, we do not know whether identification of the word starts only when the appropriate set is in place or whether word identification starts during the time that the cue for the task is being interpreted and the appropriate task set is being implemented. This kind of experiment is simply not analytic enough. Two published reports are important exceptions to this issue of ambiguity with respect to where at least some of the switch cost arises in the processing sequence. In both, factors that affect early processing were manipulated.

Oriet and Jolicoeur (2003) used the paradigm described above: Arabic numerals and magnitude versus parity judgments in the alternating runs paradigm. Critically, they also varied stimulus quality (i.e., how hard vs. easy it was to see the stimulus, accomplished by varying contrast). They reasoned that if the component of stimulus identification that is influenced by stimulus quality (minimally, the earliest possible locus in the information processing sequence—the feature level) could be carried out while the participant was readying the appropriate set in the switch condition, then stimulus quality and switch/no-switch ought to produce an interaction. That interaction was expected to be underadditive: The stimulus quality effect should be smaller (or absent) on switch trials relative to no-switch trials (an effect described as "absorption into slack"; e.g., see Besner & Care, 2003; Pashler, 1984, 1994). This logic and outcome are depicted graphically in Figure 2 (Panel B) and are used throughout this section. Note that Task 1 does not require an explicit response but Task 2 does.

The data reported by Oriet and Jolicoeur (2003) were unambiguous in showing no such interaction: Relative to the no-switch condition, the time cost of a switch was additive with the effect of stimulus quality (see Panel A of Figure 2). Evidently, participants could not both prepare the task set on switch trials and simultaneously deal with stimulus quality during the time taken to implement the new task set. Oriet and Jolicoeur observed this pattern in response time in two experiments across a range of intervals between the prior stimulus and the predictable onset of the target stimulus that required the same versus a different kind of categorization.

Oriet and Jolicoeur's (2003) final experiment used the same logic (looking for underadditivity) but this time in the context of the psychological refractory period paradigm (see Pashler's, 1994 review). Here, stimulus quality again was manipulated but the target task was held constant across trials. Now there was an additional variable: A tone appeared before the target with the time between the tone and the target—the stimulus onset asynchrony, or SOA—varying. The task was to classify the tone as high or low, requiring an explicit response, before responding to the target. Typically, responses to a target under these conditions are considerably slowed (by several hundred ms) because of having to first respond to the tone: The shorter the SOA, the longer is the delay in responding to the target. But if the target can be processed to any extent while preparing the response to the tone (within the period of "cognitive slack" noted earlier), then stimulus quality should have a smaller effect as SOA decreases. Oriet and Jolicoeur now observed strong underadditivity between stimulus quality and SOA (indeed, there was no remaining effect of stimulus quality at the shortest SOA). This experiment shows that it is possible to see underadditivity of this particular factor, stimulus quality, in concert with SOA in a paradigm closely related to the task-switching paradigm.

Taken together with the results of their first two experiments using the alternating runs paradigm, Oriet and Jolicoeur (2003) reasoned that the additive effects of stimulus quality and switch/no-switch implied a structural bottleneck in which even the earliest elements of stimulus identification had to wait until a task set was in place before such identification could begin. In short, there was no evidence of automatic processing.

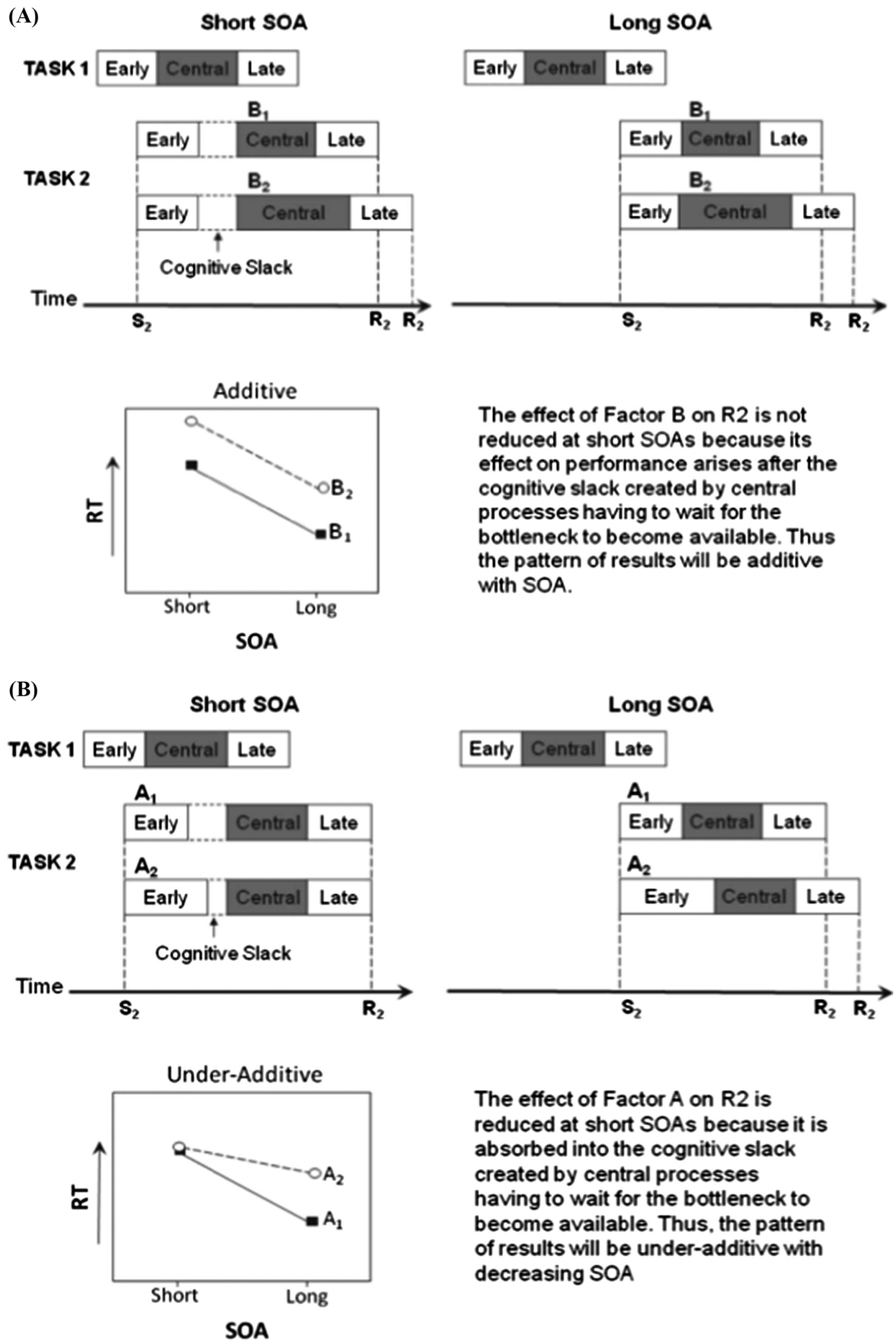
One other report merits particular consideration because it used words rather than Arabic numerals. Elchlepp et al. (2015) used two tasks: In Experiment 1, symmetry decisions about colour were paired with a semantic judgment of the printed word (living/nonliving); in Experiment 2, symmetry judgments of colour were paired with lexical decision in response to the letter string (word/nonword). In both experiments, these manipulations were combined factorially with switch/no-switch using a long SOA. Their behavioural data yielded strongly additive effects of switch/no-switch and word frequency (in Experiment 1) and of word/nonword (in Experiment 2). On its face, the absence of an interaction in which the word frequency effect in Experiment 1 was attenuated under the switch condition as compared to the no-switch condition and the associated failure to see any such attenuation of the word/nonword difference in Experiment 2 are most easily understood as reflecting a delay in lexical processing until the task set is put into place. Elchlepp and colleagues also reported evoked response potential data that largely converged with their behavioural data, again showing delays in these measures for the switch condition relative to the no-switch condition.[2]

