



**Is there semantic conflict in the Stroop task ?
Investigation of distinct components of the Stroop effect
in the two-to-one Stroop paradigm**

Mariana Burca

► **To cite this version:**

Mariana Burca. Is there semantic conflict in the Stroop task ? Investigation of distinct components of the Stroop effect in the two-to-one Stroop paradigm. Psychology. Normandie Université, 2022. English. NNT : 2022NORMR041 . tel-03894640

HAL Id: tel-03894640

<https://theses.hal.science/tel-03894640>

Submitted on 12 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Normandie Université

THÈSE

Pour obtenir le diplôme de doctorat

Spécialité PSYCHOLOGIE

Préparée au sein de l'Université de Rouen Normandie

Is There Semantic Conflict in the Stroop Task ? Investigation of Distinct Components of the Stroop effect in the Two-to-One Stroop Paradigm

**Présentée et soutenue par
MARIANA BURCA**

**Thèse soutenue le 13/06/2022
devant le jury composé de**

M. JAMES SCHMIDT	PROFESSEUR DES UNIVERSITES, UNIVERSITE DE BOURGOGNE	Rapporteur du jury
M. PIERPAOLO IODICE	MAITRE DE CONFERENCES HDR, Université de Rouen Normandie	Membre du jury
M. FRANCOIS MAQUESTIAUX	PROFESSEUR DES UNIVERSITES, UNIVERSITE BESANCON FRANCHE COMTE	Président du jury
MME MARIA AUGUSTINOVA	PROFESSEUR DES UNIVERSITES, Université de Rouen Normandie	Directeur de thèse

Thèse dirigée par MARIA AUGUSTINOVA (CENTRE DE RECHERCHE SUR LES FONCTIONNEMENTS ET LES DYSFONCTIONNEMENTS PSYCHOLOGIQUES)

ABSTRACT

By assessing whether semantic (in addition to response) conflict contributes to the Stroop interference effect observed in the two-to-one Stroop paradigm (De Houwer, 2003), the present thesis was aimed at contributing to the ongoing debate about the locus (vs. loci) of the Stroop interference effect. To this end, the two-to-one Stroop paradigm administered with an additional (i.e., color-neutral) baseline was combined with Single-letter coloring and cueing (SLCC) in Chapter 1 and Chapter 4; with healthy aging in Chapter 2; and with variations in Response-stimulus intervals in Chapters 3 and 4. In complete contrast to dominant single-stage response competition models of the Stroop interference effect, results of experiments reported in this dissertation leave no doubt that selection occurs at the semantic level in the Stroop task. The immediate implications for the composite (as opposed to unitary) nature of the Stroop interference effect and other still unresolved issues in the Stroop literature are discussed.

Key words: Stroop Interference Effect, Facilitation, Interference, Semantic and Response Conflict

RESUME

En examinant dans quelle mesure le conflit sémantique (en plus de celui de réponse) contribue dans l'effet d'interférence Stroop observé avec le paradigme Stroop dit two-to-one (De Houwer, 2003), la présente thèse vise à contribuer au débat scientifique – actuellement en cours – sur le caractère unitaire ou possiblement composite de l'interférence Stroop. Pour ce faire, le paradigme Stroop two-to-one – administré avec une condition contrôle supplémentaire – a été combiné avec des modérateurs tels que la coloration et l'indication spatiale d'une seule lettre (Chapitre 1 et Chapitre 4), le vieillissement normal (Chapitre 2) et avec les variations d'intervalles entre la réponse et le nouveau stimulus (Chapitre 3 et Chapitre 4). En parfait contraste avec les modèles unitaires de l'effet d'interférence Stroop – modèles actuellement dominant dans la littérature Stroop – les résultats des expériences rapportés dans cette thèse ne laissent aucun doute que la sélection opère dans la tâche Stroop au niveau sémantique. Les implications pour le caractère composite de l'interférence Stroop, tout comme pour d'autres questions actuellement sans réponse dans la littérature Stroop sont discutées.

Mots clefs : L'effet d'interférence Stroop, Facilitation, Interférence, Conflit sémantique, Conflit de réponse



Acknowledgements

First, I would like to thank the jury that took the time to evaluate my dissertation: Pr. François Maquestiaux, Pr. James Schmidt, and (associate) Pr. PierPaolo Iodice. Thank you for accepting to be part of my defense committee, it is an honor for me.

I want to thank my advisor, Pr. Maria Augustinova. Working under your supervision, has taught me a lot, both on a professional and personal level. I am grateful for all the advice and guidance I received from you. The patience that you had with me, my silly mistakes, and periods of stress, are highly and exceedingly appreciated. Doing a PhD is not always a joyful ride, but I am glad and proud I had you as an advisor. Without you, I am not sure I would have completed this journey. Thank you for always finding the time for me, for teaching me what research is, thank you for the personal insights, and advice. I have grown up and got to know myself better, both professionally and personally, and that is thanks to you. I am glad I shared this PhD adventure with you.

I am grateful and proud to be a part of the OptiCog project (funded by ANR and Normandy Region that also funded my PhD). Working with Ludovic Ferrand (Directeur de recherche, HDR at LAPSCO, Université de Clermont-Auvergne), Dr. Ben Parris (Assistant Professor at Bournemouth University, UK) and Pierre Chausse (research engineer at LAPSCO, Université de Clermont-Auvergne) was an amazing experience. Thus, I would like to thank you all for the fruitful collaboration we had so far. I am extremely happy I had the chance to participate in scientific talks, meetings and debates with all of you. I got to learn from you, and your expertise, and that is priceless.

I would, also, like to thank my mother, Tamara Burca, you have passed through this PhD journey together with me. Thank you for your encouragements, and for believing in me, it meant the world to me. Even though, you were far away, you were always ready to help, support and reassure. Thank you for everything, there are no words that could express how much you did and mean for me.

As well, I would like to thank all my friends. First of all, my PhD fellows, that understood the struggles that come during a PhD, and were always there for me: Clémence, Léa, Laura, Marion, Camille and François. Thank you for picking me up when I was down, but most of all thank you for all the laughs we shared, and the good times that helped us all to get through. Now, I can proudly say: “Dobby is a free elf” (cf., Harry Potter).

Certainly, I want to thank as well my friends from Moldova: Victor, Sergiu, Cezar, Kira, Lida, Lenuta and Cris. You never got tired to encourage me, thank you for believing in me.

The last but not the least, I would like to thank myself, for listening to all your advices and for not giving up.

Everything was beautiful and nothing hurt. (Kurt Vonnegut, “Slaughterhouse-Five”, 1969).

CONTENTS

General Introduction.....	6
Locus vs. loci of Stroop interference	6
How to assess the contribution of a semantic conflict to the Stroop interference?	11
Overview of the present work	16
References.....	18
Chapter One: 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.....	23
1. Introduction.....	23
The present study	26
2. Method	29
Participants.....	29
Design	29
Stimuli.....	29
Apparatus and Procedure	29
3. Results and Discussion	31
Is there semantic conflict in the Stroop task?.....	33
How does SLCC affect semantic conflict?	35
4. Discussion and Conclusions.....	36
5. References.....	40
6. Supplementary Materials	44
Chapter Two: Some further clarifications on age-related differences in the Stroop task: new evidence from the two-to-one Stroop paradigm 48	
1. Introduction.....	48

The Present Study	52
2. Method	52
Participants and Design.....	52
Apparatus, Stimuli and Procedure.....	53
3. Results and Discussion	55
Is there any Semantic Conflict in the two-to-one Stroop paradigm?	56
How does healthy aging influence semantic vs. response conflict in the Stroop task?	58
4. General Discussion and Conclusion.....	59
5. References.....	64
6. Supplementary Materials	69
7. References for Supplementary Materials	79
Chapter Three: Further investigation of distinct components of Stroop interference and of their reduction by short response-stimulus intervals: new evidence from the two-to-one Stroop paradigm.	81
1. Introduction.....	81
At what level of processing the benefit of a short (as opposed to long) RSI takes its effect?	82
Is there any semantic conflict in the Stroop task?	86
The Present study.....	89
2. Method	91
Participants and Design.....	91
Procedure, apparatus, stimuli and design	91
3. Results and discussion	93
Is there a semantic conflict in the 2-to-1 paradigm?	94
At what level of processing the benefit of a short (as opposed to long) RSI takes its effect?	97
4. General Discussion and Conclusion.....	99
5. References.....	104

Chapter Four: How to effectively improve older adults' performance in the Stroop task?.....	109
1. Introduction.....	109
2. Experiment 1: The influence of variations in response-stimulus intervals on the elderly's performance in the Stroop task	110
2.1 Method	112
Participants and Inclusion process and criteria	112
Procedure, apparatus, stimuli and design.	113
2.2 Results and Discussion	114
Is there semantic conflict in the Stroop task?.....	116
Does RSI manipulation improve older adults' performance in the Stroop task?.....	117
3.1 Experiment 2: The influence of Single letter coloring and cueing	118
3.2 Method	122
Participants and Inclusion process and criteria	122
Procedure, Apparatus, Stimuli and Design	122
Is there semantic conflict in the Stroop task?.....	123
Does SLCC manipulation improve older adults' performance in the Stroop task?.....	125
4. General Discussion and Conclusion.....	127
5. References.....	132
6. Supplementary Materials	140
General Discussion.....	148
The overall Stroop effect is a composite phenomenon	151
The Stroop interference effect is itself a composite phenomenon.....	155
Semantic conflict is permeable to moderators	158
Concluding remarks	162
References.....	164

GENERAL INTRODUCTION

In the Stroop Task – construed by John Ridley Stroop in 1935 –, individuals are asked to identify the ink color of different stimuli. Perhaps the most known stimuli are color words presented in an ink color that is different from that the word actually designates (e.g., the word BLUE presented in a green font, hereafter $BLUE_{\text{green}}$). Given this, they are often designated as *color-incongruent* Stroop stimuli. Typically, individuals' color-identification times are substantially longer and often more error-prone for these latter stimuli as compared to a relevant comparison or baseline stimuli (e.g., color patches, strings of symbols or letters, words that do not depict a color etc., presented in a green font). This difference – already shown in Stroop's own dissertation – is now designated in the literature as the *Stroop interference effect* or the *Stroop effect*. Over the years, it became one of the most widely used indicators of selective attention and of inhibitory abilities and this in both clinical and research practice (152.000 results in Google ScholarTM, on 3rd of March 2022). Indeed, it is now considered as the “gold standard of attentional measurers” (MacLeod, 1992). Despite this, cognitive processes that are mobilized in the Stroop task are still rather poorly understood as the still ongoing debate about the actual locus or loci of Stroop interference indicates.

Locus vs. loci of Stroop interference

The aforementioned longer response times (hereafter RTs) and more frequent errors observed for color-incongruent (e.g. $BLUE_{\text{green}}$) than for color-neutral stimuli (e.g., the word PAGE presented in a green font, hereafter $PAGE_{\text{green}}$) were initially interpreted by Stroop, (1935) as a result *interference* between color-naming and word-reading. He reasoned that « The associations that have been formed between the word stimuli and the reading response

are evidently more effective than those that have been formed between the color stimuli and the naming response. Since these associations are products of training, and since the difference in their strength corresponds roughly to the difference in training in reading words and naming colors, it seems reasonable to conclude that the difference in speed in reading names of colors and in naming colors may be satisfactorily accounted for by the difference in training in the two activities. » (Stroop, 1935b, pp. 659-660).

In line with Stroop's initial reasoning, the so-called *relative speed hypothesis* (e.g., Morton & Chambers, 1973; see also Dyer, 1973; Fraisse, 1969; Hommel, 1997) argues that reading is indeed faster than color naming such that the meaning of the stimulus is processed before the ink color that has to be named. Because the word-dimension of an incongruent Stroop stimulus – that is processed first – provides evidence towards a response (i.e., blue for *BLUE_{green}*) that is different from that subsequently triggered by the (relevant) color-dimension of this stimulus (i.e., green for *BLUE_{green}*) and because there is a *bottleneck* in attention (see (Posner & Snyder, 1975 for understanding this metaphor; and e.g., Maquestiaux, 2017 for a review of its more recent developments) a queue (and subsequent delay or interference) occurs at the level of response choice/production. According to the so-called *automaticity hypothesis* (e.g., (Hunt & Lansman, 1986; LaBerge & Samuels, 1974; Logan, 1978; Posner & Snyder, 1975; Shiffrin & Schneider, 1977), the interference does not result from relative differences in speed but rather these undeniable differences in speed reflect the fact that reading is more automatic than color-naming.

On the basis of these latter ideas, although placing the explanatory weight on parallel processing of the irrelevant and the relevant dimensions (as opposed to a single bottleneck in attention), many prominent accounts of the Stroop effect (e.g., Cohen et al., 1990; Glaser & Glaser, 1989; Logan & Zbrodoff, 1979; Melara & Algom, 2003; Phaf et al., 1990; Roelofs, 2003) subsequently proposed that Stroop interference occurs late in processing (i.e. at the

level of response output). These accounts are therefore termed *single-stage late selection accounts*. In this view that still dominates Stroop literature, Stroop interference results from a unique (i.e., *response*) *conflict* between two different possible responses.

In a firm contrast *early-single stage selection accounts* (Hock & Egeth, 1970; Luo, 1999; Scheibe et al., 1967; Seymour, 1974; 1977; Stirling, 1979; see also Klopfer, 1996), proposed that a single conflict responsible for the Stroop interference effect arises earlier – at the level of processing input. Although this conflict was conceptualized as perceptual (Hock & Egeth, 1970), the majority of these accounts consider this conflict to occur at a conceptual encoding level of color-incongruent words. For instance, Seymour (1977) argued that « (...) one is led towards the view that the effect arises at a level of conceptual representation and derives from a semantic conflict brought about by simultaneous analysis of both aspects of the colored word display. The principle involved here could be that delays of processing occur whenever distinct semantic codes are simultaneously activated, and that these delays become acute when the conflicting codes are values on a single dimension or a closely related dimension.” (Seymour, 1977, p. 263; see also Seymour, 1973). In this view, the Stroop interference effect results from a conflict between these two different semantic representations and is, therefore, often termed as *semantic conflict*.

In summary, both early and late selection accounts agree on the fact that the Stroop interference is a unitary phenomenon such that it results from a unique conflict. They both agree on the fact that the Stroop interference effect is a result of the automaticity of the word-reading. However, the early and late single-stage selection accounts clearly disagree in regards to the level of processing at which this unique conflict is located (at the level of stimulus vs. response) and, consequently, on the type of conflict (semantic vs. response conflict) that is actually responsible for the Stroop interference. In 1998, Logan and Zbrodoff

argued that « There is some support for each position, but the bulk of the evidence points to a locus in response selection (p. 979; see also MacLeod, 1991 for similar conclusion).

This latter conclusion does not take into account a *multi-stage selection model* of Zhang and Kornblum (1998) that was also proposed the same year. Its starting point is the aforementioned long-term opposition between the early and late (single-stage) accounts. According to Zhang and Kornblum « These two proposals (...) focus on one particular aspect of the Stroop task to the exclusion of the other. The early-selection account focuses on the similarity between the relevant stimulus and the irrelevant stimulus, whereas the late-selection account focuses on the similarity between the irrelevant stimulus and the response. Both similarity relationships are, of course, present in the Stroop task – in fact, they constitute a confounding that makes distinguishing empirically between the two accounts difficult. » (Zhang & Kornblum, 1998, p. 4).

This latter idea is rooted in the theory of *Dimensional Overlap* (DO, Kornblum, 1992; see also e.g., Kornblum et al., 1990; Kornblum & Lee, 1995) that is defined as the occurrence of perceptual, conceptual, or structural similarity between stimulus sets, stimulus and response sets, or both.

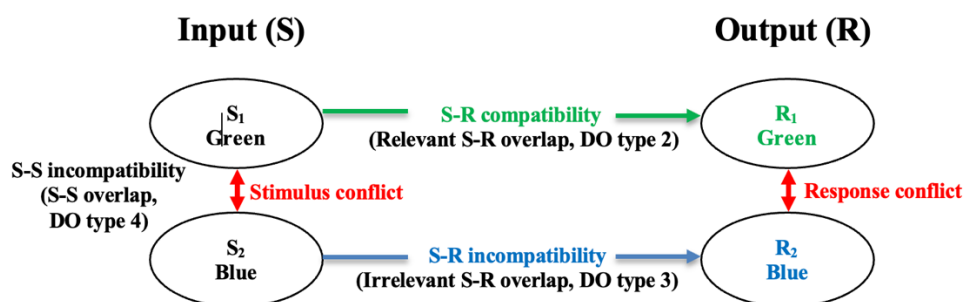


Figure 1. Representation of possible dimensional overlaps in incongruent Stroop stimuli

Dimensional overlap may occur between the relevant stimulus and response dimensions (called relevant SR overlap *or DO type 2, see Figure 1*), between the irrelevant stimulus and response dimensions (called irrelevant SR overlap *or DO type 3, see Figure 1*), or between the relevant and irrelevant stimulus dimensions (called SS overlap, *or DO type 4, see Figure 1*) » (1998, p. 4, text in italics added).

Therefore, in an incongruent Stroop stimulus (e.g., *BLUE*_{green}), a color stimulus (GREEN) overlaps with a written stimulus (BLUE) creating stimulus-stimulus overlap (S-S overlap or DO type 4, see Figure 1). As a result of this overlap, the meaning of the word conflicts with the meaning of the color that is to be named – creating the first source of interference (i.e., interference located at the input or stimulus (S) level, see Figure 1). Each of these two stimuli prompts a response. Specifically, because the response that is prompt by the relevant (color) stimulus (GREEN; DO type 2, see Figure 1) and that is prompt by the irrelevant (written) stimulus (BLUE; DO type 3, see Figure 1) are different, an additional (i.e., response) conflict occurs at the output or response (R) level (see Figure 1).

Despite the fact that this model is particularly convincing (as it nicely bridges two opposing perspectives), single-stage late selection accounts (e.g., Cohen et al., 1990; Glaser & Glaser, 1989; Logan & Zbrodoff, 1979; Melara & Algom, 2003; Phaf et al., 1990; Roelofs, 2003) still keep dominating the Stroop research at both theoretical (see e.g., Risko et al., 2006 for discussion) and applied level (e.g., Stroop Victoria test included in Strauss et al., 2006). Indeed, the Stroop task is now viewed as a “a prototypical inhibition task (...) in which one needs to inhibit or override the tendency to produce a more dominant or automatic response (i.e., name the color word)”, and the magnitude of Stroop interference is thought to reflect “one's ability to deliberately inhibit dominant, automatic, or prepotent responses when necessary” (Miyake et al., 2000). One of the main reasons for this dominance is linked to another ongoing debate of whether the contribution of semantic conflict to Stroop interference

can be actually be assessed (e.g., Augustinova & Ferrand, 2014; Augustinova et al., 2015; De Houwer, 2003; Parris et al., 2021).

How to assess the contribution of a semantic conflict to the Stroop interference?

To better implement Zhang and Kornblum's model (1998) empirically, De Houwer (2003) proposed the so-called *two-to-one* paradigm in which two response colors are mapped to one response button (e.g., blue and red are assigned to the 'f' key and green and yellow are assigned to the 'j' key, see Figure 2).

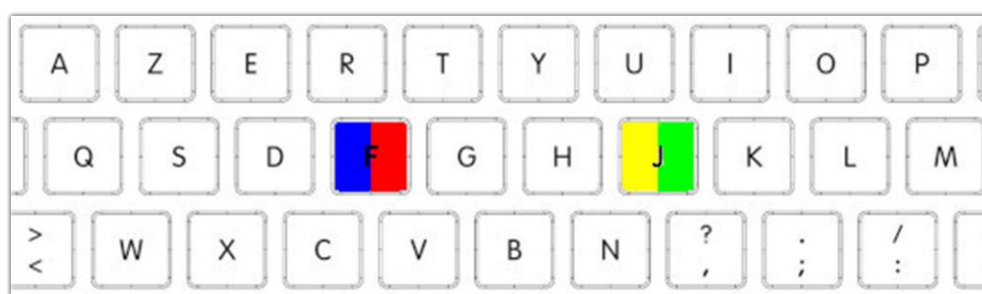


Figure 2. Response mapping in the two-to-one paradigm

As a result of this response mapping, two kinds of color-incongruent trials occur. Some trials (e.g., $BLUE_{\text{green}}$) generate an incorrect response activity toward the 'f' key (primed by the distractor " $BLUE$ ") that subsequently interferes with the production of the correct response "green" that is made by pressing the 'j' key (hence are termed *different-response trials*). Other color-incongruent trials (e.g., $BLUE_{\text{red}}$, termed *same-response trials*) do not generate this type of response interference, because the response activity is primed by both the distractor " $BLUE$ " and the target color "red" converge toward the same (here 'f') key. Consequently, according to unitary or single-stage response conflict models mentioned above, these latter trials should not produce any interference, as they do not involve any response conflict. However, in De Houwer (2003)'s study not only were both different- and same-response trials responded to slower than baseline identity or color-congruent trials

($BLUE_{blue}$), but, also, different-response trials were responded to slower than same-response trials (e.g., RTs to $BLUE_{green} > RTs$ to $BLUE_{red}$). Therefore, in line with the two-conflicts model of Zhang and colleagues (H. Zhang et al., 1999; H. Zhang & Kornblum, 1998a), these results allowed (De Houwer, 2003) to conclude that different-response trials generate both semantic and response conflicts, whereas interference produced by same-response trials results solely from semantic conflict.

These findings have largely been replicated since and the aforementioned so-called two-to-one Stroop paradigm has become, perhaps, the most popular way of distinguishing the contribution of semantic conflict from that of response conflict (e.g., A. Chen et al., 2011, 2013; Z. Chen et al., 2013; Hershman & Henik, 2019, 2020b; Jiang et al., 2015; Šaban & Schmidt, 2021; Schmidt et al., 2018; Schmidt & Cheesman, 2005a; Shichel & Tzelgov, 2018b; Van Veen & Carter, 2005). Despite this profusion of convergent empirical results, the contribution of semantic conflict to the overall color-word Stroop effect remains an open empirical issue (see Parris et al., 2021 for ample discussion).