The results from both the Oriet and Jolicoeur (2003) and the Elchlepp et al. (2015) studies thus imply that at least some of the time cost associated with the reconfiguration of the task set in the switch condition is a preliminary to lexical processing. Again, the participant must know what to do with a word before processing of it can begin.

[2] A clarification is warranted. Elchlepp et al. (2015, p. 318) claim that "Besner and colleagues (Besner et al., 1997, 2009) have argued for the automaticity of at least early stages of lexical identification." Besner and colleagues maintained that spatial attention is a necessary preliminary to lexical processing; they said nothing about whether subsequent processes also needed spatial attention (e.g., see Besner, 2001; Besner & Stolz, 1999; Labuschagne & Besner, 2015). Relatedly, in the context of the psychological refractory period paradigm, Besner et al. (2009) claimed only that lexical processing (but not semantic processing; see Besner et al., 2009; Besner & Reynolds, 2017; Reynolds & Besner, 2006; White & Besner, 2018) in Task 2 could sometimes proceed during Task 1 processing. That is, lexical processing, at least for very skilled readers, may not need a task set as a preliminary in that context. For less skilled readers, additivity of word frequency and SOA has been reported in many experiments (e.g., Lien et al., 2006; McCann et al., 2000). Older work is also relevant but has been widely ignored: Both Becker (1976) and Herdman (1992) reported that the effect of a word frequency manipulation was modulated by dual task conditions, contrary to the claim that processing is automatic in some sense.

Figure 2

Graphical Description of Processing That Is Bottlenecked (Top Panel) Versus Absorbed Into Slack (Bottom Panel)



Note. Panel A depicts an additive effect; Panel B depicts an underadditive interaction. These figures are from "When Underadditivity of Factor Effects in the Psychological Refractory Period Paradigm Implies a Bottleneck: Evidence From Psycholinguistics," by D. Besner, M. Reynolds, and S. O'Malley, 2009, *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *62*(11), 2222–2234 (https://doi.org/10.1080/17470210902747187). Copyright 2009 by Sage. SOA = stimulus onset asynchrony; RT = response time.

Stimulus Onset Asynchrony and Task Cuing

Using a variant of the task-switching paradigm where instead of a predictable task alternation, a cue indicated the task to perform on each trial, other work also set out to address the issue of whether a task set is ever a preliminary to various word recognition processes. The critical differences from the experiments outlined in the preceding section are that in these cuing studies, (a) SOA between cue and target was manipulated randomly across trials, and (b) the emphasis was on whether SOA and a second factor (e.g., stimulus quality) had additive or underadditive effects on response times. Again, the critical issue is whether "absorption into slack" was or was not seen (see Figure 2 for examples of additivity and underadditivity).

Besner and Care (2003) presented a pronounceable nonword on each trial and had participants either read it aloud or decide by pressing one of two buttons whether the nonword appeared in upper or lower case. The two critical manipulations were stimulus quality and SOA. For each trial, a task cue (a coloured frame around the word) signalled what task to do. In one condition, the task cue appeared 750 ms before the onset of the target nonword (SOA = −750); in the other condition, the onsets of the task cue and the target coincided (SOA = 0). The word was either dim or bright, representing stimulus quality. These two factors were manipulated in a 2 × 2 within-subject design, with all four conditions appearing randomly within a single block of trials.

The condition in which the task cue appeared 750 ms before the target provides an estimate of the magnitude of the stimulus quality effect when the reader is allotted sufficient time to generate the appropriate task set before the target appears. In comparison to this −750 SOA condition, the 0 SOA condition makes it possible to determine when the reader puts a task set in place. Must this task set be in place before beginning to process the target, or is the target processed to some degree during the time taken to decode the task cue and to implement the task set that is needed for other target processes? An intention-free account in which stimulus identification can start while the task cue is being interpreted and a task set is being implemented makes the straightforward prediction of an underadditive interaction in which the effect of stimulus quality is smaller (or absent) in the 0 SOA condition as compared to the −750 SOA condition. In contrast, if SOA and stimulus quality have additive effects on RT, then the inference is that processing of the word starts only after cue decoding and task set implementation have been completed in the 0 SOA condition.

The results were clear. There were large main effects of SOA and of stimulus quality (and less interestingly of task). Stimulus quality and task interacted; unsurprisingly, there was a smaller effect of stimulus quality in the letter case judgment task, where only a single letter had to be processed to judge letter case, than in the reading aloud task, where all letters had to be processed to read the nonword aloud. Critically, however, stimulus quality and SOA had additive effects on response time and on accuracy in both tasks.[3] This same critical pattern—additive effects of stimulus quality and SOA for both tasks—was also reported by Kahan et al. (2011), again when the targets were all nonwords.

In short, the interpretation of these experiments dovetails with that of the Oriet and Jolicoeur (2003) and Elchlepp et al. (2015) studies: A task set must be in place before even very early target identification begins. This is true despite the fact that the interpretation turns on the fact that switch/no-switch has additive effects with stimulus quality in the Oriet and Jolicoeur experiments and with word frequency and lexicality in the Elchlepp et al. studies, whereas in Besner and Care (2003) and Kahan et al. (2011), the additivity is between stimulus quality and SOA rather than switch/no switch. There is, consequently, some important generalization evident across these studies, but some other facts complicate this picture considerably.

Stimulus Set (Words vs. Nonwords)

The Oriet and Jolicoeur (2003) account proposed that there is a structural bottleneck. This becomes problematic when task switching is random (as in Besner & Care, 2003; Kahan et al., 2011) rather than predictably alternating, and when the manipulation is SOA rather than switch/no-switch. This is because when the targets, instead of being nonwords, are words, stimulus quality and SOA were underadditive in both the case judgment task and the reading aloud task. That is to say, the effect of stimulus quality was smaller at the zero SOA than it was at the long SOA when the cue appeared 750 ms before the target (Paulitzki et al., 2009, Experiment 1; Kahan et al., 2011, Experiment 2).

Beyond that, there are two conflicting findings in which word frequency was manipulated and either it did not yield an underadditive interaction with SOA (Paulitzki et al., 2009, Experiment 2; O'Malley & Besner, 2012) or it did yield an underadditive interaction (the two experiments in O'Malley & Besner, 2011). It is unclear why these different patterns are seen (despite replications of both patterns). Our speculation is that reader skill is important. This is certainly central in the psychological refractory period work: Lien et al. (2006) reported that, in the context of a lexical decision study, older readers with larger vocabularies showed underadditive effects of word frequency and SOA (the effect of word frequency decreased as SOA decreased), whereas university-level readers showed additive effects, appearing to need a task set in place before processing the target.

To summarize, when reading nonwords aloud or making case judgments about them, a task set must be in place before stimulus identification starts, given that stimulus quality and SOA are additive factors for both tasks. In contrast, provided that all of the targets are words, a task set does not always have to be in place before stimulus identification starts, as shown by an underadditive interaction of stimulus quality and SOA for both tasks.

There currently is no explanation for why this replicated dissociation across labs occurs between the joint effects of stimulus quality and SOA on the one hand and the lexical status of the target (high vs. low frequency words) on the other hand. Feature and letter identification are processes common to both words and nonwords, and stimulus quality surely affects both of these levels. All other things being equal, underadditivity of SOA and stimulus quality would therefore be expected regardless of whether the stimuli consisted of words or nonwords. Speculatively, there may be some

[3] It is good to be able to detect the two-way interaction of stimulus quality and task in the same experiment where stimulus quality and SOA are additive factors because this mitigates the concern that additive effects are merely Type II errors. Note also that there was no three-way interaction of stimulus quality, SOA, and task to complicate the pattern.

sort of "meta set" in play—invoked when the stimuli all are words. This warrants further investigation.

Go Versus No-Go (Words Only)

The last variation of the task-switching paradigm again involves the task varying randomly from trial to trial, depending on the pretrial cues. Following Besner and Risko (2005), Besner et al. (2021) provided university-level participants with a brief cue to make a response (Go)—to read aloud a single very high-frequency word—or to withhold the response (No-Go). The logic here was that the use of such high-frequency words should maximize the possibility of intention-free identification.

The Go/No-Go cue, with a cue–target SOA of 0, was incorporated into one block of trials; in another block of trials—the baseline block—all words were to be read aloud. Stimulus quality was again manipulated. The block where all words were to be read aloud acted as the baseline for determining the magnitude of the stimulus quality effect when the task set was already in place. In contrast, if participants in the Go versus No-Go block could start to process the target during the time that they were decoding the cue, then the effect of stimulus quality should have been smaller in that block than in the baseline block.

In Experiment 1, Besner et al. (2021) reported that the stimulus quality effect was the same size in both blocks. They inferred, therefore, that participants did not work on the target until the task cue was decoded and the task set was implemented. That is, once again, a task set had to be in place before word identification started (absorption into slack did not occur). Critically, the Go and No-Go probabilities in Experiment 1 were .5. In Experiment 2, where all of the experimental details were otherwise the same, the Go probability now was .8 and the No-Go probability now was .2. This experiment yielded an underadditive interaction in which the stimulus quality effect was smaller in the Go versus No-Go block than in the baseline block.

The simplest account is that a task set was implemented prior to stimulus identification in Experiment 1 (because stimulus quality and block type were additive factors). In contrast, stimulus identification was able to start during the time that the cue was being decoded in Experiment 2 (because the stimulus quality effect was smaller in the .80 Go condition than in the 100% baseline condition). The conclusion, again, is that there is no fixed relation between a task set and visual word identification, despite the use of only very high-frequency words as targets and university readers as participants, and despite what seems like a very simple decision about what to do on a given trial. Simply put, context matters in ways not predicted by any existing account of when a task set must precede stimulus identification and when it need not.

Stimulus Switching Without Task Switching

There are numerous experiments in the literature showing that context strongly affects how words are processed (e.g., Baluch & Besner, 1991; Tabossi & Laghi, 1992; Zevin & Balota, 2000). In those experiments, lexical versus sublexical processing is emphasized (by including exception words or nonwords, respectively). These results, of course, are not consistent with the idea that the target stimulus simply triggers processing such that it unfolds in the same way regardless of context, as an automaticity account would hold. Indeed, O'Malley and Besner (2008) and Besner et al. (2010) even challenged the widely accepted view that the processing of words is typically cascaded rather than being staged. Their data are consistent with a cascaded account of the relation between stimulus quality and word frequency when only words appear in the experiments, but staged processing of these two factors appears called for when nonwords are randomly interspersed in the list.

Below we review other work that illustrates how parameters of the task set change on the fly (across adjacent trials) despite the task—reading aloud—being held constant, illustrating just how dynamic processing can be, clearly not in line with an automaticity account.

The stimulus-switching paradigm is similar to the task-switching paradigm in its use of the alternating runs methodology (at least in the experiments described here), but only a single task is involved. What alternates between runs is the type of stimuli rather than the type of task. This approach was pioneered by Shafiullah and Monsell (1999) who examined whether there was a switch cost in the alternating sequence A1, A2, B1, B2, A1, A2, B1, B2, where A and B represented different types of stimuli. Specifically, is B1 slower than A2 and A1 slower than B2 due to a stimulus switch? The answer to this question turns out, once again, to depend—this time, on the type of stimuli.

The central argument here is that when a switch cost is seen despite the upcoming switch being perfectly predictable, participants must not have fully enabled the task set needed for the upcoming stimulus. Such accommodation occurs only after the switch stimulus has been encountered for the first time: Whatever the parameter set associated with the upcoming category is, these are modified to make that process more efficient after B1 (or A1) has been seen and worked on, hence the reduction in response time to the following B2 (or A2) trial. The task set makes processing more efficient within B or within A, but at the cost of reduced efficiency in moving from B2 to A1 or from A2 to B1. If it were simply the case that the participant adopted a general set that would allow all items to be processed equally efficiently, then no switch cost should be evident.