Indeed, all studies employing this measure of semantic conflict – including De Houwer (2003) – have used color-congruent (or identity) trials as the baseline against which interference, induced by same-response trials, is measured. Therefore, unitary or single-stage response conflict models of the Stroop effect (Cohen et al., 1990a; Glaser & Glaser, 1989; Logan, 1980; Melara & Algom, 2003; Phaf et al., 1990; Roelofs, 2003a) can still easily account for this difference as resulting from facilitation on color-congruent trials (as opposed to interference on same-response trials). In line with this possible interpretation (see also Roelofs, 2010), Hasshim and Parris (2014, 2015) consistently reported significantly longer RTs for same-response trials than for color-congruent trials, but no difference between same-response trials and trials that are free of facilitation (i.e., color-neutral word trials, $DEAL_{green}$ see e.g., Brown, 2011; MacLeod, 1991a for discussion). The absence of semantic conflict was

supported further by Bayesian evidence for the null-hypothesis and by the unchanged magnitude of associated pre-response pupillometric measures of effort (i.e., a reliable and more sensitive measure of the potential differences between conditions; Hasshim & Parris, 2015; Parris et al., 2021). Therefore, at this point, the contribution of semantic conflict to the overall Stroop interference in the two-to-one Stroop paradigm is unclear.

Another popular way to separate the contribution of semantic conflict to Stroop interference is the so-called “sky-put” design (Neely & Kahan, 2001). This design also known as the Semantic Stroop paradigm (SSP) – first empirically employed by Manwell, Roberts and Besner (2004) – supplements standard color-incongruent stimuli (*BLUE*_{yellow}) and their color neutral counterparts (*PAGE*_{yellow}) with stimuli that are only semantically associated with a color (e.g., *SKY* that is associated with color blue). The addition of these stimuli – first introduced by Klein (1964) – makes it possible to compute two kinds of Stroop interference: *standard* and *semantic* Stroop interference.

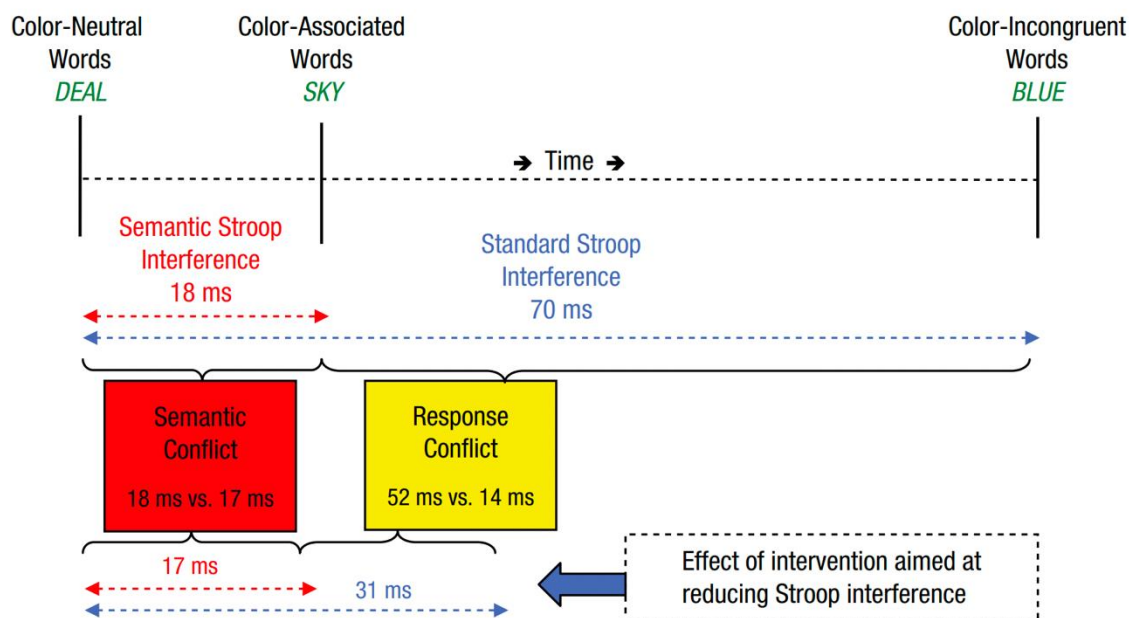


Figure 3. The semantic Stroop paradigm. Its “subtractive” logic (applied here to data from Augustinova & Ferrand, 2012b, Experiment 1, in which participants made vocal responses) makes it possible to differentiate between the respective contributions of semantic versus response conflict to the overall Stroop interference (taken from Augustinova & Ferrand, 2014b)

Standard Stroop interference corresponds to longer response latencies for standard color-incongruent as compared to color-neutral trials (e.g., $BLUE_{\text{green}} - DEAL_{\text{green}}$). As explained earlier, because standard incongruent trials are incompatible at both stimulus and response level, *standard* Stroop interference is expected to reflect the simultaneous contribution of semantic and response conflict (see Figure 3). According to Neely and Kahan (2001), the former conflict results from task-irrelevant processes involved in computing the lexical and semantic representations of the word (i.e., a written distractor to ignore), whereas the latter arises from task-relevant processes involved in the selection of a response (i.e., a color target to name).

The so-called *semantic* Stroop interference, which corresponds to longer response latencies for color-associated as compared to color-neutral trials (e.g., $SKY_{\text{green}} - DEAL_{\text{green}}$), is considered as free of this latter conflict (Augustinova & Ferrand, 2014b). Indeed, semantic associates do not generate a color response linked to the associated color (i.e., blue associated with the *SKY*) during the selection of a color to name (see e.g., Schmidt & Cheesman, 2005, for a straightforward demonstration; but see e.g., Risko et al., 2006; Šaban, 2021). However, there is still a semantic overlap between the target (i.e., green) and the distractor because blue is strongly associated to *SKY* and, therefore, activated upon presentation of this type of stimulus. Also, and importantly, thanks to the same “subtractive” logic as seen for two-to-one Stroop paradigm (see also Figure 3), this variant of the Stroop paradigm also makes it possible to evaluate the respective contributions of semantic and response conflict to overall Stroop interference.

As mentioned above, many studies subsequently confirmed the contribution of a semantic conflict (as assessed via semantic associated interference) to the overall Stroop interference (Augustinova, Clarys, et al., 2018; Augustinova & Ferrand, 2014b; Li & Bosman, 1996; Sharma & McKenna, 1998; White et al., 2016). However, single-stage response

competition models can still argue that semantic-associative interference ($SKY_{\text{yellow}} - DEAL_{\text{yellow}}$) measured in these studies results entirely from response conflict (e.g., Roelofs, 2003). According to this position, semantic associates elicit incorrect *response* activity (e.g., say “blue”/press blue for SKY_{yellow}) indirectly – through their association with the response-set colors (blue in this case) – which explains in turn the smaller magnitude of semantic-associative interference ($SKY_{\text{yellow}} - DEAL_{\text{yellow}}$) compared to its standard ($BLUE_{\text{yellow}} - DEAL_{\text{yellow}}$) counterpart. Under this account, none of these past studies satisfactorily demonstrated that the driving force behind the semantic-associative interference is indeed semantic and not (indirect) response conflict.

The very same criticism actually applied to studies that used *non-response* or *out-of-the set* trials (e.g., $PURPLE_{\text{green}}$, Klein, 1964; see e.g., Hasshim & Parris, 2018; Hasshim et al., 2019; Lamers et al., 2010; Sharma & McKenna, 1998). Indeed, as noted by Parris et al. (2021): “It has been argued that the interference measured by using non-response set trials, the *non-response set effect*, is an indirect measure of response conflict (Cohen et al., 1990; Roelofs, 2003) and is, thus, not a measure of semantic conflict.” (p. 9). Indeed, because there is a semantic link between the non-response set words and the response set colors, the indirect activation of the other response set colors leads to response competition with the target color.

Given this current state of art, Parris and colleagues (2021) have recently argued that “(...) at this point, it seems reasonable to conclude that published research conducted so far with additional color incongruent trial types (same-response, non-response, or semantic-associative trials) does not permit the unambiguous conclusion that the informational conflict generated by standard color-incongruent trials (...) can be decomposed into semantic and response conflicts. More than ever then, cumulative evidence from more time- and process-sensitive measures are required.” (Parris et al., 2021, p. 11).

Overview of the present work

In line with the latter idea, the first goal of the present dissertation was first to answer a simple question – is there any contribution of semantic conflict to the Stroop interference effect in the two-to-one Stroop paradigm as implemented by Hashim and Parris (2014, i.e., with an additional color-neutral baseline). Given that they concluded (see also Hashim & Parris, 2015) that semantic conflict in this task (i.e., significant positive difference between mean RTs to same-response and color-congruent (or identity trials) results from facilitation on the latter trials (as opposed to interference on same-response trials), this choice might come as a surprise. Recall that unitary or single-stage response conflict models of the Stroop effect (Cohen et al., 1990a; Glaser & Glaser, 1989; Logan, 1980; Melara & Algom, 2003; Phaf et al., 1990; Roelofs, 2003a) assume that same response trials do not produce any interference. Indeed, unlike semantic associates (*SKY*_{blue}) or out-of-the set trials (*PURPLE*_{blue}), they do not generate response conflict. Therefore, any interference generated by these trials can only be attributed to semantic conflict – as originally reasoned by De However (2003). Although Hashim and Parris (2014, 2015) failed to observe this interference relative to color-neutral baseline, unpublished findings generated at about the same time by Augustinova and Ferrand's lab provided there is at least some evidence in favor of this effect. Therefore, these finding motivated conceptual replications of Hashim and Parris (2014) that the present dissertation conveys.

Several studies of Augustinova and Ferrand mentioned above also indicated that semantic is spared by various experimental manipulations including hypnosis-like suggestion (Augustinova & Ferrand, 2012), priming of dyslexia (Augustinova & Ferrand, 2014a), variations in Response–Stimulus Intervals (Augustinova, Silvert et al., 2018), viewing position (Ferrand & Augustinova, 2014) and Single-letter Coloring and Cueing (Augustinova & Ferrand, 2007; Augustinova et al., 2010, 2015, Augustinova, Clarys et al., 2018). Indeed,

all of these studies (see e.g., Augustinova & Ferrand, 2014b for a review) suggested these interventions along with age-related differences (Augustinova, Clarys et al., 2018) only affect response conflict in the Semantic Stroop paradigm (see Figure 3 above). Also, and importantly, this dissociative pattern (i.e., significant semantic conflict while response conflict is reduced or even eliminated) is often viewed as indicating that the two conflicts are qualitatively distinct (Augustinova & Ferrand, 2014b; see also Parris et al., 2021 for discussion). Given this latter theoretical implication along with the criticisms of the Semantic Stroop paradigm, the second goal of this dissertation was to examine the extent to which semantic (unlike response) conflict is still preserved when isolated in the two-to-one Stroop paradigm.

To sum up, the present dissertation was aimed at contributing to the ongoing debate about the locus (vs. loci) of the Stroop interference effect and this via two main goals: to assess (1) if a semantic conflict does indeed contribute to the Stroop interference effect observed in the two-to-one Stroop paradigm (De Houwer, 2003); (2) whether it is still spared by various experimental interventions (Augustinova & Ferrand, 2014b) and age-related differences (Augustinova, Clarys et al., 2018; Li & Bosman, 1996). To this end, the two-to-one Stroop paradigm administered with an additional (i.e., color-neutral) baseline was combined with Single-letter coloring and cueing (SLCC) in Chapter 1 (Burca et al., 2021) and Chapter 4; with healthy aging in Chapter 2 (Burca et al., 2022); and with variations in Response-stimulus intervals in Chapters 3 and 4. Since each of these chapters were written as separate empirical papers, repetitions might occur throughout these empirical chapters.

References

- Augustinova, M., Clarys, D., Spatola, N., & Ferrand, L. (2018). Some further clarifications on age-related differences in Stroop interference. *Psychonomic Bulletin & Review*, 25(2), 767–774.
<https://doi.org/10.3758/s13423-017-1427-0>
- 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>
- Brown, T. L. (2011). The relationship between stroop interference and facilitation effects: Statistical artifacts, baselines, and a reassessment. *Journal of Experimental Psychology: Human Perception and Performance*, 37(1), 85–99. <https://doi.org/10.1037/a0019252>
- Burca M., Chausse P., Ferrand L., Parris B. A, Augustinova M. (2022). Some further clarifications on age-related differences in the Stroop task: new evidence from the two-to-one Stroop paradigm. *Psychonomic Bulletin and Review*, 29, 492–500. <https://doi.org/10.3758/s13423-021-02011-x>
- Burca M., Beauconsin V., Chausse P., Ferrand L., Parris B. A., Augustinova M. (2021). Is there semantic conflict in the Stroop task? Further evidence from the 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>
- Chen, A., Bailey, K., Tiernan, B. N., & West, R. (2011). Neural correlates of stimulus and response interference in a 2–1 mapping stroop task. *International Journal of Psychophysiology*, 80(2), 129–138.
<https://doi.org/10.1016/j.ijpsycho.2011.02.012>
- Chen, A., Tang, D., & Chen, X. (2013). Training Reveals the Sources of Stroop and Flanker Interference Effects. *PLoS ONE*, 8(10), e76580. <https://doi.org/10.1371/journal.pone.0076580>
- Chen, Z., Lei, X., Ding, C., Li, H., & Chen, A. (2013). The neural mechanisms of semantic and response conflicts: An fMRI study of practice-related effects in the Stroop task. *NeuroImage*, 66, 577–584.
<https://doi.org/10.1016/j.neuroimage.2012.10.028>
- Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990a). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, 97(3), 332–361.
<https://doi.org/10.1037/0033-295X.97.3.332>
- Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990b). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, 97(3), 332–361.

- De Houwer, J. (2003). On the role of stimulus-response and stimulus-stimulus compatibility in the Stroop effect. *Memory & Cognition*, 31(3), 353–359. <https://doi.org/10.3758/BF03194393>
- Duncan-Johnson, C. C., & Kopell, B. S. (1981). The Stroop effect: Brain potentials localize the source of interference. *Science (New York, N.Y.)*, 214(4523), 938–940.
- Dyer, F. N. (1973). The Stroop phenomenon and its use in the study of perceptual, cognitive, and response processes. *Memory & Cognition*, 1(2), 106–120. <https://doi.org/10.3758/BF03198078>
- Fraisse, P. (1969). Why is naming longer than reading? *Acta Psychologica*, 30, 96–103. [https://doi.org/10.1016/0001-6918\(69\)90043-2](https://doi.org/10.1016/0001-6918(69)90043-2)
- Maquestiaux, F. (2017). *Psychologie de l'attention*. De Boeck Supérieur: Bruxelles.
- Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General*, 118(1), 13–42.
- Hasshim, N., & Parris, B. A. (2014). Two-to-one color-response mapping and the presence of semantic conflict in the Stroop task. *Frontiers in Psychology*, 5, 1157. <https://doi.org/10.3389/fpsyg.2014.01157>
- Hasshim, N., & Parris, B. A. (2015). Assessing stimulus–stimulus (semantic) conflict in the Stroop task using saccadic two-to-one color response mapping and prereponse pupillary measures. *Attention, Perception, & Psychophysics*, 77(8), 2601–2610. <https://doi.org/10.3758/s13414-015-0971-9>
- Hasshim, N., & Parris, B. A. (2018). Trial type mixing substantially reduces the response set effect in the Stroop task. *Acta Psychologica*, 189, 43–53. <https://doi.org/10.1016/j.actpsy.2017.03.002>
- Hasshim, N., Downes, M., Bate, S., & Parris, B. A. (2019). Response time distribution analysis of semantic and response interference in a manual response Stroop task. *Experimental Psychology*, 66(3), 231–238. <https://doi.org/10.1027/1618-3169/a000445>
- Hershman, R., & Henik, A. (2019). Dissociation between reaction time and pupil dilation in the Stroop task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(10), 1899–1909. <https://doi.org/10.1037/xlm0000690>
- Hershman, R., & Henik, A. (2020). Pupillometric contributions to deciphering Stroop conflicts. *Memory & Cognition*, 48(2), 325–333. <https://doi.org/10.3758/s13421-019-00971-z>
- Hommel, B. (1997). Interactions between stimulus-stimulus congruence and stimulus-response compatibility. *Psychological Research*, 59(4), 248–260. <https://doi.org/10.1007/BF00439302>
- Hunt, E., & Lansman, M. (1986). Unified model of attention and problem solving. *Psychological Review*, 93(4), 446–461. <https://doi.org/10.1037/0033-295X.93.4.446>

- Jiang, J., Zhang, Q., & Van Gaal, S. (2015). EEG neural oscillatory dynamics reveal semantic and response conflict at difference levels of conflict awareness. *Scientific Reports*, 5(1), 12008.
<https://doi.org/10.1038/srep12008>
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6(2), 293–323. [https://doi.org/10.1016/0010-0285\(74\)90015-2](https://doi.org/10.1016/0010-0285(74)90015-2)
- Lamers, M. J. M., Roelofs, A., & Rabeling-Keus, I. M. (2010). Selective attention and response set in the Stroop task. *Memory & Cognition*, 38(7), 893–904. <https://doi.org/10.3758/MC.38.7.893>
- Li, K. Z. H., & Bosman, E. A. (1996). Age differences in stroop-like interference as a function of semantic relatedness. *Aging, Neuropsychology, and Cognition*, 3(4), 272–284.
<https://doi.org/10.1080/13825589608256630>
- Logan, G. D. (1978). Attention in character-classification tasks: Evidence for the automaticity of component stages. *Journal of Experimental Psychology: General*, 107(1), 32–63. <https://doi.org/10.1037/0096-3445.107.1.32>
- Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology*, 12(4), 523–553. [https://doi.org/10.1016/0010-0285\(80\)90019-5](https://doi.org/10.1016/0010-0285(80)90019-5)
- 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. (1992). The Stroop task: The “gold standard” of attentional measures. *Journal of Experimental Psychology: General*, 121(1), 12–14. <https://doi.org/10.1037/0096-3445.121.1.12>
- Maquestiaux, F. (2017). Psychologie de l'attention. De Boeck Supérieur.
- Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review*, 110(3), 422–471. <https://doi.org/10.1037/0033-295X.110.3.422>
- Morton, J., & Chambers, S. M. (1973). Selective Attention to Words and Colours. *Quarterly Journal of Experimental Psychology*, 25(3), 387–397. <https://doi.org/10.1080/14640747308400360>
- Parris, B. A., Hasshim, N., Wadsley, M., Augustinova, M., & Ferrand, L. (2021). The loci of Stroop effects: A critical review of methods and evidence for levels of processing contributing to color-word Stroop effects and the implications for the loci of attentional selection. *Psychological Research*, 0123456789.
<https://doi.org/10.1007/s00426-021-01554-x>

- Phaf, R. H., Van der Heijden, A. H. C., & Hudson, P. T. W. (1990). SLAM: A connectionist model for attention in visual selection tasks. *Cognitive Psychology*, 22(3), 273–341. [https://doi.org/10.1016/0010-0285\(90\)90006-P](https://doi.org/10.1016/0010-0285(90)90006-P)
- Posner, M. I., & Snyder, C. R. R. (n.d.). Attention and cognitive control. In R. L. Solso (Ed.) *Information processing and cognition: The Loyola symposium* (pp. 55-85). Hillsdale, NJ: Erlbaum.
- Risko, E. F., Schmidt, J. R., & Besner, D. (2006). Filling a gap in the semantic gradient: Color associates and response set effects in the Stroop task. *Psychonomic Bulletin & Review*, 13(2), 310–315. <https://doi.org/10.3758/BF03193849>
- Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review*, 110(1), 88–125. <https://doi.org/10.1037/0033-295X.110.1.88>
- Šaban, I., & Schmidt, J. R. (2021). Stimulus and response conflict from a second language: Stroop interference in weakly-bilingual and recently-trained languages. *Acta Psychologica*, 218, 103360. <https://doi.org/10.1016/j.actpsy.2021.103360>
- Schmidt, J. R., & Cheesman, J. (2005). Dissociating stimulus-stimulus and response-response effects in the stroop task. *Canadian Journal of Experimental Psychology*, 59(2), 132–138. <https://doi.org/10.1037/H0087468>
- Schmidt, J. R., Hartsuiker, R. J., & De Houwer, J. (2018). Interference in Dutch–French Bilinguals. *Experimental Psychology*, 65(1), 13–22. <https://doi.org/10.1027/1618-3169/a000384>
- Sharma, D., & McKenna, F. P. (1998). Differential components of the manual and vocal Stroop tasks. *Memory & Cognition*, 26(5), 1033–1040. <https://doi.org/10.3758/BF03201181>
- Shichel, I., & Tzelgov, J. (2018). Modulation of conflicts in the Stroop effect. *Acta Psychologica*, 189, 93–102. <https://doi.org/10.1016/j.actpsy.2017.10.007>
- 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>
- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). A Compendium of Neuropsychological Tests: Administration, Norms and Commentary. New York: Oxford University Press.
- 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>

- Van Veen, V., & Carter, C. S. (2005). Separating semantic conflict and response conflict in the Stroop task: A functional MRI study. *NeuroImage*, 27(3), 497–504.
<https://doi.org/10.1016/J.NEUROIMAGE.2005.04.042>
- White, D., Risko, E. F., & Besner, D. (2016). The semantic Stroop effect: An ex-Gaussian analysis. *Psychonomic Bulletin & Review*, 23(5), 1576–1581. <https://doi.org/10.3758/s13423-016-1014-9>
- Zhang, H., & Kornblum, S. (1998). The Effects of Stimulus-Response Mapping and Irrelevant Stimulus-Response and Stimulus-Stimulus Overlap in Four-Choice Stroop Tasks with Single-Carrier Stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 24(1), 3–19.
<https://doi.org/10.1037/0096-1523.24.1.3>
- Zhang, H., Zhang, J., & Kornblum, S. (1999). A Parallel Distributed Processing Model of Stimulus–Stimulus and Stimulus–Response Compatibility. *Cognitive Psychology*, 38(3), 386–432.
<https://doi.org/10.1006/COGP.1998.0703>

CHAPTER ONE: IS THERE SEMANTIC CONFLICT IN THE STROOP TASK? FURTHER EVIDENCE FROM A MODIFIED TWO-TO-ONE STROOP PARADIGM COMBINED WITH SINGLE-LETTER COLORING AND CUEING¹

1. Introduction

The typical result in the well-known Stroop task (Stroop, 1935) is that individuals' color-identification times are longer for *color-incongruent* (e.g., “BLUE” displayed in yellow; hereafter *BLUE*_{yellow}), than for *color-neutral* items (e.g., the word “DEAL” displayed in yellow, hereafter *DEAL*_{yellow}). This difference – called *the Stroop interference effect* – reveals that, despite being explicitly instructed not to, individuals still attend to the irrelevant word-dimension of compound Stroop items (i.e., process its meaning), which in turn slows their performance. Therefore, numerous studies have tried to identify experimental manipulations that improve attentional selectivity in the Stroop task.

Among such manipulations, the effects of *single-letter coloring and cueing* (SLCC; Besner et al., 1997) have perhaps received the most attention (e.g., Küper & Heil, 2012 for a review). Despite differences in the methodologies employed², all of these studies have reported a substantial reduction or even elimination of Stroop interference when only a single letter (as opposed to all letters) of a target display was colored in an incongruent color from the response set and spatially pre-cued (hereafter SLCC vs. ALCC [all-letters colored and

¹ Corresponds to 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 the 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>

² In some studies, the coloring manipulation was coupled with variations in spatial pre-cueing such that, for instance, (a) small arrow(s) appeared on the screen to indicate the position(s) subsequently occupied by the color-carrier(s) (i.e., target letter(s)) whose color was to be named. Some studies (e.g., Manwell et al., 2004) also added empty spaces between letters that were filled – in other studies (e.g., Labuschagne & Besner, 2015) – with non-letter characters.

cued]). Because SLCC (unlike ALCC) is thought to keep two informational sources (i.e., color and word) separate (i.e., *odd-one-out effect* of SLCC, Besner et al., 1997), a response induced by the irrelevant word-dimension of compound Stroop items (e.g., say “blue”/press blue for $BLUE_{yellow}$) interferes less with that induced by the relevant color-dimension (e.g., say “yellow”/press yellow for $BLUE_{yellow}$) in SLCC than in ALCC. Indeed, according to dominant *single-stage response competition* models (e.g., Roelofs, 2003), a single conflict between responses (hence *response conflict*) is the driving force behind the Stroop interference effect.

In contrast to this unitary view of Stroop interference, other (i.e., *multi-stage*) models anticipate that incidental processing of the irrelevant word-dimension of Stroop items actually generates an additional (i.e., *semantic or stimulus*) conflict (e.g., Zhang et al., 1999; Zhang & Kornblum, 1998). On the basis of this idea, a handful of studies set out to investigate whether the SLCC-manipulation affects semantic and/or response conflict.

Manwell and colleagues (2004) consequently reported that SLCC reduced Stroop interference depicted above (e.g., $BLUE_{yellow} - DEAL_{yellow}$) but eliminated *semantic-associative* Stroop interference (e.g., $SKY_{yellow} - DEAL_{yellow}$, see also Labuschagne & Besner, 2015). Because these studies viewed the former (hereafter *standard* Stroop interference) as a product of both semantic and response conflicts, and viewed semantic-associative Stroop interference as being free of response conflict (e.g., Neely & Kahan, 2001³; but see e.g., Klein, 1964), they consequently concluded that SLCC affects the Stroop interference effect early (i.e., it curtails semantic conflict altogether, Manwell et al., 2004, Account 2; see also Besner et al., 2016 for discussion). But this conclusion contrasts with other studies reporting that SLCC reduces standard but leaves semantic-associative Stroop interference (Augustinova

³ Since SKY is not part of the response set, it does not activate (pre-)motor responses linked to the associated color (e.g., press a blue button on seeing SKY; see Schmidt & Cheesman, 2005 for a direct demonstration).

& Ferrand, 2007; Augustinova et al., 2010, 2018), and the associated N400-like ERP-activity⁴ (Augustinova et al., 2015) unaffected. These studies subsequently claimed that SLCC affects the Stroop interference effect late (i.e., it reduces or even eliminates response conflict, see Augustinova & Ferrand, 2014 for discussions) but leaves semantic conflict unaffected. In sum, the processing stage at which SLCC-manipulation actually takes its effect (semantic vs. response) remains unclear.

However, this lack of clarity is not entirely due to the contrasting findings across the aforementioned studies. Indeed, as noted above, all of these studies have employed the semantic-associative Stroop trials (e.g., *SKY_{yellow}*) to investigate the extent to which SLCC affects semantic conflict. Because the irrelevant word-dimension “sky” is strongly associated with “blue”, attending to its meaning is thought to interfere with the meaning of the relevant color-dimension (yellow for *SKY_{yellow}*). Indeed, “(...) delays of processing occur whenever distinct semantic codes are simultaneously activated, and that these delays become acute when the conflicting codes are values on a single dimension or closely related dimensions” (Seymour, 1977, p. 263). Yet, this assumption is at odds with the aforementioned unitary models of Stroop interference such as that proposed by Roelofs (2003). This model predicts that both standard and semantic-associative Stroop interference are not qualitatively distinct but both result from different quantities of a single (i.e., response) conflict taking place in the language production unit. Specifically, for semantic-associative items, the incorrect response (e.g., say “blue”/press blue for *SKY_{yellow}*) is activated through the association of the distractor *SKY* with the response set colors (i.e., blue is a part of the response set) in such a way that it subsequently interferes with the production of the correct response (e.g., “yellow”/press yellow button for *SKY_{yellow}*). As a result, the influence of SLCC on semantic conflict remains an open issue because the extent to which semantic conflict is the actual driving force behind

⁴ N400-like corresponded to fronto-central negativity occurring from 380 to 480msec after the presentation of a Stroop item taken as evidence of semantic processing in the Stroop task.

semantic-associative interference (e.g., $SKY_{\text{yellow}} - DEAL_{\text{yellow}}$) is itself unclear (see Parris et al., 2021 for discussion). To address this more fundamental issue, the present study substituted semantic-associative items with items producing semantic interference in such a way that cannot be accounted for by *single-stage response competition models* (i.e., *cannot be the result of response conflict*).

The present study

To this end, the study only used color-words as distractors (i.e., BLUE, RED, GREEN, YELLOW). Importantly, following De Houwer (2003), the same response-key (e.g., actioned with left hand) was used for making responses to blue and red (i.e., target) items, and the other response-key (e.g., actioned with right hand) was used for making responses to green and yellow items. This so-called *two-to-one* response-mapping therefore generated two kinds of color-incongruent trials. For *different-response* trials like $BLUE_{\text{yellow}}$ the correct response (“BLUE”) is indicated using the pointing finger on the right hand, and the incorrect response is indicated by using the pointing finger of the left hand. There is no such (response) interference on *same-response* trials like $BLUE_{\text{red}}$, because the responses primed by both the distractor “BLUE” and the target “RED” are indicated using the pointing finger on the left hand⁵. Consequently, any significant interference caused by same-response trials cannot be attributed to response conflict. Indeed, in line with *multi-stage models* of Stroop interference (e.g., Zhang & Kornblum, 1998), it is commonly attributed to semantic conflict (De Houwer, 2003, see e.g., Hershman & Henik, 2020 for the most recent example).

However, despite this consensus, the extent to which same-response trials actually induce semantic conflict is still an open empirical issue. Indeed, with only a few exceptions

⁵ Note that these items are likely to involve at least some response facilitation (Hasshim & Parris, 2014; 2015 for a discussion).

(see below), all studies employing same-response trials to induce semantic conflict – including De Houwer (2003) – have used color-congruent trials as the baseline. This means that in line with what unitary models of Stroop interference would predict (Roelofs, 2003), the difference between same-response and color-congruent trials might not involve any semantic conflict but be entirely driven by facilitation on color-congruent trials. Consistent with this idea, Hasshim and Parris (2014, 2015) reported significantly longer RTs for same-response trials than for color-congruent trials, but no difference between same-response and color-neutral word trials that are free from facilitation (Brown, 2011; MacLeod, 1991 for discussion)⁶. Somewhat in contrast, Hershman and Henik (2020) reported evidence for a difference between same-response and *non-word* color-neutral trials but in pupillometric measures only⁷, whereas Burca found preliminary – as yet unpublished – evidence suggesting the presence of semantic conflict in RTs even when color-neutral *words* are used as baseline. In sum, because of the theoretical implications for composite (e.g., Zhang & Kornblum, 1998) as opposed to unitary (e.g., Roelofs, 2003) Stroop interference, the unambiguous demonstration of same-response trials inducing (or not) *semantic conflict* still remains an empirical void to fill.

Therefore, the present study used color-neutral words to measure the magnitude of interference induced respectively by *same-* (e.g., *BLUE*_{red}) and *different-response* (*BLUE*_{yellow}) trials. Specifically, color-neutral trials were matched with color-incongruent ones in such a way that they appeared in the same colors an equal number of times (e.g., *BLUE* only appeared in red, green, yellow, and so did its counterpart *DEAL*). This allowed for

⁶ The absence of semantic conflict was supported further by Bayesian evidence for the null-hypothesis and by the unchanged magnitude of associated pre-response pupillometric measures of effort (i.e., a reliable measure of the potential differences between conditions, Hasshim & Parris, 2015).

⁷ In the case of RTs, this study perfectly replicated Hasshim and Parris (2014, 2015) including conclusions based on Bayes factors. Based on this dissociation, these authors concluded that RTs are not the best-suited measures to address the more fine-grained components of the Stroop interference effect (see also Hershman & Henik, 2019).

comparisons of *same-* (e.g., $BLUE_{red}$) and *different-response* ($BLUE_{yellow}$) trials with their corresponding (in terms of response contingency) color-neutral baselines (e.g., $DEAL_{red}$, and $DEAL_{yellow}$ respectively). In sum, as in past studies of SLCC-effect outlined above, the magnitudes of both *semantic* (e.g., $BLUE_{red} - DEAL_{red}$) and *standard Stroop interference* ($BLUE_{yellow} - DEAL_{yellow}$) were derived. However, this was done without using the same color-neutral baseline twice (see Shichel & Tzelgov, 2018 for discussion). If, in agreement with single-stage response competition models (e.g., Roelofs, 2003), there is no contribution of semantic conflict to the overall Stroop interference effect, same-response trials will not induce any Stroop interference (i.e., their RTs will be comparable to those observed for color-neutral trials – as previously reported by Hasshim and Parris (2014, 2015)). Different-response trials will be responded to slower than same-response trials (De Houwer, 2003), and this in the absence of a significant difference between the same-response and their neutral counterparts. In this case – confirming that the response conflict is the unique driving force of the Stroop interference effect –, only the magnitude of standard Stroop interference (the significant difference between different-response and color-neutral trials) will be reduced in SLCC compared to ALCC.

In contrast, if the multi-stage models of Stroop interference (e.g., Zhang & Kornblum, 1998) are correct and semantic conflict genuinely contributes to overall Stroop interference, then same-response trials will produce significant semantic Stroop interference such that the difference between same-response and color-neutral trials will be significant (Burca, unpublished; see also Hershman and Henik's pupillometry data). The semantic conflict could then be affected by SLCC in several ways. In line with past studies reviewed above, SLCC might *eliminate* semantic interference altogether (Manwell et al.'s Account 2; Labuschagne & Besner, 2015), or leave its magnitude unaffected (e.g., Augustinova et al., 2010, 2015, 2018). Finally, SLCC might simultaneously *reduce* both semantic and standard interference

(Manwell et al.'s Account 1)⁸ – a finding that would suggest that SLCC affects both semantic and response conflicts. The aim of the experiment reported below was to directly test these different hypotheses regarding the contribution of semantic conflict to Stroop interference and regarding its reduction by SLCC.

2. Method

Participants

Ninety-two native French-speakers (64 females and 28 males; $M_{\text{age}}=20.95$; $SD=1.612$) with normal or corrected-to-normal color-vision volunteered to take part in this experiment, which was approved by the local ethics committee (2018-02-A).

Design

The study used a 4 (Stimulus-Type: different response vs. same response vs. neutral different response vs. neutral same response) \times 2 (Coloring: ALCC vs. SLCC) within-participant design for data collection.

Stimuli

The (French) stimuli were presented in lowercase Courier font, size 18, on a black background and subtended an average visual angle of 0.9° high \times 3.0° wide. They consisted of four color-words: *rouge* [red], *jaune* [yellow], *bleu* [blue], and *vert* [green]; and four non-color counterparts: *plomb* [lead], *liste* [list], *page* [page], *cave* [basement], that were paired in length and frequency via *Lexique 3.38* (New et al., 2004).

Apparatus and Procedure

The participants were seated 70cm in front of a 17" computer screen. Eprime 2.0 software was used for data presentation and recording. Because the experimental trials only

⁸ This causes the elimination of semantic-associative Stroop interference ($SKY_{\text{yellow}} - DEAL_{\text{yellow}}$) as it is usually proportionally smaller than its standard counterpart ($BLUE_{\text{yellow}} - DEAL_{\text{yellow}}$).

used words as distractors, the participants were instructed to identify the color of the letter indicated by the arrow as quickly and accurately as possible by pushing the appropriate color-button and to ignore everything else in the display. To this end, and similarly to Augustinova et al. (2010, Experiment 1), the participants were instructed to concentrate on the fixation cross (“+”) that appeared for 2000msec in the center of the screen at the beginning of each trial. A small white arrow (height of 1.2° of visual angle, displayed 0.6° below) then appeared for 150msec. As it served as a spatial pre-cue, it was located at the position that was subsequently occupied by the target letter. This position varied randomly from trial-to-trial, being located at the initial, optimal viewing position (OVP), middle, or final letter of the distractor-word (Parris et al., 2007). In order to avoid an additional color-color interference in the SLCC-condition⁹, the spatially pre-cued letter was the only one that appeared in an incongruent color from the response set and the rest of the letters appeared in white (i.e., a color that was not part of the response set; see e.g., Augustinova & Ferrand, 2007; Besner et al., 1997; Brown et al., 2002 for the use of gray or black). In the ALCC-condition, all letters (including the one that was spatially pre-cued) appeared in incongruent colors from the response set. The entire display remained on screen until the participant responded or until 3500msec had elapsed.

The participants responded manually using a modified SRBox® consisting of two handles, each of which had a single response button at the top (placement of the handles in the right or left hand, respectively, was counterbalanced across participants). The participants pushed these response buttons with their thumbs. This allowed them to hold each handle comfortably in their palm with the remaining four fingers (see Supplementary Materials, Figure S1, left panel) while resting their arms on an armrest.

⁹ An additional color-color interference occurs when remaining letters are colored in other (incongruent) colors from the response set. Since this is known to increase RTs on color-neutral trials in the SLCC-condition (e.g. Küper & Heil, 2012; Monahan, 2001) it might inflate the SLCC-induced reduction of both standard and semantic Stroop interference (e.g. Manwell et al., 2004).

One response button was flanked by blue and red color-stickers and the other by yellow and green color-stickers (see Supplementary Materials, Figure S1, middle and right panels). Because of this color assignment, together with the fact that color-stimuli only appeared in incongruent colors, a single presentation of the full set of color-stimuli in all possible colors resulted in eight same-response items (i.e., 4 in ALCC and 4 in SLCC; e.g., “blue” presented entirely in red, “red” presented entirely in blue, “green” presented entirely in yellow and “yellow” presented entirely in green in ALCC) and sixteen different-response items. To control for contingency, the same presentation was used for the items' color-neutral counterparts (e.g., *PAGE* only appeared in red, yellow and green, exactly like its color-incongruent counterpart *BLUE*). All stimulus types were therefore seen an equal number of times, meaning that 5 repetitions (making it possible to fully control for the letter-position effects, see above) of the full set of 48 different stimuli resulted in a total of 240 experimental stimuli. In each coloring-condition, we therefore collected RTs on 20 same-response (SR) incongruent trials, 20 color-neutral trials paired with SR-incongruent trials, 40 different-response (DR) incongruent trials, and 40 color-neutral trials paired with DR-incongruent trials all included in a single block. Before completing this experimental block, the participants first completed 128 practice trials (MacLeod, 2005) consisting of asterisks in order to learn the color-button correspondence. They proceeded to the experimental block once their accuracy rate was above 95% (2 participants had to redo the training, one of them was later excluded from further analyses).

3. Results and Discussion

Four participants were excluded from the analysis due to faulty data recordings. The data of the remaining eighty-eight participants were first analyzed in an omnibus 4 (Stimulus-Type: DR-incongruent trials vs. SR-incongruent trials vs. DR-neutral trials vs. SR-neutral

trials) \times 2 (Coloring: ALCC vs. SLCC) ANOVA. This analysis on RTs was conducted on mean correct latencies using both 2.5 (e.g., Labuschagne & Besner, 2015) and 3SD cut-offs (Augustinova and colleagues' studies). Given their convergence and to permit comparisons with our past studies, only analyses using the 3SD cut-off (leading to the exclusion of 1.12% of the total data) are reported below. In omnibus ANOVA, this latter analysis revealed main effects of Stimulus-Type [$F(3,261) = 38.09$; $p < .001$, $\eta_p^2 = 0.304$; $BF_{10} = 4.160e+18^{10}$, indicating extreme evidence of an effect of Stimulus-Type] and of Coloring [$F(1,87) = 11.57$; $p = .001$, $\eta_p^2 = 0.117$; $BF_{10} = 29.223$, indicating strong evidence of an effect of Coloring]. It also revealed a significant Stimulus-Type \times Coloring interaction, [$F(3,261) = 3.236$; $p = .023$, $\eta_p^2 = 0.036$; $BF_{10} = 2.229e+20$]. This extreme Bayesian evidence, is in favor the model where the both main effects and their interaction are significant. Nevertheless, an effect of the interaction alone was not supported by Bayesian evidence ($BF_{10} = 0.606/BF_{01} = 1.648^{11}$, see Table S1 in Supplementary Materials for a comparison of models in a Bayesian repeated-measures ANOVA].

These same analyses of error percentages (see Table 1) revealed only significant main effects of Stimulus-Type [$F(3, 261) = 6.91$; $p < .001$, $\eta_p^2 = 0.74$; $BF_{10} = 182.463$, indicating extreme evidence of an effect of Stimulus-Type]. Both the main effect of Coloring [$F(1,87) = 2.83$; $p = .096$, $\eta_p^2 = 0.32$; $BF_{10} = 0.210/BF_{01} = 4.810$], and Stimulus-Type \times Coloring interaction [$F(3, 261) = .12$; $p = .950$, $\eta_p^2 = 0.01$] remained non-significant. Indeed, BFs provided very strong evidence against the effect of interaction alone ($BF_{10} = 0.013/BF_{01} = 68.52$, see Table S2 in the Supplementary Materials for a comparison of the models).

¹⁰ BF_{10} corresponds to the Bayesian probability of the occurrence of a hypothesis (H1) and the likelihood of another null hypothesis (H0). It was calculated with JASP (JASP Team, 2017) and interpreted according to Lee and Wagenmakers (2013 adjusted from Jeffreys, 1961). All priors were equal.

¹¹ BF value reported for interaction only is the one associated with the model including the interaction and the two main effects divided by BF value observed for the model including the two main effects only.

Is there semantic conflict in the Stroop task?

To answer this key question, we first decomposed the aforementioned main effect of the Stimulus-Type on mean correct latencies. It revealed the presence of both semantic and standard Stroop interference. The SR-incongruent trials were indeed responded to slower than their SR-neutral counterparts ($p < .001$; $BF_{10} = 44197.61$) and so were the DR-incongruent trials compared to their DR-neutral counterparts ($p < .001$; $BF_{10} = 1.342e+12$)¹². To explore this issue of semantic conflict further, we then decomposed the aforementioned Stimulus-Type \times Coloring interaction (i.e., interaction supported at least by frequentist statistics) by testing the simple main-effect of Stimulus-Type at each level of Coloring. This simple main-effect was significant in both ALCC [$F(3,85) = 22.762$; $p < .001$, $\eta_p^2 = 0.445$; $BF_{10} = 6.986e+11$] and SLCC [$F(3,85) = 12.898$; $p < .001$, $\eta_p^2 = 0.313$; $BF_{10} = 368930.025$]¹³. In addition to significant standard Stroop interference (DR-incongruent – DR-neutral trials) in both ALCC- and SLCC-condition (all $p_s < .01$, see Table 1 for descriptive statistics), further pairwise comparisons, again unequivocally revealed the presence of significant semantic Stroop interference (SR-incongruent – SR-neutral trials) in both ALCC- and SLCC-conditions (see Table 1 for magnitudes).

¹² The difference between the RT's of DR and SR incongruent trials (10msec), which is representative of the response conflict, did not reach significance ($p = .135$). However, $BF_{10} = 2.962/BF_{01} = 0.338$ values of paired samples T-Tests provided at least anecdotal evidence in favor of such an effect.

¹³ A Bayesian repeated measures ANOVA compares a series of different models against a null model. It therefore cannot be decomposed by testing simple main effects per se. Therefore, BFs provided test main effect of Stimulus-Type in ALCC and SLCC respectively – which explains why a Bayesian approach that is generally more conservative, actually yields larger effects.