Switch Cost

Shafiullah and Monsell (1999) applied the above logic to reading Japanese aloud. Written Japanese has three scripts. Two Kana scripts—Hiragana and Katakana—are mora based (each mora character is associated with a syllable), with Katakana reserved for foreign loan words (e.g., "computer") and Hiragana used for grammatical morphemes. The third script, Kanji, is ideographic, with the whole character representing its meaning. Shafiullah and Monsell saw no cost when switching from one Kana script to the other, arguably because although the characters in these two scripts are physically different, both are processed using the same information-processing machinery—the application of sublexical spelling-sound correspondences. They did, however, observe a switch cost when switching from Kanji to Kana or vice versa, arguably because the information-processing machinery needed for these two scripts is different: The sublexical spelling-sound correspondences in play when generating a phonological code for Kana cannot be used when reading ideographic Kanji, and vice versa. In short, costs are seen when the upcoming trial requires the use of a different information processing routine, despite the fact that the type of the upcoming stimulus is 100% predictable.

The Dual-Route Cascaded Model

Related work has been reported by Reynolds and Besner (2005, 2008). The key to understanding their results lies in the theoretical framework that they used. Briefly, they adopted the dual route model of reading aloud (for items in English) that has been articulated by Coltheart et al. (2001) and implemented in a highly successful computational model—the dual-route cascaded (DRC) model. The two routes are the lexical and the sublexical routes. Along the lexical route, print activates (through feature and letter levels) the orthographic input lexicon, which consists of lexical entries, each of which corresponds to a word that readers know how to spell. Likewise, the phonological output lexicon consists of lexical entries corresponding to the whole word pronunciations for all of the words that the reader knows—the reader's lexicon. Connections from the orthographic input lexicon to the phonological output lexicon are word-specific.

The sublexical route, in contrast, employs spelling to sound correspondences (grapheme–phoneme rules) and is required to read aloud items that do not exist in the orthographic input lexicon (e.g., SLINT, DAP, ISH). Regular words (those whose spelling-sound correspondences are typical, such as TINT, LINT, SPLINT, and MINT) can be read aloud correctly via both the sublexical route and the lexical route. In contrast, exception words, defined as those whose pronunciations violate the typical spelling-to-sound correspondences (e.g., PINT as compared to TINT, LINT, SPLINT, and MINT) can only be read aloud correctly via the lexical route because the sublexical route would regularize them (i.e., it would read PINT so as to rhyme with TINT, LINT, SPLINT, and MINT).

The DRC model has been remarkably successful in simulating various benchmarks seen with skilled readers, as well as in being able to simulate both developmental and acquired forms of surface dyslexia (difficulty with whole word recognition, especially for irregular words) and phonological dyslexia (difficulty reading aloud nonwords due to sublexical phonological impairment). For present purposes, the critical issue concerns the weights reflected in the operation of the two routes—how strong one of the routes is compared to the other. If the weights are too strong on the sublexical route relative to the lexical route, then nonwords will be read aloud correctly, but exception words like PINT risk being regularized. In contrast, if the weights are too strong on the lexical route relative to the sublexical route, then exception words will be read aloud correctly, but nonwords run the risk of being lexicalized (read aloud as words). Consequently, the balance between these two routes is critical. Coltheart and colleagues paid close attention to this issue when implementing the DRC model and were able to find a set of weights that yielded equally accurate reading aloud of regular words, exception words, and nonwords, as well as yielding the correct ordering of response times for these three stimulus classes.

Reynolds and Besner (2005, 2008) reported experiments in which the task was held constant across trials: Participants read aloud a letter string on every trial. What varied predictably was the nature of the stimuli across trials. For example, in several experiments, two successive trials consisted of low-frequency exception words, and the next two trials consisted of nonwords, and this cycle repeated. If the routines needed to read aloud these different classes of stimuli relied on participants using a single set of weights (as implemented in the lexical and sublexical routines that yielded correct reading in the DRC model), then there would be no reason to expect that switching from one class of stimuli to the other would yield a switch cost from A2 to B1 (or from B2 to A1) nor a speed up from B1 to B2 (or from A1 to A2). The absence of a switch cost is the prediction that an automatic processing account of visual word recognition would endorse. On this account, word identification is simply stimulus triggered: There is no role for a task set during identification.

On the other hand, Reynolds and Besner (2005, 2008) proposed that it was possible to tune the relative weights on each of the routes to maximize efficiency. In essence, when B1 is encountered for the first time, the weights that maximized processing on A2 (an exception word, and hence an emphasis on the lexical route) would now serve to slow reading of B1 (a nonword), but, in anticipation of B2 (a second nonword), the weights could be altered so as to maximize processing along the sublexical route. This is exactly what Reynolds and Besner reported: When exception words (two successive trials) alternated with nonwords (two successive trials), B1 was slowed relative to A2, and B2 was faster than B1 (and A1 was slower than B2 and A2 was faster than A1) in Experiments 1–3. The argument is that this tuning was driven by the presentation of B1 rather than being done following A2 but before B1 was presented (ditto for A2 following A1).

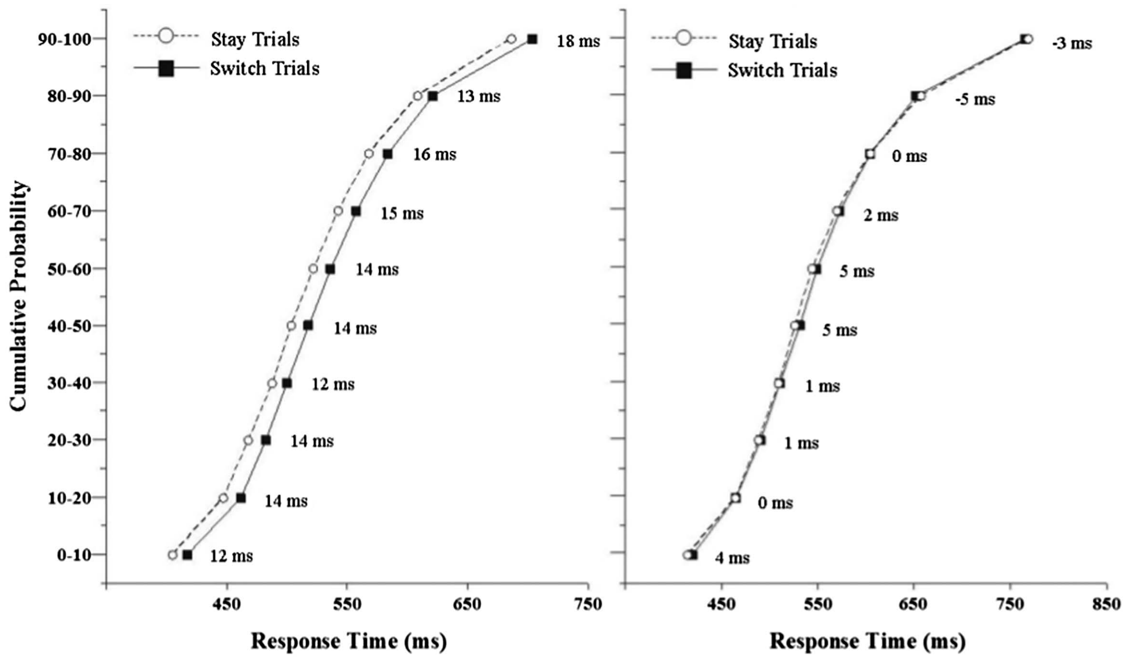
In another experiment, Reynolds and Besner (2008, Experiment 4), the stimuli alternated between exception words and regular words. Here, the expectation was that no tuning had to be applied because the default settings of the weights on the two routes would suffice given that the lexical route can correctly process both types of stimuli. Indeed, no switch cost was evident. In yet another experiment (Experiment 5), the stimuli consisted of regular words and nonwords. No switch cost was observed here either, consistent with the interpretation that participants relied to a large extent on the sublexical route to read aloud both types of stimuli. Data from all five of these experiments (where there was a switch cost and where there was not) can be seen in Figure 3. The data are entirely orderly in showing a switch cost (in Experiments 1–3; left panel), or the absence of such switch costs (in Experiments 4 and 5; right panel) across the response time distribution, just as predicted based on the dual routes in the DRC model.