Table 1. *Color-Identification Performance (Means, Standard Deviation, 95%-Confidence Intervals and Error Percentages) observed as a Function of Stimulus- or Interference- Type and Coloring*

	Coloring					
	All Letters Colored and Cued (ALCC)			Single Letter Colored and Cued (SLCC)		
	M (SD)	CI	%ER	M (SD)	CI	%ER
						SLCC effect (RT)
						SLCC effect (ER)
Color-Incongruent Different Response (DR)	720 (117)	[695, 744]	2.61	697 (95)	[677, 717]	2.81
						-23 ^{†/†}
						+ .2 ^{ns/ns*}
Color-Neutral Different Response (Neutral DR)	669 (86)	[651, 687]	1.68	670 (83)	[653, 688]	1.88
						+2 ^{ns/ns*}
						+ .2 ^{ns/ns*}
Color-Incongruent Same Response (SR)	708 (113)	[683, 732]	1.25	691 (99)	[669, 721]	1.53
						-17 ^{*ns}
						+ .3 ^{ns/ns*}
Color-Neutral Same Response (Neutral SR)	679 (157)	[656, 701]	2.10	671 (93)	[651, 690]	2.61
						-8 ^{ns/ns*}
						+ .5 ^{ns/ns*}
	M (SD)	CI		M (SD)	CI	
Standard Stroop Interference (DR – Neutral DR)	51 ^{†/††} (62)	[38, 64]		27 ^{†/††} (43)	[18, 36]	
Semantic Stroop Interference (SR – Neutral SR)	29 ^{†/††} (68)	[14, 43]		20 ^{**/**} (56)	[8, 31]	

Note. Inferential approach: ^{ns}non-significant; *significant at $p < .05$; **significant at $p < .01$; †significant at $p < .001$; Bayesian approach (presented after the slash): ^{ns}no evidence of an effect, with both BF_{10} and BF_{01} values between 0–1; ^{ns}anecdotal evidence in favor of H1, but moderate evidence in favor of H0, with BF_{10} value between 0–1, and BF_{01} value between 3–10; *moderate evidence in favor of H1, with BF_{10} value between 3–10; **strong evidence in favor of H1, with BF_{10} value between 10–30; †very strong evidence in favor of H1, with BF_{10} between 30–100; ††extreme evidence in favor of H1/of an effect, with $BF_{10} < 100$.

This latter conclusion is reinforced by $BF_{10} = 130.036$, providing extreme evidence of semantic interference in the ALCC-condition and $BF_{10} = 19.354$, providing strong evidence of semantic interference in the SLCC-condition¹⁴.

The presence of significant semantic interference overall, and in both coloring conditions, is thus consistent with the idea that a semantic conflict indeed genuinely contributes to the overall Stroop interference effect even when color-neutral *words* are used as baseline (see also Hershman & Henik, 2020' pupillometric measures, but see Hashim & Parris, 2014, 2015; Hershman & Henik, 2020' RTs).

How does SLCC affect semantic conflict?

The aforementioned significant semantic interference in SLCC-conditions runs counter to Manwell et al.'s Account 2 (see also Labuschagne & Besner, 2015) suggesting that SLCC *eliminates* semantic interference altogether. However, in line with Manwell and colleagues' Account 1 (and counter to past results observed by Augustinova, Ferrand and colleagues), SLCC (as opposed to ALCC) shortened RTs in both DR-incongruent trials [$F(1,87) = 13.396$; $p < .001$, $\eta_p^2 = 0.133$; $BF_{10} = 50.129$, which indicates very strong evidence of Coloring] and SR-incongruent trials [$F(1,87) = 5.912$; $p = .017$, $\eta_p^2 = 0.064$; $BF_{10} = 1.981/BF_{01} = 0.529$]. It should be noted however, that the latter Bayesian evidence for SLCC

¹⁴ In line with a specific contribution of response conflict to standard Stroop interference – latencies for DR-incongruent trials tended to be longer than those for SR-incongruent trials (see Table 1). But although the observed 12msec difference was indeed marginally significant ($p = .072$) in ALCC, $BF_{10} = .569/BF_{01} = 1.758$ tend to suggest otherwise. Despite this, response conflict in ALCC is still evident in percentages of errors [$t(87) = 3.80$, $p < .001$; $BF_{10} = 77.45$]. In SLCC, the 7msec difference between mean response latencies for DR-incongruent and SR-incongruent trials was not significant ($p = .173$, $BF_{10} = .291/BF_{01} = 3.436$), but response conflict was however present in percentages of errors [$t(87) = 2.79$, $p < .05$; $BF_{10} = 4.34$]. The additional sequential analyses of response conflict in ALCC vs. SLCC condition suggested that the probability of finding this type conflict if more observations were added, is unlikely (see Figure 2 in Supplementary Materials). Therefore, for each level of Stimulus and of Coloring, we binned slow vs. fast trials. The additional analyses (see Supplementary Materials) suggest that the lack of response conflict reported above is driven by the lack of response conflict in the fast trials (see Table 3S in Supplementary Materials). Again, the additional sequential analyses of response conflict in fast vs. slow trials suggested that the probability of finding this type conflict if more observations were added, is unlikely (see Figure 3S in Supplementary Materials).

shortening SR-incongruent trials was anecdotal. Finally, this simple main effect of Coloring on both DR-neutral [$F(1,87) = 0.117$; $p = .733$, $\eta_p^2 = 0.001$; $BF_{10} = 0.125/BF_{01} = 8.025$] and SR-neutral trials [$F(1,87) = 1.519$; $p = .221$, $\eta_p^2 = 0.017$; $BF_{10} = 0.244/BF_{01} = 4.091$] revealed that color-neutral trials remained unaffected by Coloring manipulation. To further examine the SLCC-effect in the two-to-one Stroop paradigm, we then conducted a 2 (Interference-Type: Standard vs. Semantic) \times 2 (Coloring: ALCC vs. SLCC) ANOVA to analyze the extent to which magnitudes of both standard and semantic Stroop interference were affected by SLCC (see Table 1). This analysis revealed main effects of Interference-Type [$F(1,87) = 7.285$; $p = .008$, $\eta_p^2 = 0.077$; $BF_{10} = 2.207$], and Coloring [$F(1,87) = 6.066$; $p = .016$, $\eta_p^2 = 0.065$; $BF_{10} = 5.243$] but only a non-significant Interference-Type \times Coloring interaction, [$F(1,87) = 1.925$; $p = .169$, $\eta_p^2 = 0.0322$]. Indeed, in line with the aforementioned Stimulus \times Coloring interaction, the specific BF value for the interaction (see Table S4 in Supplementary Materials for a comparison of the models) was only $BF_{10} = 0.401$ and again $BF_{01} = 2.491$ actually provided anecdotal evidence for a null effect of the Interference-Type \times Coloring interaction.

These results are therefore clearly inconsistent with Augustinova and colleagues' past results showing that SLCC leaves the magnitude of semantic Stroop interference unaffected (Augustinova et al., 2010, 2015, 2018). Indeed, the present results suggest that SLCC reduces both semantic and standard Stroop interference in tandem.

4. Discussion and Conclusions

With regards to the first aim of addressing controversies concerning the role of semantic conflict in contributing to Stroop interference, the results reported above provided extreme (ALCC) and strong (SLCC) Bayesian evidence for differences between same-response trials and color-neutral word baselines (e.g., *DEAL*_{red}). This means that the present study reports – for the first time – a genuine contribution of semantic conflict to overall

Stroop interference. Indeed, in all past Stroop studies this contribution was possibly confounded either with response conflict (e.g., when semantic-associative items [*SKY*_{blue}] were used to induce semantic conflict) or facilitation (when color-congruent items [*BLUE*_{blue}] were used as the baseline from which to derive the magnitude of semantic conflict generated by same-response items). Also, and importantly, magnitudes of semantic interference reported in the present paper are not inflated either by the use of color-neutral *non-word* baselines (see Brown, 2011 for discussion of this latter issue) or by the inclusion of color-congruent trials that are also known to amplify interference (e.g., Roelofs, 2014). This unambiguous contribution of a genuine semantic conflict to the (overall) Stroop interference effect – at least with manual responses that the two-to-one paradigm (De Houwer, 2003) necessarily involves – constitutes the most significant result reported in the present study.

Whilst it is unclear why it was observed in the present study and not in other studies comparing the interfering effect of same-response trials against a color-neutral word base-line (Hasshim & Parris, 2014, 2015), the unambiguous presence of a genuine semantic conflict has at least two immediate implications for several still unresolved issues in the Stroop literature. First, it provides clear evidence that so-called *informational conflict* (e.g., MacLeod & MacDonald, 2000) in the Stroop task includes semantic conflict. Indeed, because of the aforementioned confounds, the evidence in favor of this contribution from semantic conflict available was still inconclusive prior to the present paper (see Parris et al., 2021, for a thorough discussion of this issue). Second, the unambiguous presence of genuine semantic conflict clearly implies that selection occurs at the level of semantics – a finding that runs counter to the dominant single-stage response competition models of Stroop interference. Therefore, it is likely that these latter models need to be modified to make room for this type of conflict.

Despite the data being collected in a considerable sample (Brysbaert & Stevens, 2018) and analyzed with both classical (frequentist) and Bayesian inferential approaches, the results reported above are clearly less conclusive with regards to the second aim which was to address controversies concerning the reduction of semantic conflict by SLCC-manipulation. Indeed, the lack of Interference-Type \times Coloring interaction (along with a Stimulus-Type \times Coloring only being present in standard frequentist ANOVA) seems most consistent with Manwell et al.'s Account 1 suggesting that SLCC reduces both semantic and standard interference in tandem¹⁵. In line with Kinoshita et al. (2018; see also e.g., Besner et al., 2016), this means that SLCC reduces semantic conflict in the Stroop task. However, the results with regards to the SLCC-effect on SR-incongruent trials come with caveats. Indeed, the Bayesian analyses reported above provided anecdotal evidence for the simple main-effect of Coloring on these trials whereas this effect was strong for DR-incongruent trials. This points to the fact the while the interpretation in terms of reduced semantic conflict (i.e., the common denominator of both semantic and standard Stroop interference) is possible, it is likely to be incomplete. In line with past studies (e.g., Augustinova et al., 2010, 2015, 2018), it therefore remains possible that SLCC also reduces response conflict as the aforementioned strong simple main-effect of Coloring on DR-incongruent trials suggest. Although the response conflict was not significant overall¹⁶, this possibility cannot be ruled out entirely, as it was significant on slow trials (see Supplementary Materials). This reduction is again probably too

¹⁵ This conclusion is reinforced by the exploratory 2 (Interference-Types: Stroop interference vs. Semantic interference) \times 2 (Coloring: ALCC vs. SLCC) ANOVA conducted on slow trials only (i.e., trials revealing response conflict). Indeed, it yielded a marginally significant main effect of Interference-Type [$F(1,87)=3.50$; $p < .065$, $\eta_p^2 = 0.039$], a significant main effect of Coloring [$F(1,87) = 5.39$; $p = .023$, $\eta_p^2 = 0.058$], but a non-significant Interference-Type \times Coloring interaction [$F(1,87)=.294$; $p < .589$, $\eta_p^2 = 0.003$] suggesting again that Coloring manipulation affects both types of interference in tandem.

¹⁶ This unexpected absence of significant response conflict (that our group has actually replicated since) is at odds with past studies using the two-to-one Stroop paradigm – including De Houwer's initial study (2003). This could be explained by differences in methodology applied in the present study (e.g., absence of color-congruent items that are known to amplify the interfering effects of color-incongruent items (e.g., Roelofs, 2014), the equal percentage of color-incongruent vs color-neutral trials (MacLeod, 2005) with no contingency issues involved (Hasshim & Parris, 2014; Schmidt, 2013). More studies are needed to assess these different possibilities directly.

small to fully explain the present pattern of results. The remaining possibility is that SLCC actually reduces task conflict (i.e., a more general conflict that – for all readable Stroop items including color-neutral ones – derives from the simultaneous preparation of two task sets: word-reading vs. color-naming, e.g., Goldfarb & Henik, 2007; Hershman & Henik, 2019). The lack of SLCC-effects on color-neutral stimuli in the present study is seemingly at odds with this latter idea (but see e.g., Manwell et al., 2004). Indeed, any reduction of task conflict should in principle facilitate processing of all compound stimuli (i.e., including the color-neutral items) and even more so in the present study given that only one color from the response set was used to implement the SLCC-manipulation. It is, however, noteworthy that two colors are still present in our SLCC-condition (a color from the response set and white) – making it plausible that an additional color-color interference still occurs in SLCC as compared to ALCC. Its known lengthening effect on color-neutral items (e.g. Küper & Heil, 2012; Monahan, 2001) might be cancelled in the RTs by the concomitant shift in attentional focus toward the relevant color-dimension in SLCC, which in turn reduces the task conflict. In sum, future studies – which should include a direct measure of task conflict – need to address these possibilities directly, perhaps with a more fine-grained measures than RTs (e.g., Hershman & Henik, 2019). The decrease in the magnitude of overall Stroop interference that is specifically due to the reduction of task conflict implies – contrary to the reduction of semantic conflict – a simultaneous increase in Stroop facilitation (Parris, 2014), which was also not measured in the present study. It is of course possible that SLCC actually affects all the components of the overall Stroop interference (response, semantic and task conflict).

In conclusion, whilst the influence of SLCC on semantic conflict (but also on other types of conflicts) still remains an open issue, the presence of semantic conflict in the Stroop task – at least when administered in the form of the two-to-one Stroop paradigm –, is no longer one. Therefore, the present study provides impetus for future empirical work on SLCC

(see above). Furthermore, the present work also strongly encourages the development of new integrative models of the Stroop inference effect, as only one existing model effectively accounts for either semantic (Zhang & Kornblum, 1998; Zhang et al., 1999) or task conflict (Kalanthroff et al., 2018), respectively, but none currently accounts for the probable coexistence of task, semantic and response conflicts (see Parris et al., 2021 for further discussion).

5. References

- Augustinova, M., & Ferrand, L. (2007). Influence de la présentation bicolore des mots sur l'effet Stroop. *L'Année Psychologique*, 107(02), 163-179. <https://doi.org/10.4074/S0003503307002011>
- Augustinova, M., & Ferrand, L. (2014). Automaticity of Word Reading: Evidence from the Semantic Stroop Paradigm. *Current Directions in Psychological Science*, 23, 343–348.
- Augustinova, M., Flaudias, V., & Ferrand, L. (2010). Single-letter coloring and spatial cuing do not eliminate or reduce a semantic contribution to the Stroop effect. *Psychonomic Bulletin & Review*, 17(6), 827-833. <https://doi.org/10.3758/PBR.17.6.827>
- Augustinova, M., Clarys, D., Spatola, N., & Ferrand, L. (2018). Some further clarifications on age-related differences in Stroop interference. *Psychonomic Bulletin & Review*, 25(2), 767-774. <https://doi.org/10.3758/s13423-017-1427-0>
- Augustinova, M., Silvert, L., Ferrand, L., Llorca, P. M., & Flaudias, V. (2015). Behavioral and electrophysiological investigation of semantic and response conflict in the Stroop task. *Psychonomic Bulletin & Review*, 22(2), 543-549. <https://doi.org/10.3758/s13423-014-0697-z>
- 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., & Boutilier, C. (1997). The stroop effect and the myth of automaticity. *Psychonomic Bulletin & Review*, 4(2), 221-225. <https://doi.org/10.3758/BF03209396>
- Brown, T. L. (2011). The relationship between Stroop interference and facilitation effects: Statistical artifacts, baselines, and a reassessment. *Journal of Experimental Psychology: Human Perception and Performance*, 37(1), 85-99.

- Brown, T. L., Joneleit, K., Robinson, C. S., & Brown, C. R. (2002). Automaticity in Reading and the Stroop Task: Testing the Limits of Involuntary Word Processing. *The American Journal of Psychology*, 115(4), 515. <https://doi.org/10.2307/1423526>
- Brysbaert, M., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition*, 1(1). <https://doi.org/10.5334/joc.10>
- Burca, M. (unpublished). *Is there any Semantic Conflict in the two-to-one Stroop paradigm? (Working title)*. Doctoral dissertation in progress. University of Rouen.
- De Houwer, J. (2003). On the role of stimulus-response and stimulus-stimulus compatibility in the Stroop effect. *Memory & Cognition*, 31(3), 353-359. <https://doi.org/10.3758/BF03194393>
- Goldfarb, L., & Henik, A. (2007). Evidence for task conflict in the Stroop effect. *Journal of Experimental Psychology: Human Perception and Performance*, 33(5), 1170-1176. <https://doi.org/10.1037/0096-1523.33.5.1170>
- Hasshim, N., & Parris, B. A. (2014). Two-to-one color-response mapping and the presence of semantic conflict in the Stroop task. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.01157>
- Hasshim, N., & Parris, B. A. (2015). Assessing stimulus–stimulus (semantic) conflict in the Stroop task using saccadic two-to-one color response mapping and prereponse pupillary measures. *Attention, Perception, & Psychophysics*, 77(8), 2601-2610. <https://doi.org/10.3758/s13414-015-0971-9>
- Hershman, R., & Henik, A. (2019). Dissociation between reaction time and pupil dilation in the Stroop task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(10), 1899–1909. <https://doi.org/10.1037/xlm0000690>
- Hershman, R., & Henik, A. (2020). Pupillometric contributions to deciphering Stroop conflicts. *Memory & Cognition*, 48(2), 325–333. <https://doi.org/10.3758/s13421-019-00971-z>
- Kalanthroff, E., Davelaar, E. J., Henik, A., Goldfarb, L., & Usher, M. (2018). Task conflict and proactive control: A computational theory of the Stroop task. *Psychological Review*, 125(1), 59–82. <https://doi.org/10.1037/rev0000083>
- Kinoshita, S., Mills, L., & Norris, D. (2018). The semantic Stroop effect is controlled by endogenous attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(11), 1730–1742. <https://doi.org/10.1037/xlm0000552>
- Klein, G. S. (1964). Semantic power measured through the interference of words with color-naming. *The American journal of psychology*, 77(4), 576-588. 588. <https://doi.org/10.2307/1420768>

- Küper, K., & Heil, M. (2012). Attentional Focus Manipulations Affect Naming Latencies of Neutral But Not of Incongruent Stroop Trials. *Swiss Journal of Psychology*, 71(2), 93-100. <https://doi.org/10.1024/1421-0185/a000075>
- Labuschagne, E. M., & Besner, D. (2015). Automaticity revisited: When print doesn't activate semantics. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00117>
- Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139087759>
- 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. (2005). The Stroop task in cognitive research. In A. Wenzel & D. C. Rubin (Eds.), *Cognitive methods and their application to clinical research* (pp. 17–40). American Psychological Association. <https://doi.org/10.1037/10870-002>
- MacLeod, C. M., & MacDonald, P. A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in cognitive sciences*, 4(10), 383–391. [https://doi.org/10.1016/s1364-6613\(00\)01530-8](https://doi.org/10.1016/s1364-6613(00)01530-8)
- 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>
- Monahan, J. S. (2001). Coloring Single Stroop Elements: Reducing Automaticity or Slowing Color Processing? *The Journal of General Psychology*, 128(1), 98-112. <https://doi.org/10.1080/00221300109598901>
- 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*. (p. 69-93). American Psychological Association. <https://doi.org/10.1037/10394-005>
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, 36(3), 516-524. <https://doi.org/10.3758/BF03195598>
- Parris, B. A. (2014). Task conflict in the Stroop task: When Stroop interference decreases as Stroop facilitation increases in a low task conflict context. *Frontiers in Psychology*, 5, 1182. <https://doi.org/10.3389/fpsyg.2014.01182>

- Parris, B. A., Sharma, D., & Weekes, B. (2007). An Optimal Viewing Position Effect in the Stroop Task When Only One Letter Is the Color Carrier. *Experimental Psychology*, 54(4), 273-280.
<https://doi.org/10.1027/1618-3169.54.4.273>
- Parris, B. A., Augustinova, M., & Ferrand, L. (2020). *The Locus of the Stroop Effect*. Frontiers Media SA, Lausanne. <https://doi.org/10.3389/978-2-88963-445-3>
- Parris, B. A., Hasshim, N., Wadsley, M., Augustinova, M., & Ferrand, L. (2021). The loci of Stroop effects: A critical review of methods and evidence for levels of processing contributing to colour-word Stroop effects and the implications for the loci of attentional selection. *Psychological Research*.
<https://doi.org/10.1007/s00426-021-01554-x>
- Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review*, 110(1), 88-125. <https://doi.org/10.1037/0033-295X.110.1.88>
- Roelofs, A. (2014). Tracking eye movements to localize Stroop interference in naming: Word planning versus articulatory buffering. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(5), 1332-1347.
- Schmidt, J. R. (2013). Questioning conflict adaptation: Proportion congruent and Gratton effects reconsidered. *Psychonomic Bulletin & Review*, 20(4), 615–630. <https://doi.org/10.3758/s13423-012-0373-0>
- Schmidt, J. R., & Cheesman, J. (2005). Dissociating stimulus stimulus and response-response effects in the Stroop task. *Canadian Journal of Experimental Psychology*, 59(2), 132–138.
<https://doi.org/10.1037/h0087468>
- Seymour, P. H. K. (1977). Conceptual encoding and locus of the Stroop effect. *Quarterly Journal of Experimental Psychology*, 29(2), 245–265. <https://doi.org/10.1080/14640747708400601>
- Shichel, I., & Tzelgov, J. (2018). Modulation of conflicts in the Stroop effect. *Acta Psychologica*, 189, 93-102.
<https://doi.org/10.1016/j.actpsy.2017.10.007>
- 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>
- Zhang, H., & Kornblum, S. (1998). The effects of stimulus-response mapping and irrelevant stimulus-response and stimulus-stimulus overlap in four-choice Stroop tasks with single-carrier stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 24(1), 3-19.
- Zhang, H. H., Zhang, J., & Kornblum, S. (1999). A parallel distributed processing model of stimulus–stimulus and stimulus–response compatibility. *Cognitive Psychology*, 38(3), 386- 432.

6. Supplementary Materials

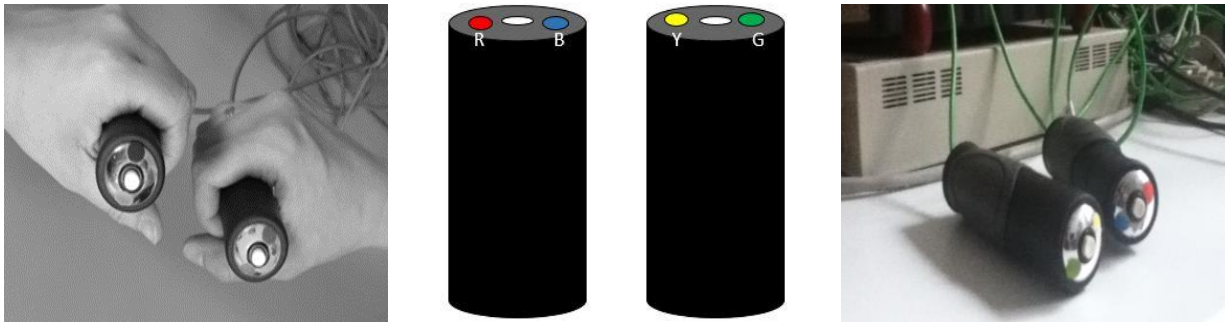


Figure S1. Illustration of Modified SRBox® called Chambox

Table S1. Analysis of RTs in 4 (Stimulus-Type: DR-incongruent trials vs. SR-incongruent trials vs. neutral DR-trials vs. neutral SR-trials) \times 2 (Coloring: ALCC vs. SLCC) Bayesian Repeated Measures ANOVA

Model Comparison					
Models	P(M)	P(M data)	BF _M	BF ₁₀	error %
Null model (incl. subject)	0.200	1.738e -21	6.953e -21	1.000	
Coloring + Stimulus_Type	0.200	0.618	6.467	3.555e +20	2.028
Coloring + Stimulus_Type + Coloring * Stimulus_Type	0.200	0.375	2.399	2.157e +20	1.592
Stimulus_Type	0.200	0.007	0.029	4.160e +18	0.590
Coloring	0.200	5.080e -20	2.032e -19	29.223	1.062

Note. All models include subject

Table S2. Analysis of percentages of errors in 4 (Stimulus-Type: DR-incongruent trials vs. SR-incongruent trials vs. neutral DR-trials vs. neutral SR-trials) \times 2 (Coloring: ALCC vs. SLCC) Bayesian Repeated Measures ANOVA

Model Comparison					
Models	P(M)	P(M data)	BF _M	BF ₁₀	error %
Null model (incl. subject)	0.200	0.004	0.017	1.000	
Stimulus_Type	0.200	0.779	14.074	182.463	0.744
Stimulus_Type + Coloring	0.200	0.213	1.084	49.974	20.607
Stimulus_Type + Coloring + Stimulus_Type * Coloring	0.200	0.003	0.012	0.672	15.472
Coloring	0.200	8.983e -4	0.004	0.210	1.665

Note. All models include subject

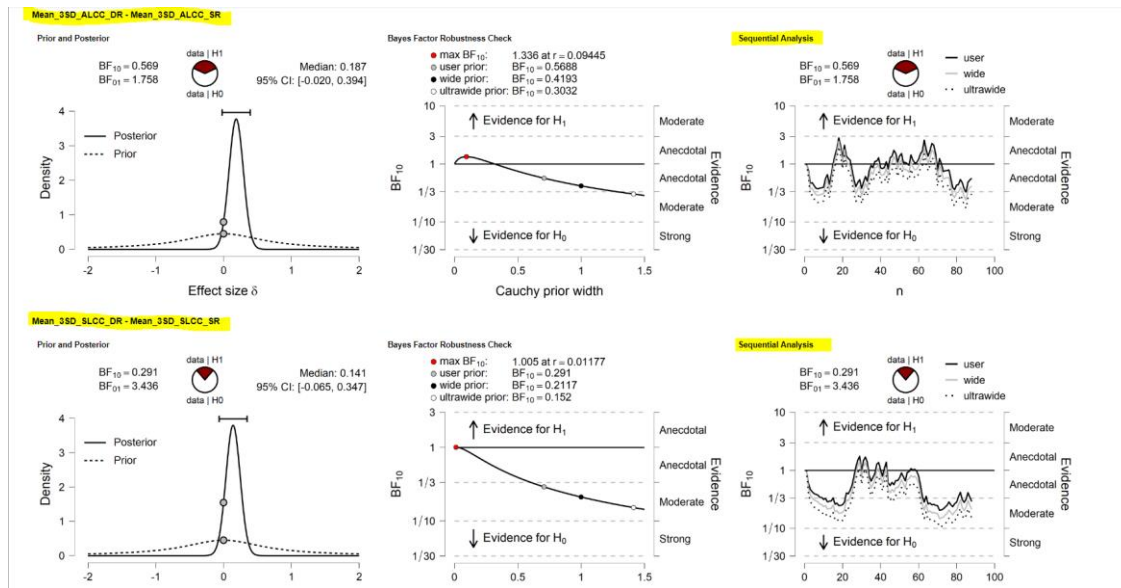


Figure S2. Sequential analysis of RC in ALCC vs. SLCC condition (probability of finding a RC if more observations are added)

Analysis of slow vs. fast trials

Given that interference (i.e., conflict effect) is known to increase proportionally with slow RTs (De Jong et al., 1999; Hasshim, Bate, Downes & Parris 2019) for each level of Stimulus and of Coloring we binned slow vs. fast trials. While the 3-way Stimulus-Type \times Coloring \times Bin interaction remained non-significant [$F(3,261) = .981$; $p = .402$, $\eta_p^2 = 0.011$], the main effect of Stimulus-Type [$F(3,261) = 21.57$; $p < .001$, $\eta_p^2 = 0.199$] and that of Bin [$F(3,87) = 881.03$; $p < .001$, $\eta_p^2 = 0.910$] were.

We also observed a Stimulus-Type \times Bin interaction [$F(3,261) = 15.60$; $p < .001$, $\eta_p^2 = 0.152$]. Once decomposed, in slow trials, both conflicts are present, whereas in fast trials no conflict is significant (see the Table 3S below). Thus, the lack of response conflict in the averaged data was driven by the lack of response conflict in the fast trials. Again, the additional sequential analyses of response conflict in fast vs. slow trials suggested that the probability of finding this type conflict if more observations were added, is unlikely (see Figure 3 here below).

Table 3S. *Color-Identification Performance (Means, Standard Error, 95%-Confidence Intervals) observed as a Function of Stimulus × Bin interaction*

	Fast Bin		Slow Bin		Bin effect
	M (SE)	CI	M (SE)	CI	
Color Incongruent Different Response (DR)	570 (9)	[553,587]	878 (17)	[844,912]	-308**
Color Neutral Different Response (Neutral DR)	552 (7)	[538,566]	818 (14)	[790,845]	-266**
Color Incongruent Same Response (SR)	563 (8)	[548,579]	863 (17)	[829,898]	-300**
Color Neutral Same Response (Neutral SR)	558 (8)	[543,574]	830 (17)	[796,864]	-272**
	M (SD)	CI	M (SD)	CI	
Standard Stroop Interference (DR – Neutral DR)	18**/†† (3)	[-1,14]	60**/†† (7)	[45,75]	
Semantic Stroop Interference (SR – Neutral SR)	5 ^{ns} / ^{ns} (4)	[-2,13]	33**/† (10)	[14,53]	
Response Conflict (DR-SR)	7 [°] / ^{ns} (4)	[-1,14]	15 [*] / ^{ns} * (7)	[.57, 29]	

Note. Inferential approach: ^{ns}non-significant; [°]marginally significant ($p < .10$), ^{*}significant at $p < .05$; ^{**}significant at $p < .01$; [†]significant at $p < .001$; Bayesian approach (presented after the slash): ^{ns}no evidence of an effect, with BF_{01} value between 1–3; ^{ns*}anecdotal evidence in favor of H1, with BF_{10} value between 1–3; ^{*}moderate evidence in favor of H1, with BF_{10} value between 3–10; ^{**}strong evidence in favor of H1, with BF_{10} value between 10–30; [†]very strong evidence in favor of H1, with BF_{10} between 30–100; ^{††}extreme evidence in favor of H1/of an effect, with $BF_{10} < 100$.

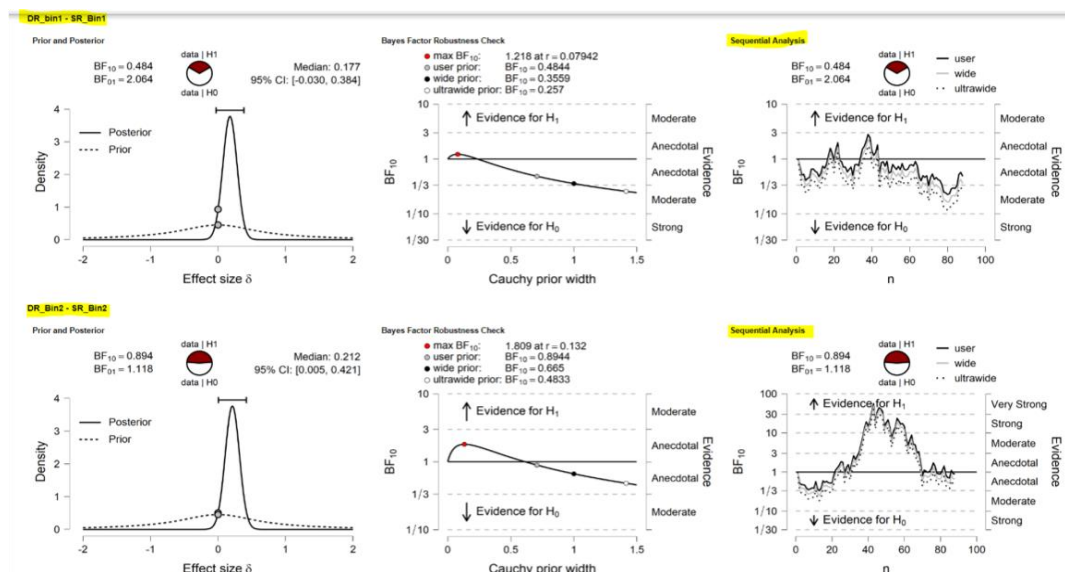


Figure S3. *Sequential analysis of RC in fact vs. slow trials (probability of finding a RC if more observations are added)*

Table S3. *Analysis of Magnitudes of Interference in 2 (Interference-Type: Standard vs. Semantic) × 2*

(Coloring: ALCC vs. SLCC) Bayesian Repeated Measures ANOVA

Models	P(M)	P(M data)	BF _M	BF ₁₀	error %
Null model (incl. subject)	0.200	0.039	0.162	1.000	
Interference_Type + Coloring	0.200	0.478	3.669	12.269	1.232
Coloring	0.200	0.204	1.028	5.243	0.777
Interference_Type + Coloring + Interference_Type * Coloring	0.200	0.192	0.951	4.926	4.855
Interference_Type	0.200	0.086	0.377	2.207	1.297

Note. All models include subject

CHAPTER TWO: SOME FURTHER CLARIFICATIONS ON AGE-RELATED DIFFERENCES IN THE STROOP TASK: NEW EVIDENCE FROM THE TWO-TO-ONE STROOP PARADIGM¹

1. Introduction

The *Stroop interference effect* (i.e., longer color-identification times for color-incongruent [e.g., “BLUE” displayed in yellow] than for color-neutral words [e.g., the word “DEAL” displayed in yellow]) is generally larger in healthy older adults than in their younger counterparts (see Comalli et al., 1962, for the first empirical demonstration). Also, and importantly, this age-effect in the Stroop task (Stroop, 1935) persists even after controlling for differences in processing-speed (e.g., Aschenbrenner et al., 2015, 2017; Bugg et al., 2007; Jackson & Balota, 2013; Nicosia & Balota, 2020; Spieler et al., 1996). It is therefore thought to reflect an inhibition deficit (e.g., Hasher & Zacks, 1988) due to which older adults are less efficient at suppressing the word-dimension of color-incongruent Stroop words, leading them to experience greater competition at the response output stage (Spieler et al., 1996).

Indeed, according to dominant single-stage response competition models (e.g., Roelofs, 2003), incidental semantic processing of the irrelevant word-dimension of color-incongruent Stroop items generates a single type of conflict: response conflict. According to this view, the Stroop interference effect is considered a unitary phenomenon due solely to competition between two alternative responses indicated by the two dimensions of the Stroop stimulus. In contrast, *multi-stage* models anticipate this incidental processing to generate an additional level of conflict at the level of semantics: *semantic* conflict (e.g., Zhang et al.,

¹ Corresponds to Burca M., Chausse P., Ferrand L., Parris B. A., Augustinova M. (2022). Some further clarifications on age-related differences in the Stroop task: new evidence from the two-to-one Stroop paradigm. *Psychonomic Bulletin and Review*, 29, 492–500. <https://doi.org/10.3758/s13423-021-02011-x>

1999; Zhang & Kornblum, 1998). They therefore view the Stroop interference effect as a composite phenomenon comprising both response and semantic conflict.

Taking this idea as their starting point, several studies have set out to investigate the level of processing (e.g., response and/or semantic) at which the age-related differences in the Stroop task take their effects and, more specifically, whether semantic conflict is or is not affected by healthy aging. Indeed, the idea proffered by Spieler and colleagues that older adults are less efficient in suppressing the word-dimension of Stroop stimuli leads to the somewhat straightforward prediction that they should (also) experience a greater amount of semantic conflict. This is not what studies have found.

Li and Bosman (1996) and, later, Augustinova and colleagues (2018) reported greater magnitudes of standard Stroop interference (e.g., $BLUE_{\text{yellow}} - DEAL/****_{\text{yellow}}$) in healthy older adults, but neither study reported age-related differences in the magnitude of semantic-associative Stroop interference (e.g., $SKY_{\text{yellow}} - DEAL/****_{\text{yellow}}$)². Augustinova and colleagues (2018) subsequently claimed that the locus of the age-effect in the Stroop task is at the level of response conflict rather than the level of semantic conflict or a combination of the two. Contrary to past conceptualizations (e.g., Spieler et al., 1996), these results imply that both older and younger participants are actually equally (in)efficient at suppressing the word-dimension of Stroop stimuli. In line with the most recent contributions to the literature on the above-mentioned inhibition deficit (e.g., Rey-Mermet & Gade, 2018), it further implies that older participants are rather less efficient in inhibiting the irrelevant response that is primed by the (irrelevant) word-dimension. This in turn reinforces the idea that the age-related deficit in inhibition (e.g., Andrés et al., 2008), or, more broadly, the age-related deficit in cognitive control, is not general (e.g., Bugg, 2014).

² To control for differences in processing-speed, raw naming latencies were proportionally transformed in these studies into percentages of standard ($[(M_{\text{standard color-incongruent RT}} - M_{\text{color-neutral RT}}) / M_{\text{color-neutral RT}}] * 100$) and semantic Stroop interference ($[(M_{\text{color-associated incongruent RT}} - M_{\text{color-neutral RT}}) / M_{\text{color-neutral RT}}] * 100$).

However, single-stage response competition models argue that semantic-associative interference ($SKY_{\text{yellow}} - DEAL_{\text{yellow}}$) measured in these prior studies results entirely from response conflict (e.g., Roelofs, 2003). According to this position, semantic associates elicit incorrect *response* activity (e.g., say “blue”/press blue for SKY_{yellow}) indirectly – through their association with the response-set colors (blue in this case) – which explains in turn the smaller magnitude of semantic-associative interference ($SKY_{\text{yellow}} - DEAL_{\text{yellow}}$) compared to its standard ($BLUE_{\text{yellow}} - DEAL_{\text{yellow}}$) counterpart. Under this account, neither Li and Bosman’s (1996) nor Augustinova and colleagues’ (2018) studies satisfactorily demonstrated that the type of conflict that is spared by healthy aging is *semantic* (i.e., due specifically to a slowdown that occurs whenever two distinct, yet closely related semantic representations are simultaneously activated in an amodal semantic network, see e.g., Seymour, 1977 for discussion).

To address this issue directly, the present study replaced semantic-associative items with items which induce semantic conflict in a way that cannot be accounted for by single-stage response competition models. Specifically, the study employed the *two-to-one Stroop paradigm* (De Houwer, 2003, hereafter 2-to-1). In this paradigm, all the distractors are part of the response set (e.g., BLUE, RED, GREEN, YELLOW), while responses for paired target colors are mapped to only one response-key (e.g., ‘f’ for blue and red and ‘j’ for green and yellow). As a result of this response-mapping, standard incongruent Stroop trials like $BLUE_{\text{yellow}}$ provide evidence toward two different responses (they are therefore termed *different-response trials*). Indeed, relevant color-dimension (YELLOW) prompts the correct response activity toward ‘j’ key, whilst the irrelevant word-dimension (BLUE) prompts the incorrect response activity toward ‘f’ key. There is no such (response) conflict on trials like $BLUE_{\text{red}}$ since both dimensions of the Stroop stimulus provide evidence toward the same response. Consequently, significant *interference* generated by these so-called *same-response*

trials is interpreted as representing the independent contribution of semantic conflict to overall Stroop interference (De Houwer, 2003, see e.g., Hershman & Henik, 2020 for the most recent example).

However, with the exception of a few notable studies (see below), all studies employing this measure of semantic conflict – including De Houwer (2003) – have used color-congruent trials as the baseline against which semantic conflict is measured. Problematically, the difference between same-response and color-congruent trials could be entirely driven by facilitation on color-congruent trials and thus not involve any semantic conflict (Hasshim and Parris 2014, 2015) – as unitary models of Stroop interference (Roelofs, 2003) would predict. In line with this interpretation, Hasshim and Parris consistently reported significantly longer RTs for same-response trials than for color-congruent trials, but no difference between same-response trials and trials that were free of facilitation (i.e., color-neutral word trials, see e.g., Brown, 2012; MacLeod, 1991 for discussion)³.

In contrast to Hasshim and Parris, Burca and colleagues' study (2021, see Chapter 1) reported a significant difference between same-response and color-neutral trials. This suggests that the difference between same-response and color-congruent trials (i.e., when no color-neutral baseline is included) simply confounds the (*semantic*) *conflict* produced by same-response trials and *facilitation* produced by color-congruent trials (MacLeod, 1991). However, the extent to which this is actually the case remains uncertain, since Burca et al.'s study did not include color-congruent trials. As a result, no study has so far demonstrated that semantic conflict contributes to overall Stroop interference in the 2-to-1 Stroop paradigm independently of both response conflict and facilitation. Considering this as a necessary prerequisite for any

³ The absence of semantic conflict was supported further by Bayesian evidence for the null-hypothesis and by the unchanged magnitude of associated pre-response pupillometric measures of effort (i.e., a reliable measure of the potential differences between conditions, Hasshim & Parris, 2015).

empirical demonstration of the specific age-effect (or lack thereof) on semantic vs. response conflict in the Stroop task, the present study aimed to address this more fundamental issue.

The Present Study

To this end, items that are traditionally included in the 2-to-1 Stroop paradigm (De Houwer, 2003) were supplemented by color-neutral word trials (Hasshim & Parris, 2014). This addition enabled us to test adequately for the presence of semantic conflict predicted by the multi-stage models of Stroop interference (e.g., Zhang & Kornblum, 1998) that were favored a priori in the current study over the still-dominant single-stage response competition models (e.g., Roelofs, 2003). With this design, the study was able to more unambiguously measure age-related differences in response and semantic conflict. Consequently, if, as reported by past studies (Augustinova et al., 2018; Li & Bosman, 1996), semantic conflict (same-response trials – color-neutral trials) is indeed spared in healthy aging, its magnitude will not differ between young and old adults. In contrast, response conflict (different-response – same-response trials) will be greater in healthy aging adults as compared to their younger counterparts.

2. Method

Participants and Design

Fifty-one older (i.e., over 65 years old) and fifty younger (i.e., below 35 years old) native French-speakers reporting normal or corrected-to-normal vision and presenting no impairment in color discrimination initially volunteered to participate in the study approved by the local ethics committee. One older participant presented a medical history that included a head injury and one other was undergoing a medical treatment for depression. Six months prior to inclusion in the study, none of the other participants suffered from other psychiatric

and/or neurological disorders. None of them declared taking any drug and/or following any medical treatment that is known to impact the nervous system during the 48h prior to inclusion. To ensure that the remaining participants fitted the inclusion criteria, they completed a psychometric evaluation battery. To this end, the older adults completed the Mini Mental State Examination (Folstein, 1975). The scores of two participants were lower than the cutoff score of 25 points. The older adults also completed the Frontal Assessment Battery (Dubois et al., 2000). None of them presented with a cutoff score of 16 (or 15 depending on the participant's sociocultural level). A depression scale was then administered to both the older and younger adults. No older adults reached the cutoff score of 7 on the short version (15 items) of the Geriatric Depression Scale (Sheikh & Yesavage, 1986). In addition, none of the younger adults reached the cutoff score of 8 on Beck's Depression Inventory (Beck, 1988). In both groups, working memory was assessed with the forward and backward digit span (WAIS, Wechsler, 2008). All participants had scores within the norm, recalling seven plus or minus two items. Finally, to further assess differences in processing speed, the French equivalent (Bugaiska et al., 2007) of the letter-comparison test (Salthouse, 1990) was administered in both age-groups. After the exclusion of 5 participants in total (one was unable to perform the manual 2-to-1 Stroop task due to reduced hand mobility), the Stroop data of forty-six healthy older (36 females and 10 males; $M_{\text{age}} = 74.04$) and fifty younger adults (41 females and 9 males; $M_{\text{age}} = 21.48$) were analyzed in a 4 (Stimulus-Type: different-response vs. same-response vs. neutral vs. congruent) \times 2 (Age-Group: older vs. younger) ANOVA, with the former factor as within-participants factor.