These data are thus consistent with the interpretation that adjustments to the relative weights along the two routes are applied (tuning of the task set parameters) when two different stimulus classes (exception words vs. nonwords) each emphasize processing along one of the two routes, and that no such tuning need be applied when the default settings suffice to produce a pronunciation (as in the cases of exception words vs. regular words and regular words vs. nonwords). The interesting aspect of this task set influence is when it did not occur: It did not occur in anticipation of a stimulus class shift (e.g., from A2 to B1) but came into play only after encountering the first stimulus that changed class. We assume that an executive routine monitors the feedback from having generated a response on each trial and decides whether parameter changes are called for with respect to the upcoming trial.

Once again, this research points to greater complexity than a straightforward automaticity account would anticipate and, in particular, calls for a pivotal role for task set. How words are processed depends on the context in which they occur, and the fact that context is critical stands fundamentally in opposition to the idea

Figure 3

Distribution of Reading Aloud Times Under Conditions That Promote Switching (Left Panel) and That Do Not Promote Switching (Right Panel)



Note. The left panel depicts Experiments 1–3; the right panel depicts Experiments 4 and 5. These data appear as Figure 2 in "Contextual Effects on Reading Aloud: Evidence for Pathway Control," M. Reynolds and D. Besner, 2008, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 50–64 (https://doi.org/10.1037/0278-7393.34.1.50). Copyright 2008 by the American Psychological Association.

that visual word processing is automatic in the sense of not needing a task set.

Summary

We have covered considerable ground in reviewing experiments whose results bear on the issue of whether visual word identification is, as widely assumed, "automatic"—both in the specific sense of not needing any form of task set as a preliminary and more generally. Trying to bring order to these results is complicated in part because of the details of the different paradigms that we have considered, but more so because the results from these different paradigms are influenced by various forms of context not explained by any account except in a post hoc way (if at all). We therefore conclude with a brief recapitulation of these different lines of investigation, highlighting a few of the central findings from these paradigms for our argument about the central role played by task set as an implementation of intention, a perspective that contrasts sharply with the automaticity perspective.

Stroop

Despite the widespread view that the Stroop effect arises because readers are unable to prevent themselves from reading the word, a large number of observations undercut that view. Various factors were reviewed, each of which reduced or eliminated a Stroop effect; for example, salience of the word relative to the colour (Sabri et al., 2001), various forms of instructions (e.g., Bauer & Besner, 1997), and hypnosis (e.g., Raz et al., 2002). It is also the case that readers do follow the instructions to avoid "reading" the irrelevant word to a considerable extent in that they rarely make a mistake and read the incongruent word aloud.

Readers also make (unconscious) use of the typical correlation between word and colour such that they identify the word because it helps them to identify the colour. These contingencies are learned quickly, extinguished quickly, and relearned quickly too (see MacLeod, 2019; Schmidt, 2021; Schmidt & Besner, 2008; Schmidt et al., 2007, 2010). In short, the words, although nominally irrelevant, are always being processed because there is often a payoff to doing so: They are constantly being considered as data that might be relevant to determining what the colour is.

Importantly, the distribution of spatial attention is also critical. The default set is to spread it across the letters in a word (e.g., when the task is explicit, as in lexical decision; see McCann et al., 1992), but this mental set is easily overridden by drawing the reader's attention to individual letters by colouring only a single letter and providing spatial cues; this procedure serves to reduce or eliminate the Stroop effect (Besner, 2001; Besner & Stolz, 1999). This is also true for the semantic Stroop task variant using colour-related words, although this claim is more controversial (see Besner & Young, 2024; Labuschagne & Besner, 2015; Kahan et al., 2006; Manwell et al., 2004 vs. Augustinova & Ferrand, 2014; Burca et al., 2021). The need for spatial attention has also been well documented in tasks other than Stroop (e.g., lexical decision and reading aloud; e.g., see

Besner et al., 2005; Lachter et al., 2004; McCann et al., 1992). Other work shows that spatial attention can be set to be wide or narrow in response to various changes in context (e.g., Besner et al., 2005; Waechter et al., 2011).

Nonword Versus Word Processing

The data derived from nonword processing are quite clear: They support the conclusion that a task set must be in place before very early processing can begin (for task set with multiple tasks see Besner & Care, 2003; Kahan et al., 2011; Paulitzki et al., 2009; for letter search on the prime, see Ferguson & Besner, 2006). We know of no exceptions to the claim that the processing of nonwords across these paradigms requires a task set to be in place as a preliminary to target processing.

The evidence from the processing of words is more mixed. Word processing (and Arabic numeral processing) is sometimes delayed until a task set is in place. For experiments in which switch time from one task to another was the measure of interest, see Oriet and Jolicoeur (2003) for parity and odd/even judgments; for words in semantic categorization and lexical decision tasks, see Elchlepp et al. (2015). When the measure of interest involves SOA, for case judgments and reading aloud, see Paulitzki et al. (2009, Experiment 2), Kahan et al. (2011), and O'Malley and Besner (2012); for reading aloud in the context of the Go-No-Go paradigm, when the Go probability is 50%, see Besner et al. (2021). On the other hand, when the measure of interest is the effect of SOA, there is some evidence for the processing of words being able to begin without a task set being in place. For reading aloud words and making case judgments, see the effect of stimulus quality (Paulitzki et al., 2009, Experiment 1); for the effect of word frequency when reading aloud, see O'Malley and Besner (2011, Experiments 1 and 2). See also Besner et al. (2021) with regard to the Go-No-Go paradigm when the Go probability is .80.

Why are these different outcomes observed across experiments? Some appear to be context-driven (see Besner et al.'s, 2021, discussion with regard to the Go-No-Go paradigm). But for others, we imagine that reader skill plays an important but neglected role as an individual difference factor has never been studied in any of the contexts noted here. Lower frequency words might not exist in the orthographic lexicon, or may be less firmly established there, necessitating the use of sublexical spelling-sound correspondences to address phonological lexical memory. As we have seen, the use of such correspondences requires that a task set be in place. Failures to see the processing of words in the absence of a task set being in place could be attributed to lower levels of reader skill. Further study of reader skill in the context of the paradigms considered here is clearly warranted. Individual differences in complex skills like reading are unlikely to be simply "error variance" (see, e.g., Palmer et al., 1985).

Stimulus Switching

In this paradigm, the task is held constant across the experiment, and what varies is the nature of the stimuli across alternating pairs of trials (e.g., as in exception word [A1], exception word [A2], nonword [B1], nonword [B2]). The results of this small literature are flatly inconsistent with a strong claim that a task set typically is not in play. Instead, B1 typically is slower than A2, and following (theoretically motivated) parameter adjustments, B2 is confirmed to be faster than B1, as is A2 following A1. Additional theoretical predictions as to when no switch costs would be expected were also confirmed (Reynolds & Besner, 2008; Shafiullah & Monsell, 1999). Contrary to an automaticity account, these data illustrate that a task set and visual word recognition are intimately intertwined.

Conclusions

These findings from different paradigms make it clear that there is no simple association between an intention in the form of a task set as a preliminary to visual word identification and how and when the word will be read. Our selective review can thus be considered both as brush clearing and as a challenge. We have looked hard for consistent evidence that stimulus identification can start in the absence of a task set and have generally come up almost empty-handed.

Where, then, do we stand after all this? We have focused here on the role of an intention in the form of a task set with regard to early processes in visual word identification. It seems to us uncontentious to suggest that the field is likely to make more progress if investigators are less dogmatic in asserting that visual word identification is "automatic" in various senses beyond the need for a task set. Indeed, there are many processes involved in "identifying" a single well-formed letter string. It strains the imagination (at least ours) to think that all of them are automatic, and that they are automatic in the same way (see Moors & De Houwer, 2006 for extended discussion on this point). Instead, we should concern ourselves with trying to understand how different contexts, stimuli, and especially reader skill produce the effects that they do. We recognize that this is a tall order, but without paying attention to these issues, we see genuine progress as being hampered. One thing seems certain: We will have to learn to live with considerably more complexity than has been acknowledged to date.

Résumé

L'une des distinctions les plus fondamentales en psychologie cognitive est celle entre le traitement « contrôlé » et le traitement « automatique ». Le concept du traitement automatique de l'identification visuelle des mots, largement répandu, affirme que, parmi d'autres caractéristiques, la présentation d'une chaîne de lettres bien formée déclenche une activation sublexicale, lexicale et sémantique en l'absence de toute intention de le faire. Au contraire, le rôle de l'intention est considéré comme indépendant de l'identification du stimulus et limité à la sélection d'une action utilisant les produits de l'identification (par exemple, freiner en apercevant un panneau indiquant « PONT EN RÉFECTION »). Nous examinons quatre paradigmes concernant le rôle d'une intention – définie ici comme un « ensemble de tâches » indiquant comment agir dans la situation actuelle – lors de l'identification de chaînes de lettres simples et bien formées. Contrairement à l'idée reçue de l'automaticité, la littérature concernant chacun de ces paradigmes démontre que la relation entre une intention et l'identification d'un stimulus est limitée de multiples façons, dont beaucoup ne sont pas bien comprises à l'heure actuelle. Une chose est claire : il n'existe pas de relation simple entre une intention, sous la forme d'un ensemble de tâches, et l'identification du stimulus. Le traitement automatique des mots, s'il a lieu, n'est certainement pas un défaut du système.

Mots-clés : identification visuelle des mots, intention, ensemble de tâches, automaticité

References

Algom, D., & Chajut, E. (2019). Reclaiming the Stroop effect back from control to input-driven attention and perception. *Frontiers in Psychology*, *10*, Article 1683. https://doi.org/10.3389/fpsyg.2019.01683

Algom, D., & Fitousi, D. (2016). Half a century of research on Garner interference and the separability-integrality distinction. *Psychological Bulletin*, *142*(12), 1352–1383. https://doi.org/10.1037/bul0000072

Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance 15: Conscious and nonconscious information processing* (pp. 421–452). The MIT Press.

Augustinova, M., & Ferrand, L. (2012a). Suggestion does not de-automatize word reading: Evidence from the semantically based Stroop task. *Psychonomic Bulletin & Review*, *19*(3), 521–527. https://doi.org/10.3758/s13423-012-0217-y

Augustinova, M., & Ferrand, L. (2012b). The influence of mere social presence on Stroop interference: New evidence from the semantically-based Stroop task. *Journal of Experimental Social Psychology*, *48*(5), 1213–1216. https://doi.org/10.1016/j.jesp.2012.04.014

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/10.1177/0963721414540169

Baluch, B., & Besner, D. (1991). Visual word recognition: Evidence for strategic control of lexical and nonlexical routines in oral reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(4), 644–652. https://doi.org/10.1037/0278-7393.17.4.644

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

Becker, C. A. (1976). Allocation of attention during visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(4), 556–566. https://doi.org/10.1037/0096-1523.2.4.556

Besner, D. (2001). The myth of ballistic processing: Evidence from Stroop's paradigm. *Psychonomic Bulletin & Review*, *8*(2), 324–330. https://doi.org/10.3758/BF03196168

Besner, D., & Care, S. (2003). A paradigm for exploring what the mind does while deciding what it should do. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *57*(4), 311–320. https://doi.org/10.1037/h0087434

Besner, D., McLean, D., Young, T., & Risko, E. (2021). Intention and performance when reading aloud: Context is everything. *Consciousness and Cognition*, *95*, Article 103211. https://doi.org/10.1016/j.concog.2021.103211

Besner, D., O'Malley, S., & Robidoux, S. (2010). On the joint effects of stimulus quality, regularity, and lexicality when reading aloud: New challenges. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(3), 750–764. https://doi.org/10.1037/a0019178