Apparatus, Stimuli and Procedure

After the psychometric evaluation presented above, the participants completed a computerized version of the Stroop Task run using Eprime 2.0 software (Schneider et al.,

2002). The participants were seated 70cm in front of a 13" portable computer and instructed to identify the color of the stimulus presented on the screen, as quickly and accurately as possible, by pressing the appropriate color-button and to ignore everything else in the display. To this end, they were instructed to concentrate on the fixation cross ("+") that appeared for 2000msec in the center of the screen at the beginning of each trial. The stimulus remained on the screen until the participant responded or until 3500msec had elapsed.

All stimuli were presented in lowercase Courier font, size 18, on a black background and subtended an average visual angle of 0.9° high × 3.0° wide. The participants responded manually using a modified SRBox® consisting of two handles, each of which had a single response button at the top flanked by two color-stickers (blue and red on one handle, yellow and green on the other, see Supplementary Materials). The participants pushed these response buttons with their thumbs. This allowed them to hold each handle comfortably in their palms with the remaining four fingers. The placement of the handles in the right or left hand, respectively, was counterbalanced across participants.

To familiarize themselves with the color-button correspondence before completing the experimental block, the participants first completed 96 practice trials consisting of asterisks. Due to the low accuracy rate, 8 older participants had to repeat this practice block (3 of them were later excluded from further analyses) before proceeding to the experimental trials. As in Hasshim and Parris (2014, Exp. 2A), these consisted of 96 different-response, 48 same-response, 48 color-neutral and 48 color-congruent trials. The trials were randomly intermixed in a single block. To this end, four (French) color-words: *rouge* [red], *jaune* [yellow], *bleu* [blue], and *vert* [green] presented in both congruent and incongruent colors and four non-color words: *plomb* [lead], *liste* [list], *page* [page], *cave* [basement] presented in all the colors were used. They were paired on length and frequency via *Lexique 3.38* (New et al., 2004).

3. Results and Discussion

Five older participants were excluded from further analyses: one due to faulty recording, and the four others due to the fact that more than 33% of their data was removed from the analysis after the 3SD correction and the exclusion of the wrong answers (see Table 1S in the Supplementary Materials for demographic and psychometric data of the remaining participants). *RTs* greater than 3SDs above or below each participant's mean latency for each condition were excluded from the analysis (i.e., less than 2% of the total data, corresponding to 0.9% of younger adults' data and 1.5% of older adults' data). Consequently, *RTs* and errors of the remaining 91 participants (41 older and 50 younger) were first analyzed in an omnibus 4 (Stimulus-Type: different-response vs. same-response vs. neutral vs. congruent) \times 2 (Age-Group: older vs. younger) standard and Bayesian ANOVA. The values for this latter ANOVA were calculated with JASP (JASP Team, 2017) and interpreted according to Lee and Wagenmakers (2013 adjusted from Jeffreys, 1961). All priors were equal. Recall that further reported BF_{10} is the Bayes factor giving the evidence for H1 over the null hypothesis (H0), whereas BF_{01} is evidence for H0 over H1.

For errors (see Table 1), these analyses revealed a main effect of Stimulus-Type [$F(3,267) = 19.03$; $p < .001$, $\eta_p^2 = 0.176$; $BF_{10} = 4.450e+7$], but not of Age-Group [$F(1,89) = .018$; $p = .894$, $\eta_p^2 < .000$; $BF_{10} = 0.227/BF_{01} = 4.396$]. The Stimulus Type \times Age-Group interaction was also significant [$F(3,267) = 3.11$; $p = .041$, Greenhouse-Geisser corrected, $\eta_p^2 = 0.034$; $BF_{10} = 1.130/BF_{01} = 0.884^4$]. However, the BF evidence in favor of an interaction was only anecdotal⁵.

⁴ Since the different models that a Bayesian ANOVA compares against a null model never include interaction alone, the BF values reported for all the interactions correspond to values obtained by dividing the BF value of the model containing the two main effects and their interaction by the BF value of the model with the two main effects only.

⁵ The simple-main effect of Stimulus-Type was significant in both older [$F(3,87) = 9.09$; $p < .001$, $\eta_p^2 = 0.239$; $BF_{10} = 66195.29$] and younger [$F(3,87) = 3.26$; $p = .025$, $\eta_p^2 = 0.101$; $BF_{10} = 35.97$] participants. Unsurprisingly, and in line with the main effect of Stimulus-Type, further pairwise comparisons revealed that the most errors in both age-groups were committed for different-response incongruent items (see Table 1 for descriptive statistics and simple main-effects of Age-group). However, in the younger adults, %ER for these

Given that the analysis of *RTs* showed a considerable but expected (see Table 2S in the Supplementary Materials) general slowing in older adults (i.e., the significant Stimulus-Type \times Age-Group interaction [$F(3, 267) = 14.78$; $p < .001$; $\eta_p^2 = 0.142$; $BF_{10} = 1.378e+6$]), which was qualified by a significant simple main effect of Age-Group for each type of Stimulus (all p 's $< .001$, see Table 2S in the Supplementary Materials), these *RTs* were z-scored (e.g., Jackson & Balota, 2013). The same omnibus ANOVA then revealed a main effect of Stimulus-Type [$F(3,267) = 128.59$; $p < .001$, $\eta_p^2 = 0.591$; $BF_{10} = 1.459e+59$] that was also included in the significant Stimulus-Type \times Age-Group interaction [$F(3,267) = 10.36$; $p < .001$, $\eta_p^2 = 0.104$, $BF_{10} = 706286.31$], thus indicating that age-related differences persist even after controlling for generalized slowing (see Table 1).

Is there any Semantic Conflict in the two-to-one Stroop paradigm?

To answer this key question, we first analyzed the aforementioned main effect of Stimulus-Type. This analysis revealed that, as in De Houwer's original study, the total Stroop effect ($M_{\text{different-response}} - M_{\text{congruent}}$, $p < .001$; $BF_{10} = 1.814e+24$) resulted from a significant contribution of both response conflict ($M_{\text{different-response}} - M_{\text{same-response}}$; $p < .001$; $BF_{10} = 4.134e+11$) and the difference between same-response and congruent trials ($p < .001$; $BF_{10} = 2.880e+10$) – taken in previous studies as evidence for semantic conflict. However, the crucial addition of color-neutral trials enabled us to show that, overall, this latter difference did indeed confound the contribution of semantic conflict ($M_{\text{same-response}} - M_{\text{neutral}}$; $p < .001$; $BF_{10} = 27038.729$) and that of Stroop facilitation ($M_{\text{neutral}} - M_{\text{congruent}}$, $p = .016$), which was moderate ($BF_{10} = 7.835$).

latter items differed only marginally from those observed for color-congruent ones – yielding only a marginally significant overall Stroop effect ($M_{\text{different-response}} - M_{\text{congruent}}$) on errors ($p = .062$; $BF_{10} = 1.751/BF_{01} = 0.571$).

Table 1. *Color-Identification Performance (Means, Standard Error, 95%-Confidence Intervals and Percent Errors) observed as a Function of Stimulus and Age*

Z-Scored RT's	Younger Adults			Older Adults			Age effect on %ER
	M (SE)	CI	%ER	M (SE)	CI	%ER	
Different Response (DR) Color-Incongruent items	.208 (.020)	[.168, .248]	3.90	.306 (.022)	[.261, .350]	5.29	1.39 ^{ns}
Same Response (SR) Color-Incongruent items	.000 (.021)	[-.042, .041]	2.04	.093 (.023)	[.047, .138]	2.13	-.09 ^{ns}
Color-Neutral items	-.080 (.021)	[-.121, -.038]	2.46	-.154 (.023)	[-.200, -.109]	2.19	.27 ^{ns}
Color-Congruent items	-.127 (.019)	[-.165, -.090]	2.71	-.244 (.021)	[-.285, -.203]	1.83	.88 ^{ns}
<i>Amplitudes (based on Z-Scored RT's)</i>	M (SE)	CI		M (SE)	CI		Age effect
Stroop Effect (DR - Congruent)	.335 ^{**/††} (.034)	[.267, .403]		.550 ^{**/††} (.038)	[.475, .624]		+ .214 ^{**/††}
Facilitation Effect (Neutral - Congruent)	.047 ^{ns} (.030)	[-.012, .106]		.089 ^{*/ns*} (.033)	[.024, .155]		+ .042 ^{ns}
Stroop Interference Effect (DR - Neutrals)	.288 ^{**/††} (.033)	[.221, .354]		.460 ^{**/††} (.037)	[.387, .534]		+ .172 ^{*/†}
Response Conflict (DR - SR)	.208 ^{**/††} (.031)	[.146, .270]		.213 ^{**/††} (.034)	[.144, .281]		+ .005 ^{ns}
Semantic Conflict (SR - Neutrals)	.079 ^{*/ns*} (.037)	[.006, .153]		.247 ^{**/††} (.041)	[.166, .328]		+ .168 ^{*/**}

Note: ^{ns}non-significant; * significant at $p < .05$; ** significant at $p < .01$; † significant at $p < .001$; Bayesian approach (presented after the slash): ^{ns}no evidence of an effect, with BF_{10} value between 0–1; ^{ns*} anecdotal evidence in favor of H1, with BF_{10} value between 1–3; * moderate evidence in favor of H1, with BF_{10} value between 3–10; ** strong evidence in favor of H1, with BF_{10} value between 10–30; † very strong evidence in favor of H1, with BF_{10} between 30–100; †† extreme evidence in favor of H1/of an effect, with $BF_{10} < 100$.

This finding is consistent with MacLeod's reasoning (1991) that in the absence of color-neutral trials, the total Stroop effect ($M_{\text{different-response}} - M_{\text{congruent}}$) is likely to confound two qualitatively distinct phenomena: the Stroop interference ($M_{\text{different-response}} - M_{\text{neutral}}$) and facilitation ($M_{\text{neutral}} - M_{\text{congruent}}$) effects.

The decomposition of the Stimulus-Type \times Age-Group interaction further revealed that the simple main effect of Stimulus-Type was significant in both older [$F(3,87) = 76.86$; $p < .001$, $\eta_p^2 = 0.726$; $BF_{10} = 1.876e+33$] and younger [$F(3,87) = 35.65$; $p < .001$, $\eta_p^2 = 0.551$; $BF_{10} = 3.019e+26$] participants. Further pairwise comparisons conducted in both age-groups revealed that the significant total Stroop effect had the same structure, although excluding Stroop facilitation (see Table 1 for descriptive statistics and magnitudes), which was no longer significant in younger adults ($p = .114$, $BF_{10} = 0.621/BF_{01} = 1.610$).

The Stroop interference effect – which was significant in both age-groups (young group: $p < .001$; $BF_{10} = 5.746e+10$; older group: $p < .001$; $BF_{10} = 3.195e+10$) – again resulted from the significant contribution of semantic ($M_{\text{same-response}} - M_{\text{neutral}}$) and response ($M_{\text{different-response}} - M_{\text{same-response}}$) conflicts (see Table 1)⁶.

Taken together, these results are therefore consistent with the idea that both semantic conflict and response conflict contribute to Stroop interference. This pre-requisite being satisfied (see Introduction), we can now go on to investigate the extent to which these independent components of Stroop interference are influenced by healthy aging.

How does healthy aging influence semantic vs. response conflict in the Stroop task?

To address this issue, the magnitudes of semantic and response conflicts (see Table 1) were analyzed in a 2 (Conflict-Type) \times 2 (Age-Group: older vs. younger) ANOVA. This

⁶ It should be noted, however, that Bayesian evidence for semantic conflict in younger adults remained anecdotal ($BF_{10} = 1.893/BF_{01} = 0.528$) despite the fact that moderate evidence in support of such a conflict was found in a recent study by our research group (Burca et al., 2022)

revealed a non-significant main effect of Conflict-Type [$F(1,89) = 1.13$; $p = .292$, $\eta_p^2 = 0.012$; $BF_{10} = 0.481/BF_{01} = 2.078$] as well as a significant [$F(1,89) = 11.94$; $p = .001$, $\eta_p^2 = .118$], although anecdotal ($BF_{10} = 1.529/BF_{01} = 0.654$), main effect of Age-Group. It also revealed a marginally significant [$F(1,89) = 3.38$; $p = .069$, $\eta_p^2 = 0.037$], although anecdotal ($BF_{10} = 2.330/BF_{01} = 0.429$), Conflict-Type \times Age-Group interaction. Even though evidence for this interaction was only anecdotal, we decomposed it further by testing the simple main effect of Age-Group at each level of Conflict-Type. Contrary to our expectations, this effect was significant for semantic conflict [$F(1,89) = 9.288$; $p = .003$, $\eta_p^2 = 0.094$; $BF_{10} = 11.683/BF_{01} = 0.086$], with older adults presenting a much greater magnitude of semantic conflict than young adults. Additionally, and also contrary to our expectations, the simple main effect of Age-Group remained non-significant for response conflict [$F(1,89) = 0.010$; $p = .922$, $\eta_p^2 = 0.000$; with evidence for the null effect of aging, $BF_{10} = 0.222/BF_{01} = 4.512$ (see Table 1)]⁷. Thus, the present study clearly extends the dissociative nature of the age-effect to the two-to-one Stroop paradigm. However, completely unlike past studies using the semantic Stroop paradigm (Augustinova et al., 2018; Li & Bosman, 1996), it points to a greater magnitude of semantic conflict in older adults.

4. General Discussion and Conclusion

Given that in all past Stroop studies, *semantic* conflict was potentially confounded with either *response conflict* (e.g., when semantic-associative items [SKY_{blue}] are used to induce semantic conflict) or with *facilitation* (when color-congruent items [$BLUE_{blue}$] are used as a baseline to derive a magnitude for semantic conflict), its contribution to the Stroop

⁷ Additionally, the simple main effect of Conflict-Type was significant in younger adults [$F(1,89) = 4.67$; $p = .034$, $\eta_p^2 = 0.050$; $BF_{10} = 9.318$] but not in older adults [$F(1,89) = 0.275$; $p = .602$, $\eta_p^2 = 0.003$; $BF_{10} = 0.278/BF_{01} = 3.597$]. The additional $BF_{+0} = 3.338$ in younger adults indicates that they displayed more response than semantic conflict, whereas $BF_{01} = 5.298$ in older adults indicates comparable magnitudes of both conflicts (see Table 1).

interference effect has so far been uncertain. Using the 2-to-1 Stroop paradigm (De Houwer, 2003) with a color-neutral baseline, the present study clearly demonstrated that the contribution of semantic conflict is independent of both response conflict and Stroop facilitation. Therefore, the present study provides an unambiguous empirical basis for the composite nature of Stroop interference – as originally claimed by De Houwer (2003) based on the multi-stage models of Stroop interference (Zhang & Kornblum, 1998; Zhang et al., 1999)⁸.

Given that no such basis was available in past studies of age-related differences in the Stroop task (Augustinova et al., 2018; Li & Bosman, 1996), the present study also investigated the extent to which healthy aging influences these independent constituents of Stroop interference. The reported results suggest a dissociative pattern opposite to that reported in past studies: whilst response conflict was not affected by healthy aging, greater semantic conflict was found in older adults. Therefore, the pattern that these studies report could be due to less efficient control of this phonological processing in older adults. Such an effect would not have been observed in the present study due to the use of manual responses. Despite this, the issue surrounding the use of semantic-associative Stroop trials remains.

If, according to single-stage models of the Stroop task, the semantic associative Stroop trials used in these previous studies induce only indirect response conflict (e.g., Roelofs, 2003), then the only conclusion that can be drawn from the studies by Augustinova et al. (2018) and Li and Bosman's (1996) is that overall response conflict is greater in older adults but its indirect portion is unaffected by healthy aging. However, since the present study unequivocally documented the existence of semantic conflict for the first time, it now seems

⁸Note that the unambiguous presence of Stroop facilitation additionally implies that magnitudes of semantic conflict observed without color-neutral baseline are clearly inflated. This also concerns magnitudes of a general conflict – central in cognitive control studies (e.g., Egner et al., 2010) – since this type of conflict is inferred from the so-called Stroop congruency effect ($BLUE_{yellow} - BLUE_{blue}$) using the same color-congruent baseline.

reasonable to assume that both semantic-associative and same-response trials actually induce semantic conflict (but in unknown quantities for the former).

If we thus assume that the present and past studies mobilized the same processes (i.e., induced comparable levels of semantic conflict; Augustinova & Ferrand, 2014), the absence of an age-effect on semantic associative interference could be potentially linked to the method used to control for age-related general slowing. Indeed, proportional transformation – applied first by Li and Bosman (1996) and later by Augustinova et al. (2018) – might actually (and counterintuitively) create an advantage for older adults in the presence of slower RTs (Hedge et al., 2018). This spurious advantage is no longer present when general slowing is controlled by means of a more suitable transformation (i.e., z-scores; Faust et al., 1999; Hedge et al., 2018) applied in the present study. To address this possibility directly, the data from Augustinova et al. (2018) were z-scored and re-analyzed in the same way as the 2-to-1 data reported above (see Supplementary Materials for a full description and results of these analyses). In line with Hedge et al.'s reasoning about proportional transformation, not only did the originally significant Conflict-Type \times Aging interaction become non-significant, but the additional Bayesian analyses actually provided moderate evidence against this interaction. This suggests that the magnitudes of both semantic and response conflict in Augustinova et al.'s z-scored data tended to be greater in older adults than in their younger counterparts (see Table 3S).

While the results regarding semantic conflict are in line with those reported above, discrepancies remain regarding the effect of healthy aging on response conflict. Although these differences could be accounted for by the response mode difference highlighted above, we also conducted cross-study analyses on the merged data sets (see Supplementary Materials for a full description and results of the analyses). Again, Bayesian analyses provided moderate evidence against a Conflict-Type \times Aging interaction, suggesting that across two studies,

healthy aging affected both the semantic and response conflicts. It should, however, be noted that a Bayesian Independent Samples t-Test conducted for exploratory purposes actually revealed anecdotal evidence against the age-effect on response conflict (see Table 4S), a finding that appears consistent with the results obtained using the 2-to-1 paradigm reported above. Alternatively, it also remains plausible that response conflict is unaffected in the 2-to-1 Stroop paradigm, not because of its specific nature but simply because its magnitude (i.e., smaller in the manual task than in the vocal tasks used in past studies) is too small to be affected.

Although not our favored a priori hypothesis, the fact that the present study could have mobilized different processes compared to past studies emphasizes the importance of choosing the correct critical and control trials for measuring the variable under test. Of course, no measure is perfect and we must therefore consider a limitation of the 2-to-1 paradigm that could provide an alternative explanation for the apparently greater semantic conflict in older adults. Because both dimensions of same-response trials provide evidence towards the same response, they cannot (unlike semantic associates) generate response conflict. However, they can still produce response facilitation. This opens up the possibility that the larger difference between same-response and color-neutral trials observed in older adults in the present study could actually be driven by greater response *facilitation* in younger adults, and not greater semantic *conflict* in older adults. Nevertheless, while this account would directly predict greater Stroop facilitation (which involves both response and semantic facilitation) in younger adults, the present study actually reports the opposite – rendering this latter account unlikely.

To sum up, the present study has provided the clearest evidence yet of a contribution of semantic conflict to overall Stroop interference (see also Parris et al., 2021, for a thorough discussion of this issue). Moreover, this has enabled us to investigate the effect of healthy aging on the independent constituents of the composite Stroop interference effect. In contrast

to previous studies, the present study showed that semantic conflict is affected by healthy aging. This finding prompted a re-analysis of the data from a previous study (Augustinova et al. 2018) using a more suitable method of controlling for the effect of general slowing in healthy aging (the same method as that employed in the present study). This re-analysis revealed that, as indicated by the present study, there is evidence of modified semantic conflict in healthy aging. Whilst the two studies diverge on the issue of the effect of aging on response conflict, the difference might be explained by the fact that a vocal response mode was used in both Augustinova et al. (2018) and Li and Bosman (1996), giving rise to the possibility that the control of phonological processing is reduced in healthy aging. Although both studies converged on the issue of semantic conflict, we would still recommend that future studies use the 2-to-1 paradigm rather than the semantic-associates method given that only the results from the present study show an unambiguous effect of aging on semantic conflict. However, to address the still-open issue of the characteristics shared (or otherwise) between same-response and semantically associated trials, future studies could combine the two (Schmidt & Cheesman, 2005) and measure the interference they generate against a color-neutral word baseline with more response-sensitive measures (e.g., EMG, mouse-tracking). Given that these latter measures are also more sensitive to the actual time course of interference, they are particularly suitable for further addressing the age-related differences in the Stroop task. Indeed, the issue of the extent to which a greater magnitude of a given conflict is due specifically to its greater *activation* (i.e, lower attentional selectivity, also implying an age-related deficit in proactive control) or to its less efficient *resolution* (i.e, less efficient inhibitory control, also implying an age-related deficit in reactive control) as yet remains unresolved (see e.g., Coderre et al., 2011 for this type of distinction). In the light of past research demonstrating an age-related deficit in proactive (e.g., Braver et al., 2001) as opposed to reactive (e.g., Bugg, 2014) cognitive control, the first possibility seems more

plausible than the second. This reasoning is reinforced by the fact that healthy aging might actually amplify task conflict (i.e., a more general conflict that – for all readable Stroop items including color-neutral ones – derives from the simultaneous preparation of two task sets: word-reading vs. color-naming, e.g., Goldfarb & Henik, 2007; Kalanthroff et al., 2018). Although, the significant age-effect on z-scored color-neutral stimuli observed in the present study is consistent with this idea, future studies – which should include more appropriate measures of task conflict – will need to address these possibilities directly.

The significant magnitudes of both semantic and response conflict observed in both younger and older adults clearly suggest that the historically favored single-stage response accounts of the Stroop interference effect are likely to be obsolete (e.g., Augustinova et al., 2018; De Houwer, 2003; Risko et al., 2006). Also, and importantly, so too are the customary implementations of Stroop interference/effect ($BLUE_{green} - DEAL_{green} / BLUE_{blue}$) that are rooted in these unitary models and from which the involvement of response and semantic processes and their modulation are merely inferred. Thus, in conclusion, the present study strongly encourages both the development of new integrative models of the Stroop interference effect (i.e., models that make room for relatively new types of conflict, e.g., Parris et al., 2021 for discussion) and further empirical work addressing the processes underlying age-related differences in the Stroop task based on such integrative models.