Besner, D., & Reynolds, M. (2017). Is semantic activation from print capacity limited? Evidence from the psychological refractory period paradigm. *Psychonomic Bulletin & Review*, *24*(3), 907–913. https://doi.org/10.3758/s13423-016-1178-3

Besner, D., Reynolds, M., & O'Malley, S. (2009). When underadditivity of factor effects in the Psychological Refractory Period paradigm implies a bottleneck: Evidence from psycholinguistics. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *62*(11), 2222–2234. https://doi.org/10.1080/17470210902747187

Besner, D., & Risko, E. F. (2005). Stimulus–response compatible orienting and the effect of an action not taken: Perception delayed is automaticity denied. *Psychonomic Bulletin & Review*, *12*(2), 271–275. https://doi.org/10.3758/BF03196371

Besner, D., Risko, E. F., & Sklair, N. (2005). Spatial attention as a necessary preliminary to early processes in reading. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *59*(2), 99–108. https://doi.org/10.1037/h0087465

Besner, D., Risko, E. F., Stolz, J. A., White, D., Reynolds, M., O'Malley, S., & Robidoux, S. (2016). Varieties of attention: Their roles in visual word identification. *Current Directions in Psychological Science*, *25*(3), 162–168. https://doi.org/10.1177/0963721416639351

Besner, D., & Stolz, J. A. (1999). What kind of attention modulates the Stroop effect? *Psychonomic Bulletin & Review*, *6*(1), 99–104. https://doi.org/10.3758/BF03210815

Besner, D., Stolz, J. A., & Boutilier, C. (1997). The stroop effect and the myth of automaticity. *Psychonomic Bulletin & Review*, *4*(2), 221–225. https://doi.org/10.3758/BF03209396

Besner, D., & Young, T. (2024). *Directed spatial attention derails a semantic contribution to a Stroop effect* [Manuscript submitted for publication].

Broadbent, D. E. (1958). *Perception and communication*. Pergamon Press. https://doi.org/10.1037/10037-000

Brown, T. L., Gore, C. L., & Carr, T. H. (2002). Visual attention and word recognition in stroop color naming: Is word recognition "automatic"? *Journal of Experimental Psychology: General*, *131*(2), 220–240. https://doi.org/10.1037/0096-3445.131.2.220

Bryan, W. L., & Harter, N. (1899). Studies on the telegraphic language: The acquisition of a hierarchy of habits. *Psychological Review*, *6*(4), 345–375. https://doi.org/10.1037/h0073117

Burca, M., Beaucousin, V., Chausse, P., Ferrand, L., Parris, B. A., & Augustinova, M. (2021). Is there semantic conflict in the Stroop task? Further evidence from a modified two-to-one Stroop paradigm combined with single-letter coloring and cueing. *Experimental Psychology*, *68*(5), 274–283. https://doi.org/10.1027/1618-3169/a000530

Cattell, J. M. (1886). The time it takes to see and name objects. *Mind*, *os-XI*(41), 63–65. https://doi.org/10.1093/mind/os-XI.41.63

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, *97*(3), 332–361. https://doi.org/10.1037/0033-295X.97.3.332

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256. https://doi.org/10.1037/0033-295X.108.1.204

Elchlepp, H., Lavric, A., & Monsell, S. (2015). A change of task prolongs early processes: Evidence from ERPs in lexical tasks. *Journal of Experimental Psychology: General*, *144*(2), 299–325. https://doi.org/10.1037/a0038740

Ferguson, R., & Besner, D. (2006). Basic processes in reading: Can functional phonological recoding be blocked? *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *60*(2), 148–158. https://doi.org/10.1037/cjep2006014

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

Herdman, C. M. (1992). Attentional resource demands of visual word recognition in naming and lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(2), 460–470. https://doi.org/10.1037/0096-1523.18.2.460

Jacoby, L. L., Lindsay, D. S., & Hessels, S. (2003). Item-specific control of automatic processes: Stroop process dissociations. *Psychonomic Bulletin & Review*, *10*(3), 638–644. https://doi.org/10.3758/BF03196526

Kahan, T. A., Hengen, K. B., & Mathis, K. M. (2011). An examination of orthographic and phonological processing using the task-choice procedure. *Language and Cognitive Processes*, *26*(2), 171–190. https://doi.org/10.1080/01690961003752355

Kahan, T. A., Sellinger, J. J., & Broman-Fulks, J. J. (2006). Associative and phonological priming effects after letter search on the prime. *American Journal of Psychology*, *119*(2), 239–254. https://doi.org/10.2307/20445337

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*(2), 293–332. https://doi.org/10.1016/0010-0285(74)90015-2

Labuschagne, E. M., & Besner, D. (2015). Automaticity revisited: When print doesn't activate semantics. *Frontiers in Psychology*, *6*, Article 117. https://doi.org/10.3389/fpsyg.2015.00117

Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty-five years after Broadbent (1958): Still no identification without attention. *Psychological Review*, *111*(4), 880–913. https://doi.org/10.1037/0033-295X.111.4.880

Lien, M.-C., Allen, P. A., Ruthruff, E., Grabbe, J., McCann, R. S., & Remington, R. W. (2006). Visual word recognition without central attention: Evidence for greater automaticity with advancing age. *Psychology and Aging*, *21*(3), 431–447. https://doi.org/10.1037/0882-7974.21.3.431

Lin, O. Y.-H., & MacLeod, C. M. (2018). The acquisition of simple associations as observed in color-word contingency learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(1), 99–106. https://doi.org/10.1037/xlm0000436

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

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*(2), 163–203. https://doi.org/10.1037/0033-2909.109.2.163

MacLeod, C. M. (2019). Learning simple associations. *Canadian Psychology/Psychologie canadienne*, *60*(1), 3–13. https://doi.org/10.1037/cap0000170

MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(1), 126–135. https://doi.org/10.1037/0278-7393.14.1.126

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*(3), 458–462. https://doi.org/10.3758/BF03196595

Marcel, A. J. (1983). Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes. *Cognitive Psychology*, *15*(2), 238–300. https://doi.org/10.1016/0010-0285(83)90010-5

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

Melara, R. D., & Mounts, J. R. (1993). Selective attention to Stroop dimensions: Effects of baseline discriminability, response mode, and practice. *Memory & Cognition*, *21*(5), 627–645. https://doi.org/10.3758/BF03197195

Monsell, S. (2017). Task set regulation. In T. Egner (Ed.), *The Wiley handbook of cognitive control* (pp. 29–49). Wiley. https://doi.org/10.1002/9781118920497.ch2

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, *132*(2), 297–326. https://doi.org/10.1037/0033-2909.132.2.297

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, 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

O'Malley, S., & Besner, D. (2008). Reading aloud: Qualitative differences in the relation between stimulus quality and word frequency as a function of context. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(6), 1400–1411. https://doi.org/10.1037/a0013084

O'Malley, S., & Besner, D. (2011). Lexical processing while deciding what task to perform: Reading aloud in the context of the task set paradigm. *Consciousness and Cognition*, *20*(4), 1594–1603. https://doi.org/10.1016/j.concog.2011.08.006

O'Malley, S., & Besner, D. (2012). Reading aloud and the question of intent. *Consciousness and Cognition*, *21*(3), 1298–1310. https://doi.org/10.1016/j.concog.2012.06.011

Oriet, C., & Jolicoeur, P. (2003). Absence of perceptual processing during reconfiguration of task set. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(5), 1036–1049. https://doi.org/10.1037/0096-1523.29.5.1036

Paap, K. R., & Herdman, C. M. (1998). Highly skilled participants and failures to redirect attention: Two plausible reasons for failing to replicate Paap and Noel's effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(4), 845–861. https://doi.org/10.1037/0278-7393.24.4.845

Palmer, J., MacLeod, C. M., Hunt, E., & Davidson, J. E. (1985). Information processing correlates of reading. *Journal of Memory and Language*, *24*(1), 59–88. https://doi.org/10.1016/0749-596X(85)90016-6

Pashler, H. (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(3), 358–377. https://doi.org/10.1037/0096-1523.10.3.358

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

Pashler, H. (2000). Task switching and multitask performance. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 275–423). The MIT Press.

Paulitzki, J. R., Risko, E. F., O'Malley, S., Stolz, J. A., & Besner, D. (2009). On the role of set when reading aloud: A dissociation between prelexical and lexical processing. *Consciousness and Cognition*, *18*(1), 135–144. https://doi.org/10.1016/j.concog.2008.11.007

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review*, *114*(2), 273–315. https://doi.org/10.1037/0033-295X.114.2.273

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

Posner, M. I., & Snyder, C. R. (1975). Facilitation and inhibition in the processing of signals. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance* (pp. 669–682). Academic Press. https://cir.nii.ac.jp/crid/1573387450013996288

Rastle, K., & Coltheart, M. (2006). Is there serial processing in the reading system; and are there local representations? In S. Andrews (Ed.), *From ink marks to ideas* (pp. 33–54). Psychology Press.

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*, *12*(3), 332–346. https://doi.org/10.1016/S1053-8100(03)00024-2

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

Reynolds, M., & Besner, D. (2005). Contextual control over lexical and sublexical routines when reading english aloud. *Psychonomic Bulletin & Review*, *12*(1), 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

Risko, E. F., Stolz, J. A., & Besner, D. (2005). Basic processes in reading: Is visual word recognition obligatory? *Psychonomic Bulletin & Review*, *12*(1), 119–124. https://doi.org/10.3758/BF03196356

Robidoux, S., & Besner, D. (2011). When benefits outweigh costs: Reconsidering "automatic" phonological recoding when reading aloud. *Canadian Journal of Experimental Psychology*, *65*(2), 105–108. https://doi.org/10.1037/a0021642

Rogers, R., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, *124*(2), 207–231. https://doi.org/10.1037/0096-3445.124.2.207

Sabri, M., Melara, R. D., & Algom, D. (2001). A confluence of contexts: Asymmetric versus global failures of selective attention to stroop dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(3), 515–537. https://doi.org/10.1037/0096-1523.27.3.515

Schmidt, J. R. (2021). Apprentissage incident des associations simples de stimulus-réponse: Revue de la recherche avec la tâche d'apprentissage de contingences couleur-mot [Incidental learning of simple stimulus–response associations: Review of research using the color-word contingency learning task]. *L'Année Psychologique*, *121*(2), 77–127. https://doi.org/10.3917/anpsy1.212.0077

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

Schmidt, J. R., Crump, M. J. C., Cheesman, J., & Besner, D. (2007). Contingency learning without awareness: Evidence for implicit control. *Consciousness and Cognition*, *16*(2), 421–435. https://doi.org/10.1016/j.concog.2006.06.010

Schmidt, J. R., Houwer, J. D., & Besner, D. (2010). Contingency learning and unlearning in the blink of an eye: A resource dependent process. *Consciousness and Cognition*, *19*(1), 235–250. https://doi.org/10.1016/j.concog.2009.12.016

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66. https://doi.org/10.1037/0033-295X.84.1.1

Shafiullah, M., & Monsell, S. (1999). The cost of switching between Kanji and Kana while reading Japanese. *Language and Cognitive Processes*, *14*(5–6), 567–607. https://doi.org/10.1080/016909699386194

Sharma, D., Booth, R., Brown, R., & Huguet, P. (2010). Exploring the temporal dynamics of social facilitation in the Stroop task. *Psychonomic Bulletin & Review*, *17*(1), 52–58. https://doi.org/10.3758/PBR.17.1.52

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), 127–190. https://doi.org/10.1037/0033-295X.84.2.127

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

Tabossi, P., & Laghi, L. (1992). Semantic priming in the pronunciation of words in two writing systems: Italian and English. *Memory & Cognition*, *20*(3), 303–313. https://doi.org/10.3758/BF03199667

Tzelgov, J., Henik, A., & Berger, J. (1992). Controlling Stroop effects by manipulating expectations for color words. *Memory & Cognition*, *20*(6), 727–735. https://doi.org/10.3758/BF03202722

Tzelgov, J., Henik, A., & Leiser, D. (1990). Controlling Stroop interference: Evidence from a bilingual task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(5), 760–771. https://doi.org/10.1037/0278-7393.16.5.760

Van Orden, G. C., Stone, G. O., Garlington, K. L., Markson, L. R., Pinnt, G. S., Simonfy, C. M., & Brichetto, T. (1992). "Assembled" phonology and reading: A case study in how theoretical perspective shapes empirical investigation. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 249–292). North-Holland. https://doi.org/10.1016/S0166-4115(08)62799-5

Waechter, S., Besner, D., & Stolz, J. A. (2011). Basic processes in reading: Spatial attention as a necessary preliminary to orthographic and semantic processing. *Visual Cognition*, *19*(2), 171–202. https://doi.org/10.1080/13506285.2010.517228

White, D., & Besner, D. (2018). Attentional constraints on semantic activation: Evidence from Stroop's paradigm. *Acta Psychologica*, *189*, 4–11. https://doi.org/10.1016/j.actpsy.2016.08.008

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

Received May 3, 2023
Revision received December 14, 2023
Accepted December 16, 2023 ■