5. References

- Andrés, P., Guerrini, C., Phillips, L. H., & Perfect, T. J. (2008). Differential Effects of Aging on Executive and Automatic Inhibition. *Developmental Neuropsychology*, 33(2), 101-123.
<https://doi.org/10.1080/87565640701884212>
- Aschenbrenner, A. J., & Balota, D. A. (2015). Interactive effects of working memory and trial history on Stroop interference in cognitively healthy aging. *Psychology and Aging*, 30(1), 1-8.
<https://doi.org/10.1037/pag0000012>

- Aschenbrenner, A. J., Balota, D. A., Weigand, A. J., Scaltritti, M., & Besner, D. (2017). The first letter position effect in visual word recognition: The role of spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, 43(4), 700-718. <https://doi.org/10.1037/xhp0000342>
- Augustinova, M., Clarys, D., Spatola, N., & Ferrand, L. (2018). Some further clarifications on age-related differences in Stroop interference. *Psychonomic Bulletin & Review*, 25(2), 767-774. <https://doi.org/10.3758/s13423-017-1427-0>
- Augustinova, M., & Ferrand, L. (2014b). Automaticity of Word Reading: Evidence from the Semantic Stroop Paradigm. *Current Directions in Psychological Science*, 23, 343-348. <https://doi.org/10.1177/0963721414540169>
- Beck A.T. (1988). Beck Hopelessness Scale. The Psychological Corporation.
- Braver, T. S., Barch, D. M., Keys, B. A., Carter, C. S., Cohen, J. D., Kaye, J. A., Janowsky, J. S., Taylor, S. F., Yesavage, J. A., Mumenthaler, M. S., Jagust, W. J., & Reed, B. R. (2001). Context processing in older adults: Evidence for a theory relating cognitive control to neurobiology in healthy aging. *Journal of Experimental Psychology: General*, 130(4), 746-763. <https://doi.org/10.1037/0096-3445.130.4.746>
- Brown, T. L. (2011). The relationship between stroop interference and facilitation effects : Statistical artifacts, baselines, and a reassessment. *Journal of Experimental Psychology: Human Perception and Performance*, 37(1), 85-99. <https://doi.org/10.1037/a0019252>
- Bugaiska, A., Clarys, D., Jarry, C., Tacconnat, L., Tapia, G., Vanneste, S., & Isingrini, M. (2007). The effect of aging in recollective experience: The processing speed and executive functioning hypothesis. *Consciousness and Cognition*, 16(4), 797-808. doi:10.1016/j.concog.2006.11.007
- Bugg, J. M. (2014). Evidence for the sparing of reactive cognitive control with age. *Psychology and Aging*, 29(1), 115–127. <https://doi.org/10.1037/a0035270>
- Bugg, J. M., DeLosh, E. L., Davalos, D. B., & Davis, H. P. (2007). Age Differences in Stroop Interference: Contributions of General Slowing and Task-Specific Deficits. *Aging, Neuropsychology, and Cognition*, 14(2), 155-167. <https://doi.org/10.1080/138255891007065>
- Burca M., Beauconsin V., Chausse P., Ferrand L., Parris B. A., Augustinova M. (2021). Is there semantic conflict in the Stroop task? Further evidence from the 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>

- Coderre, E. L., Conklin, K., & van Heuven, W. J. B. (2011). Electrophysiological measures of conflict detection and resolution in the Stroop task. *Brain Research, 1413*, 51-59.
<https://doi.org/10.1016/j.brainres.2011.07.017>
- Comalli, P. E., Wapner, S., & Werner, H. (1962). Interference Effects of Stroop Color-Word Test in Childhood, Adulthood, and Aging. *The Journal of Genetic Psychology, 100*(1), 47-53.
<https://doi.org/10.1080/00221325.1962.10533572>
- De Houwer, J. (2003). On the role of stimulus-response and stimulus-stimulus compatibility in the Stroop effect. *Memory & Cognition, 31*(3), 353-359. <https://doi.org/10.3758/BF03194393>
- Dubois, B., Slachevsky, A., Litvan, I., Pillon, B. (2000). The FAB: A Frontal Assessment Battery at bedside. *Neurology, 55*(11), 1621-1626. DOI: 10.1212/wnl.55.11.1621
- Egner, T., Ely, S. & Grinband, J. (2010). Going, going, gone: characterizing the time-course of congruency sequence effects. *Frontiers in Psychology, 1*, 154.
<https://doi.org/10.3389/fpsyg.2010.00154>
- Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information-processing rate and amount: Implications for group differences in response latency. *Psychological Bulletin, 125*, 777–799. <http://dx.doi.org/10.1037/0033-2909.125.6.777>
- Folstein, M., F. (1975). Mini Mental State: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatry Research, 12*(3), 189-198. DOI: 10.1016/0022-3956(75)90026-6
- Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193–225). San Diego, CA: Academic Press.
- Hasshim, N., & Parris, B. A. (2014). Two-to-one color-response mapping and the presence of semantic conflict in the Stroop task. *Frontiers in Psychology, 5*. <https://doi.org/10.3389/fpsyg.2014.01157>
- Hasshim, N., & Parris, B. A. (2015). Assessing stimulus–stimulus (semantic) conflict in the Stroop task using saccadic two-to-one color response mapping and prereponse pupillary measures. *Attention, Perception, & Psychophysics, 77*(8), 2601-2610. <https://doi.org/10.3758/s13414-015-0971-9>
- Hedge, C., Powell, G., & Sumner, P. (2018). The mapping between transformed reaction time costs and models of processing in aging and cognition. *Psychology and Aging, 33*(7), 1093-1104.
<http://dx.doi.org/10.1037/pag0000298>

- Hershman, R., & Henik, A. (2020). Pupillometric contributions to deciphering Stroop conflicts. *Memory & Cognition*, 48(2), 325-333. <https://doi.org/10.3758/s13421-019-00971-z>
- Jackson, J. D., & Balota, D. A. (2013). Age-related changes in attentional selection: Quality of task set or degradation of task set across time? *Psychology and Aging*, 28(3), 744-753. <https://doi.org/10.1037/a0033159>
- Li, K. Z. H., & Bosman, E. A. (1996). Age differences in stroop-like interference as a function of semantic relatedness. *Aging, Neuropsychology, and Cognition*, 3(4), 272-284. <https://doi.org/10.1080/13825589608256630>
- 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>
- 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*. (p. 69-93). American Psychological Association. <https://doi.org/10.1037/10394-005>
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, 36(3), 516-524. <https://doi.org/10.3758/BF03195598>
- Nicosia, J., & Balota, D. (2020). The consequences of processing goal-irrelevant information during the Stroop task. *Psychology and Aging*, 35(5), 663-675. <https://doi.org/10.1037/pag0000371>
- Parris, B. A., Hasshim, N., Wadsley, M., Augustinova, M., & Ferrand, L. (2021). The loci of Stroop effects: A critical review of methods and evidence for levels of processing contributing to color-word Stroop effects and the implications for the loci of attentional selection. *Psychological Research*. <https://doi.org/10.1007/s00426-021-01554-x>
- Rey-Mermet, A., & Gade, M. (2018) Inhibition in aging: What is preserved? What declines? A meta-analysis. *Psychonomic Bulletin & Review*, 25, 1695–1716. <https://doi.org/10.3758/s13423-017-1384-7>
- Risko, E. F., Schmidt, J. R., & Besner, D. (2006). Filling a gap in the semantic gradient: Color associates and response set effects in the Stroop task. *Psychonomic Bulletin & Review*, 13(2), 310-315. <https://doi.org/10.3758/BF03193849>
- Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review*, 110(1), 88-125. <https://doi.org/10.1037/0033-295X.110.1.88>
- Salthouse, T. A. (1990). Working memory as a processing resource in cognitive aging. *Developmental Review*, 10(1), 101–124. [https://doi.org/10.1016/0273-2297\(90\)90006-P](https://doi.org/10.1016/0273-2297(90)90006-P)

- Schmidt, J. R. (2013). The Parallel Episodic Processing (PEP) model: Dissociating contingency and conflict adaptation in the item-specific proportion congruent paradigm. *Acta Psychologica, 142*(1), 119-126. <https://doi.org/10.1016/j.actpsy.2012.11.004>
- Schmidt, J. R., & Cheesman, J. (2005). Dissociating stimulus-stimulus and response-response effects in the Stroop task. *Canadian Journal of Experimental Psychology, 59*(2), 132-138. <https://doi.org/10.1037/h0087468>
- Schneider W., Eschman A., Zuccolotto A. (2002). E-Prime user's guide. Pittsburgh, USA: Psychology Software Tools Inc.
- Seymour, P. H. K. (1977). Conceptual encoding and locus of the Stroop effect. *Quarterly Journal of Experimental Psychology, 29*(2), 245-265. <https://doi.org/10.1080/14640747708400601>
- Sheikh, J., I., & Yesavage, J., A. (1986). Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. *Clinical Gerontology: A Guide to Assessment and Intervention* (pp. 165-173). New York, NY: The Haworth Press.
- Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance, 22*(2), 461-479. <https://doi.org/10.1037/0096-1523.22.2.461>
- 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>
- Zhang, H., & Kornblum, S. (1998). The effects of stimulus-response mapping and irrelevant stimulus-response and stimulus-stimulus overlap in four-choice Stroop tasks with single-carrier stimuli. *Journal of Experimental Psychology. Human Perception and Performance, 24*(1), 3-19. <https://doi.org/10.1037/0096-1523.24.1.3>
- Zhang, H., Zhang, J., & Kornblum, S. (1999). A Parallel Distributed Processing Model of Stimulus-Stimulus and Stimulus-Response Compatibility. *Cognitive Psychology, 38*(3), 386-432. <https://doi.org/10.1006/cogp.1998.0703>
- Wechsler, D., Psychological Corporation., & PsychCorp (Firm). (2008). *WAIS-IV technical and interpretive manual*. San Antonio, Texas: Pearson.

6. Supplementary Materials

Table 1S. *Demographic Variables and Mean Psychometric Test Scores as a Function of Participant Group*

	Younger Adults N=50			Older Adults N=41			<i>F</i>	<i>p</i>	η_p^2
	<i>M</i>	<i>SD</i>	<i>CI</i>	<i>M</i>	<i>SD</i>	<i>CI</i>			
Age	21.48	4.03	[20.34, 22.62]	74.56	8.35	[71.92, 77.20]	1575.95	<.001	.95
Education (years)	13.84	2.07	[13.25, 14.43]	11.46	4.34	[10.09, 12.83]	11.74	=.001	.12
MMSE ¹				29.00	1.30	[28.59, 29.41]			
FAB ²				17.05	1.32	[10.06, 15.72]			
BDI/GDS ³	1.10	1.45	[.64, 1.55]	2.36	2.68	[1.60, 3.12]			
Forward Digit ⁴ Span	6.46	1.33	[6.08, 6.84]	5.51	1.79	[4.95, 6.08]	8.39	.005	.09
Backward Digit Span	4.98	1.19	[4.64, 5.32]	4.10	1.48	[3.63, 4.56]	17.54	.002	.10
Processing ⁵ speed	33.80	4.98	[32.39, 35.21]	23.88	5.14	[22.26, 25.50]	86.94	<.001	.49

Note 1. Mini Mental State Examination (Folstein, 1975). For all remaining participants, the score was higher than the cutoff score of 25 points.

Note 2. Frontal Assessment Battery (Dubois et al., 2000). For all remaining participants, the score was higher than the cutoff score of 16 (or 15 depending on the participant's sociocultural level)

Note 3. Beck's depression inventory (Beck, 1988)/Geriatric Depression Scale (Sheikh & Yesavage, 1986). For all remaining participants, the score was lower than the cutoff score of 8 and 7 points, respectively.

Note 4. The forward and backward digit span (WAIS, Wechsler, 2008) was used to assess working memory. All participants had scores within the norm, recalling seven plus or minus two items.

Note 5. The French equivalent (Bugajska et al., 2007) of the letter-comparison test (Salthouse, 1990) was used to assess differences in processing speed. These were further confirmed by RTs in the Stroop task (see Table 2S).

Table 2S. *Color-Identification Performance in raw RTs (Means, Standard Errors, 95% Confidence Intervals) observed as a Function of Stimulus and Age*

<i>RTs</i>	Younger Adults		Older Adults		Age effect (RTs)
	M (SE)	CI	M (SE)	CI	
Different response (DR) Color-Incongruent items	712 (24)	[664, 760]	977 (27)	[924, 1030]	+265 ^{**/††}
Same Response (SR) Color-Incongruent items	673 (24)	[625, 721]	933 (27)	[880, 985]	+260 ^{**/††}
Color-Neutral items	655 (18)	[619,692]	867 (23)	[827, 907]	+212 ^{**/††}
Color-Congruent items	649 (20)	[609, 690]	854 (23)	[809, 899]	+205 ^{**/††}
<i>Amplitudes (based on RT's)</i>	M (SE)	CI	M (SE)	CI	Age effect
Stroop Effect (DR – Congruent)	63 ^{**/††} (9)	[46, 79]	123 ^{**/††} (9)	[105, 141]	+60 ^{**/††}
Facilitation Effect (Neutral – Congruent)	6 ^{ns/ns} (6)	[-6, 19]	13 ^{°/ns} (7)	[-.544, 27]	+7 ^{ns/ns}
Stroop Interference Effect (DR – Neutrals)	56 ^{**/††} (9)	[39, 74]	110 ^{**/††} (10)	[90, 130]	54 ^{**/††}
Response Conflict (DR – SR)	39 ^{**/††} (7)	[26, 52]	44 ^{**/††} (8)	[30, 59]	5 ^{ns/ns}
Semantic Conflict (SR – Neutrals)	17 ^{°/*} (9)	[.112, 35]	66 ^{**/††} (10)	[46, 85]	48 ^{**/†}

Note: ^{ns}non-significant; [°]marginally significant at $p=.060$; ^{*}significant at $p < .05$; ^{**}significant at $p < .01$; [†]significant at $p < .001$; Bayesian approach (presented after the slash): ^{ns}no evidence of an effect, with BF_{10} value between 0–1; ^{ns*} anecdotal evidence in favor of H1, with BF_{10} value between 1–3; ^{*}moderate evidence in favor of H1, with BF_{10} value between 3–10; ^{**}strong evidence in favor of H1, with BF_{10} value between 10–30; [†] very strong evidence in favor of H1, with BF_{10} between 30 -100; ^{††}extreme evidence in favor of H1/of an effect, with $BF_{10} < 100$

Analyses of z-scored RT's from Augustinova et al. (2018)

An omnibus 3 (Stimulus-Type: incongruent vs. semantically-associated vs. neutral) \times 2 (Coloring: ALCC vs. SLCC) \times 2 (Age-Group: older vs. younger) ANOVA on the z-scored RTs of fifty-eight adults (29 younger vs. 29 older adults) showed a main effect of Stimulus-Type [$F(2, 112) = 140.55; p < .001; \eta_p^2 = 0.715; BF_{10} = 5.803e+41$]. The main effects of Coloring [$F(1, 56) = .000; p = 1.000; \eta_p^2 = 0.000; BF_{10} = 0.118/BF_{01} = 8.505$] and of Age-Group [$F(1, 56) = .128; p = .722; \eta_p^2 = 0.002; BF_{10} = 0.138/BF_{01} = 7.270$] proved to be non-significant. Both the Stimulus-Type \times Age-Group [$F(2, 112) = 6.88; p = .002, \eta_p^2 = 0.109, BF_{10} = 25.21$] and Stimulus \times Coloring [$F(2, 112) = 20.53; p < .001, \eta_p^2 = 0.268, BF_{10} = 34427.67$] interactions were significant, with the BFs pointing to strong and extreme evidence, respectively, in favor of these interaction effects alone¹. Finally, the 3-way Stimulus-Type \times Coloring \times Age-Group interaction was non-significant [$F(2, 112) = 1.28; p = .283, \eta_p^2 = 0.022, BF_{10} = 0.227/BF_{01} = 4.391$], with the BFs providing moderate evidence in favor of the null hypothesis of an effect of the interaction alone.

The decomposition of the main effect of Stimulus-Type showed that, in line with results originally reported by Augustinova et al. (2018), both standard (e.g., *BLUE*_{yellow} – *DEAL*/****_{yellow}) and semantic Stroop interference (e.g., *SKY*_{yellow} – *DEAL*/****_{yellow}) computed on z-scores were also significant. Indeed, z-scored RTs for incongruent and semantically-associated stimuli were both significantly ($p < .001$) longer than those for neutral stimuli (with $BF_{10} = 2.399e+9$ and $BF_{10} = 501.448$, respectively). This latter difference also means that the semantic conflict was significant overall. Since z-scored RTs for incongruent stimuli were significantly longer than those for semantically-associated stimuli ($p < .001; BF_{10} = 58824.71$), response conflict (i.e., conflict that was not specifically analyzed and reported in

¹The reported BF values for interaction alone were obtained by dividing the BF value of the model containing the interaction and the two main effects by the BF value of the model containing the main effects only.

Augustinova et al., 2018) was also found to be significant.

The decomposition of the crucial Stimulus-Type \times Age-Group (see Table 3S) interaction revealed that the simple main effect of Stimulus-type was significant for the younger adults group [$F(2,55) = 36.73$; $p < .001$, $\eta_p^2 = 0.572$, $BF_{10} = 5.893e+15$]², with significant standard Stroop interference ($p < .001$, $BF_{10} = 4.907e+8$) resulting from a significant response conflict ($p < .001$, $BF_{10} = 193266.07$) and a marginally significant semantic conflict/interference ($p = .078$), and with Bayesian evidence showing anecdotal evidence in favor of this latter effect ($BF_{10} = 1.633/BF_{01} = 0.613$).

Table 3S. *Color-Identification Performance on z-scored RTs (Means, Standard Error, 95%-Confidence Intervals) observed as a function of stimulus-type, interference-type or conflict-type, and age-group in Augustinova et al. (2018)*

	Younger Adults		Older Adults		Age-effect
	M (SE)	CI	M (SE)	CI	
Standard color incongruent ($BLUE_{red}$)	.225 (.030)	[.165, .285]	.365 (.030)	[.306, .425]	+.140 ^{*/**}
Semantically-associated incongruent color (SKY_{red})	-.143 (.030)	[-.204, -.082]	-.142 (.030)	[-.203, -.081]	+.001 ^{ns/ns}
Color neutral (PUT_{red})	-.225 (.032)	[-.289, -.160]	-.353 (.032)	[-.418, -.289]	+.129 ^{*/ns}
	M (SD)	CI	M (SD)	CI	
Standard Stroop interference ($BLUE_{red} - PUT_{red}$)	.450 ^{**/††} (.053)	[.344, .478]	.719 ^{**/††} (.053)	[.613, .825]	+.269 ^{**/†}
Semantic interference/conflict Stroop ($SKY_{red} - PUT_{red}$)	.082 ^{°/ns*} (.045)	[-.009, .173]	.212 ^{**/††} (.045)	[.121, .303]	+.129 ^{*/ns*}
Response conflict ($BLUE_{red} - SKY_{red}$)	.368 ^{**/††} (.055)	[.258, .478]	.507 ^{**/††} (.055)	[.397, .617]	+.140 ^{°/ns*}

Note. Inferential approach: ^{ns}non-significant; [°]marginally significant at $p < .10$; ^{*}significant at $p < .05$; ^{**}significant at $p < .01$; Bayesian approach (presented after the slash): ^{ns}no evidence of an effect, with BF_{01} value between 1-3; ^{ns*}anecdotal evidence in favor of H1, with BF_{10} value between 1-3; ^{*}moderate evidence in favor of H1, with BF_{10} value between 3-10; ^{**}strong evidence in favor of H1, with BF_{10} value between 10-30; [†]very strong evidence in favor of H1, with BF_{10} value between 30-100; ^{††}extreme evidence in favor of H1, with BF_{10} value < 100 .

² The BF_s of the simple main effects were obtained by conducting a simple Bayesian ANOVA in each group separately. The BF_s of the further contrasts were obtained by conducting Bayesian Paired Samples-Tests with the specified alternative hypothesis that the value 1 is higher than the value 2 (e.g., z-scored RTs for standard color-incongruent trials > color-neutral trial).

This decomposition further revealed that the simple main effect of Stimulus-Type was significant in the older adults group [$F(2,112) = 20.53$; $p < .001$, $\eta_p^2 = 0.268$, $BF_{10} = 1.392e+23$], with significant Stroop interference ($p < .001$, $BF_{10} = 6.987e+9$) being observed and combining both a significant response ($p < .001$, $BF_{10} = 2.143e+7$) and semantic conflict/interference ($p = .001$, $BF_{10} = 645.67$).

For the purposes of comparison with conclusions reported in Augustinova et al. (2018) based on analyses of proportional transformations of standard vs. semantic Stroop interference, magnitudes of both types of interference computed from z-scored RTs were further analyzed in a 2 (Interference-Type: standard vs. semantic Stroop Interference) \times 2 (Coloring: ALCC vs. SLCC) \times 2 (Age-Group: Younger vs. Older adults) ANOVA (see Table 3S). This analysis revealed a significant main effect of Interference-Type [$F(1,56) = 126.38$; $p < .001$, $\eta_p^2 = 0.693$] and of Coloring [$F(1,56) = 23.61$; $p < .001$, $\eta_p^2 = 0.297$], both supported by extreme Bayesian evidence ($BF_{10} = 3.073e+15$ and $BF_{10} = 1506.80$, respectively). It also revealed a significant main effect of Age-Group [$F(1,56) = 11.92$; $p = .001$, $\eta_p^2 = 0.175$] that was supported by moderate Bayesian evidence ($BF_{10} = 8.151$). The Interference-Type \times Coloring interaction was significant [$F(1,56) = 18.26$; $p < .001$, $\eta_p^2 = 0.246$, $BF_{10} = 35.35$]. Neither the Age-Group \times Coloring interaction [$F(1,56) = .850$; $p < .361$, $\eta_p^2 = 0.015$, $BF_{10} = 0.294/BF_{01} = 3.41$], nor the 3-way Interference-Type \times Coloring \times Age-group interaction proved to be significant [$F(1,56) = .850$; $p = .361$, $\eta_p^2 = 0.015$, $BF_{10} = 2.37683E-22/BF_{01} = 2.40$]. In sum, these different results mirror those reported in Augustinova et al. (2018) with proportional transformations of standard vs. semantic Stroop interference.

Indeed, as with proportional transformations reported in Augustinova et al. (2018), the crucial Interference-Type \times Age-Group interaction was marginally significant [$F(1,56) = 3.182$; $p = .080$, $\eta_p^2 = 0.054$]. However, Bayesian evidence ($BF_{10} = 0.566/BF_{01} = 1.766$) actually revealed anecdotal evidence in favor of the null hypothesis. This absence of

interaction suggests that when magnitudes of interference are computed from z-scored RTs, healthy aging influences both standard and semantic Stroop interference in tandem. To examine this possibility further, we analyzed the simple main effect of Age-Group at each level of interference (see Table 3S) in a simple 2 (Interference-Type: standard vs. semantic Stroop interference) \times 2 (Age-Group: younger vs. older adults) ANOVA. This effect was significant for both standard [$F(1,56) = 12.93$; $p < .001$, $\eta_p^2 = 0.188$; $BF_{10} = 85.132$) and semantic [$F(1,56) = 4.05$; $p = .049$, $\eta_p^2 = 0.067$] Stroop interference, although Bayesian evidence in favor of an effect of aging on semantic Stroop interference was only anecdotal ($BF_{10} = 2.715/BF_{01} = 0.368$). Taken together, these analyses indicate that the magnitude of semantic Stroop interference computed from z-scores is greater in older adults than in their younger counterparts. This also suggests that when computed from z-scores, both semantic and response conflict are actually affected similarly by healthy aging.

To assess this latter idea directly, magnitudes of conflicts (see Table 3S) computed from z-scored RTs were further analyzed in a 2 (Conflict type: semantic vs. response) \times 2 (Coloring: ALCC vs. SLCC) \times 2 (Age-Group: older vs. younger) ANOVA. This revealed significant main effects of Conflict-Type [$F(1,56) = 22.91$; $p < .001$, $\eta_p^2 = 0.290$, $BF_{10} = 2.388e+6$], Coloring [$F(1,56) = 36.65$; $p < .001$, $\eta_p^2 = 0.396$, $BF_{10} = 91.574$], and of Age-Group [$F(1,56) = 12.95$; $p < .001$, $\eta_p^2 = 0.188$, $BF_{10} = 1.973$]. Both the Conflict-Type \times Age-Group [$F(1,56) = 0.005$; $p = .943$, $\eta_p^2 = 0.000$, $BF_{10} = 0.199/BF_{01} = 5.011$] and Coloring \times Age-Group [$F(1,56) = 1.92$; $p = .171$, $\eta_p^2 = 0.033$, $BF_{10} = 0.270/BF_{01} = 3.666$] interactions remained non-significant (with moderate Bayesian evidence against both interactions). Finally, the Conflict type \times Coloring \times Age-Group interaction also remained non-significant [$F(1,56) = 0.579$; $p = .450$, $\eta_p^2 = 0.010$, $BF_{10} = 0.354/BF_{01} = 2.822$] (with anecdotal Bayesian evidence against the 3-way interaction).

Both the moderate Bayesian evidence against the Conflict-Type \times Age-Group

interaction, and the main effect of Age-Group are consistent with the idea that healthy aging affects semantic and response conflicts in tandem. In line with the anecdotal evidence for the main effect of Age-Group, Bayesian Paired Samples T-Tests conducted for exploratory purposes revealed anecdotal evidence for larger magnitudes of semantic ($BF_{10} = 2.715/BF_{01} = 0.368$) and response ($BF_{10} = 1.922/BF_{01} = 0.520$) conflict in older adults than in their younger counterparts.

These additional analyses using z-scored RTs are at odds with the results originally reported by Augustinova and colleagues on proportional transformations. Indeed, they showed that both standard and semantic Stroop interference are affected by healthy aging, with the magnitudes of both being greater in older adults. These effects are due to the fact that both semantic and response conflicts (calculated from z-scores) tend to be experienced more intensely by healthy older adults. It should be noted that these latter results are also at odds with those reported in the present manuscript showing that healthy aging amplifies the magnitude of semantic conflict but leaves the magnitude of response conflict unaffected.

To address these discrepancies directly, we first computed Winer's z-tests combining F values for independent interactions (Winer, 1971, pp. 49–50, see e.g., Ferrand et al., (2020); Tse & Neely, 2007 for applications). First F values for 3 (Stimulus-Type: standard incongruent/different response vs. semantic/same response vs. neutral) \times 2 (Age-Group: adults vs. older adults) ANOVA from the present study³ and from the additional analysis of data reported above were combined. Since, taken separately, both of these interferences were significant, it was not surprising that the z-score was also significant ($z = 10.127, p < .001$). This was also the case ($z = 2.227, p = .006$) for the Interference-Type \times Age-Group interaction, which was marginally significant in Augustinova et al. (2018, see above) and non-

³ The F from the present study [$F(2,178) = 7.541; p < .001, \eta_p^2 = 0.078$] used for Winer's z-test did not include color-congruent trials.

significant in the present study⁴. The final Winer's z-test combined the independent F values for the Conflict-Type \times Age-Group interaction, which was non-significant in Augustinova et al. (2018, see above) and marginally significant in the present study⁵. This also turned out to be significant ($z=2.360, p=.004$). In the light of these significant values, we conducted cross-study analyses in order to examine the effects of healthy aging across the two independent studies using a more conservative Bayesian statistical approach.

Cross-study analysis

To this end, the two data sets were averaged and analyzed in a 3 (Stimulus-Type: standard incongruent/different response vs. semantic/same response vs. neutral) \times 2 (Age-Group: younger vs. older adults) omnibus ANOVA (see Table 4S). This analysis revealed a main effect of Stimulus-Type [$F(2,294) = 229.86; p < .001, \eta_p^2 = 0.610; BF_{10} = 3.153e+77$]. The main effect of Age-Group was also significant [$F(1,147) = 7.311; p = .008, \eta_p^2 = 0.047$], but Bayesian evidence actually showed moderate evidence in favor of the null hypothesis ($BF_{10} = 0.151/BF_{01} = 6.638$). In line with the aforementioned Winer's z-score, the Stimulus-Type \times Age-Group interaction was significant [$F(2,294) = 12.79; p < .001, \eta_p^2 = 0.080; BF_{10} = 318934.59$]. Its decomposition showed that in younger adults, the significant Stroop interference ($p < .001; BF_{10} = 1.939e+18$ combined both a significant response conflict ($p < .001; BF_{10} = 7.098e+11$) and a significant semantic conflict/interference ($p = .003; BF_{10} = 12.730$). The same pattern was observed in the older adults' group, with the significant Stroop interference ($p < .001; BF_{10} = 1.814e+20$) including both a significant response conflict ($p < .001; BF_{10} = 2.089e+11$) and a significant semantic conflict ($p = .001; BF_{10} = 2.223e+7$).

⁴ The F from the present study (not including color-congruent trials) is [$F(1,89) = 0.10; p = .921, \eta_p^2 = 0.000$].

⁵ The F from the present study (not including color-congruent trials) is [$F(1,89) = 3.337; p = .069, \eta_p^2 = 0.037$]

Given that both types of interference were significant in the younger and older adults, their magnitudes were analyzed directly in a 2 (Interference-Type: standard vs. semantic) \times 2 (Age-Group: younger vs. older adults) ANOVA (see Table 4S). This revealed significant main effects of Interference-Type [$F(1,147) = 182.73$; $p < .001$, $\eta_p^2 = 0.554$; $BF_{10} = 1.780e+24$], and Age-Group [$F(1,147) = 24.60$; $p < .001$, $\eta_p^2 = 0.143$; $BF_{10} = 3102.25$]. Although, and to some extent in line with the aforementioned significant Winer's z-Test, the Interference-Type \times Age-Group interaction was marginally significant [$F(1,147) = 2.93$; $p = .089$, $\eta_p^2 = 0.020$], Bayesian evidence actually revealed anecdotal evidence against this interaction ($BF_{10} = 0.622/BF_{01} = 1.605$). This absence of interaction suggests that across the two independent experiments, standard and semantic Stroop interference computed from z-scores are affected similarly by healthy aging (see Table 4S). This is likely due to the fact that both semantic and response conflicts are experienced to a greater extent by healthy older participants (compared to their younger counterparts).

To assess this idea directly, the magnitudes of semantic and conflicts were analyzed in a 2 (Conflict-Type: semantic vs. response) \times 2 (Age-Group: younger vs. older adults) ANOVA. This revealed a main effect of Conflict-Type [$F(1,147) = 19.003$; $p < .001$, $\eta_p^2 = 0.114$; $BF_{10} = 53777.26$] and of Age-Group [$F(1,147) = 22.76$; $p < .001$, $\eta_p^2 = 0.134$; $BF_{10} = 19.689$]. Unlike Winer's z-score, the Conflict-Type \times Age-Group interaction was not significant [$F(1,147) = .742$; $p = .390$, $\eta_p^2 = 0.005$], with Bayesian evidence providing moderate evidence against the interaction ($BF_{10} = 0.277/BF_{01} = 3.603$). This absence of interaction is indeed consistent with the idea that both semantic and response conflicts are experienced to a greater extent by healthy older participants (compared to their younger counterparts). It is, however, noteworthy that as in the data from the 2-to-1 Stroop paradigm reported in the main manuscript, Bayesian Independent Samples T-Test conducted for exploratory purposes (see Table 4S) actually revealed anecdotal evidence against the age-effect ($BF_{10} = 0.668/BF_{01} = 1.497$) on response

conflict, whereas it provided very strong evidence ($BF_{10} = 47.539$) in favor of this effect on semantic conflict.

Table 4S. *Color-Identification Performance on z-scored RTs (Means, Standard Error, 95% Confidence Intervals) observed as a function of stimulus-type, interference-type or conflict-type, and age-group in the two independent experiments*

		Younger Adults		Older Adults		Age-effect
		M (SE)	CI	M (SE)	CI	
Standard color incongruent / Different response		.233 (.018)	[.197, .269]	.352 (.020)	[.313, .391]	+.119 ^{**/††}
Semantically-associated color incongruent / Same response		-.036 (.018)	[-.072, .000]	.006 (.019)	[-.033, .044]	+.041 ^{ns/ns}
Color neutral		-.117 (.017)	[-.151, -.082]	-.215 (.018)	[-.252, .178]	- +.098 ^{**/ns}
		M (SD)	CI	M (SD)	CI	
Standard Stroop		.350 ^{**/††} (.031)	[.288, .411]	.567 ^{**/†} (.033)	[.502, .633]	+.218 ^{**/††}
Semantic Stroop		.081 ^{*/†} (.027)	[.028, .134]	.221 ^{**/†} (.029)	[.164, .277]	+.140 ^{**/†}
Response conflict		.269 ^{**//} (.031)	[.207, .330]	.346 ^{**/†} (.033)	[.281, .412]	+.078 ^{°/ns}

Note. Inferential approach: °non-significant; °marginally significant at $p < .10$; *significant at $p < .05$; **significant at $p < .01$; Bayesian approach (presented after the slash): °no evidence of an effect, with BF_{01} value between 1-3; ns°anecdotal evidence in favor of H_1 , with BF_{10} value between 1-3; *moderate evidence in favor of H_1 , with BF_{10} value between 3-10; **strong evidence in favor of H_1 , with BF_{10} value between 10-30; †very strong evidence in favor of H_1 , with BF_{10} value between 30-100; ††extreme evidence in favor of H_1 , with BF_{10} value < 100 .

7. References for Supplementary Materials

Ferrand, L., Ducrot, S., Chausse, P., Maionchi-Pino, N., O'Connor, R. J., Parris, B. A., Perret, P., Riggs, K. J., &

Augustinova, M. (2020). Stroop interference is a composite phenomenon: Evidence from distinct developmental trajectories of its components. *Developmental Science*, 23(2).

<https://doi.org/10.1111/desc.12899>

Tse, C.-H., & Neely, J. H. (2007). Semantic and repetition priming effects for Deese/Roediger-McDermott (DRM) critical items and associates produced by DRM and unrelated study lists. *Memory & Cognition*, 35, 1047–1066. <https://doi.org/10.3758/BF03193477>

**CHAPTER THREE: FURTHER INVESTIGATION OF DISTINCT COMPONENTS OF STROOP
INTERFERENCE AND OF THEIR REDUCTION BY SHORT RESPONSE-STIMULUS INTERVALS: NEW
EVIDENCE FROM THE TWO-TO-ONE STROOP PARADIGM.**

1. Introduction

When performing the Stroop task (Stroop, 1935), individuals are instructed to identify – as quickly and accurately as possible – the font color of written stimuli while ignoring their actual meaning. These instructions are particularly challenging to follow when the to-be-ignored word dimension of written stimuli denotes a different color – as is the case in the emblematic color-incongruent trials (e.g., the word “BLUE” displayed in green; hereafter *BLUE_{green}*). Indeed, color-identification times are consistently longer (and more error-prone) for this type of trials compared to those observed for different baseline trials.

Given the number of studies that have employed the Stroop task since its inception in 1935 (Stroop, 1935; see also e.g., MacLeod, 1991, 2005, Parris et al., 2021 for reviews), it is no surprise that a variety of different baselines have been employed. Commonly, the aforementioned difference in color-naming times (and error rates) between color-incongruent and color-congruent trials (e.g., the word “BLUE” displayed in blue, hereafter *BLUE_{blue}*) is referred to as *the overall Stroop effect* or *congruency effect*; and that between color-incongruent and color-neutral trials (e.g., *DEAL_{green}*) as *the Stroop interference effect* (e.g., MacLeod, 1991, Parris et al., 2021). Although these two Stroop phenomena are different in the quantity of interference they actually involve, both are used as indicators that despite being explicitly instructed not to, participants still attend to the irrelevant word-dimension of Stroop items (i.e., process its meaning). Indeed, incidental semantic processing of the irrelevant word-dimension of color-incongruent Stroop items, specifically, slows in turn the

participants' performance because a response induced by this latter dimension (e.g., say blue/press blue for *BLUE*_{yellow}) interferes with that induced by the relevant color dimension (e.g., say yellow/press yellow for *BLUE*_{yellow}, see also e.g. Cohen et al., 1990; Glaser & Glaser, 1989; Logan, 1980; Melara & Algom, 2003; Phaf et al., 1990, Roelofs, 2003). Therefore, numerous studies – including those of Stroop himself – have tried to identify interventions that improve attentional selectivity in the Stroop task (see e.g., Augustinova & Ferrand, 2014; MacLeod, 1991; Parris et al., 2021 for reviews).

One intervention that seems to be particularly effective is to shorten the time that elapses between the participant's response and the presentation of a new stimulus on a computer screen (i.e., response stimulus interval, hereafter RSI; De Jong et al., 1999). Indeed, past researches have repeatedly demonstrated that this experimental manipulation – causing participants to perform the Stroop task at a much faster rate than is ordinarily the case – considerably reduces the magnitude of resulting Stroop interference (De Jong et al., 1999; see also e.g., Augustinova et al., 2018; Jackson & Balota, 2013; Parris, 2014). Despite this, the processes underlying this reduction are still poorly understood. Consequently, this paper attempted to shed more light on the level of processing at which the benefit of a short (as opposed to long) RSI takes its effect.

At what level of processing the benefit of a short (as opposed to long) RSI takes its effect?

Even though Stroop interference is often conceptualized as a measure of people's attentional selectivity (i.e., of their ability to focus on one thing while ignoring other things, see e.g., Ruthruff & Lien, 2016), the extent to which the magnitude of this inference specifically results from their focus on the *relevant color-dimension* is often disregarded. Indeed, Stroop interference is usually viewed as a consequence of people's failure to inhibit

the irrelevant word-dimension (e.g., “BLUE” for *BLUE*_{yellow}) specifically such that the Stroop task is seen as a prototypical inhibition task (see e.g., Miyake et al., 2000).

De Jong and colleagues (1999) reasoned that the failure to inhibit the irrelevant word dimension does not necessarily reflect genuine limitations in people’s inhibitory capabilities but rather certain limitations in the ability to deploy these capabilities *consistently* during the executive control of task performance. More precisely, and in line with the idea of the importance of attentional selection as emphasized above, these authors argued that the ability to inhibit the irrelevant word dimension is fully available only when the individual’s attention is intently focused on the relevant color-dimension. As mentioned above, in order to enhance this latter focus on the relevant color-dimension, De Jong and colleagues (1999) suggested to shorten a rather conventional RSI of 2000 ms to 200 ms. As a consequence, in a two-response spatial Stroop task (i.e., responding to the words “ABOVE” and “BELOW” presented either above or below the fixation point), a short RSI substantially reduced the ensuing Stroop-like interference. More specifically, the magnitude observed at a RSI of 2000 ms dropped from 47 ms to a non-significant 11 ms when a short RSI of 200 ms was used. This statistical elimination of the Stroop-like interference was therefore in line with De Jong and colleagues’ (1999) original reasoning that performing the Stroop-like task at a faster pace increases participants’ ability to focus on the relevant (i.e., color) dimension of Stroop words and thus to inhibit the irrelevant (i.e., word) dimension of these words more consistently (see also Jackson & Balota, 2013 for the same reasoning).

Although neither De Jong and colleagues (1999), nor Jackson and Balota (2013) had actually addressed this issue explicitly, according to so-called single-stage response competition models of the Stroop interference effect as that of Roelofs (2003, see also e.g. Cohen et al., 1990; Glaser & Glaser, 1989; Logan, 1980; Melara & Algom, 2003; Phaf et al., 1990), this more efficient suppressing the word-dimension of color-incongruent Stroop words

in a short RSI of 200 ms should lead to a smaller competition at the level of response output (i.e., to reduce a *response conflict* depicted above). Indeed, according to this dominant view of Stroop interference as a unitary phenomenon (i.e., phenomenon generated by a single type of conflict), a short RSI of 200 ms (exactly as other interventions) is expected to reduce a competition between two alternative responses indicated by the two dimensions of the Stroop stimulus.

In contrast, Zhang and colleagues' two-conflicts model (1999; see also Zhang & Kornblum, 1998) anticipates the Stroop interference effect to result from two distinct conflicts. From this point of view, in addition to the aforementioned conflict occurring at the level of response output (i.e., response conflict), another conflict is thought to occur at the level of semantics (hence *semantic conflict*). Therefore, if De Jong and colleagues (1999, see also Jackson & Balota, 2013) are accurate in their reasoning that the ability to inhibit the irrelevant word dimension of Stroop words is being more efficiently deployed when individuals' attention is fully focused on the color dimension (as it is with short RSI), then – in this composite view of the Stroop interference effect – both semantic and response conflicts in tandem or semantic conflict alone should be reduced in short RSI condition.

Both of these predictions contrast with the reasoning outlined in the study of Parris (2014a). It reported non-significant Stroop interference but rather large facilitation in the short RSI condition (as opposed to non-significant Stroop facilitation but rather large interference in the long RSI). This joint pattern of lowered interference and boosted facilitation was taken as an indication of reduced *task conflict* (Bench et al., 1993b; Goldfarb & Henik, 2007b; Hershman & Henik, 2020b)¹ in the short RSI condition.

¹ Recall that task conflict is a more general conflict that – for all readable Stroop items including color-neutral ones – is thought to derive from the simultaneous preparation of two task sets: word-reading vs. color-naming.

Finally, to complicate this issue further, the only published study that tackled the question of different types of conflict directly following a manipulation of RSI, Augustinova, et al. (2018) found that shortening the RSI only reduces response conflict but leaves both task and semantic conflict unaffected. Notably however, this study measured semantic conflict via semantic-associative interference (e.g., $SKY_{\text{yellow}} - DEAL_{\text{yellow}}$). It therefore remains prone to criticisms – already outlined by Burca et al. (2021, see also Burca et al., 2022 and chapters 1 & 2) – that stem from unitary or single-stage response competition models.

In line with the aforementioned idea that a unique (i.e., response) conflict is the driving force behind the Stroop interference effect, single-stage response competition models argue that semantic associates (i.e., SKY_{yellow}) elicit incorrect response activity (e.g., say “blue”/press blue for SKY_{yellow}) indirectly – through their association with the response-set colors (blue in this case). Therefore, it remains possible that semantic-associative interference (e.g., $SKY_{\text{yellow}} - DEAL_{\text{yellow}}$) measured in Augustinova and colleagues’ study results entirely from response conflict (e.g., Roelofs, 2003). As a consequence, Augustinova et al.’s (2018) study failed to demonstrate satisfactorily that the type of conflict that is spared by short RSI is *semantic* (i.e., due specifically to a slowdown that occurs whenever two distinct yet closely related semantic representations are simultaneously activated in an amodal semantic network; see e.g., Seymour, 1977, for discussion). In totality, the study of Augustinova et al. (2018) raises more issues than it actually solves. Indeed, it raises a more general and still-open issue of whether there is a semantic conflict in the Stroop task. Furthermore, the measure of task conflict it used is far from optimal (see Parris et al., 2021 for discussion) and not directly comparable with that used by Parris (2014). Therefore, the goal of the paper was to shed more light on both of these issues.

Is there any semantic conflict in the Stroop task?

To address this question – raised again² by Zhang and colleagues’ two-conflicts model (1999; see also Zhang & Kornblum, 1998) –, De Houwer (2003) ingeniously construed what it is now called the *two-to-one Stroop paradigm*. In this paradigm, all the distractors are part of the response set (e.g., BLUE, RED, GREEN, YELLOW). Also, and importantly, responses for paired target colors are mapped to only one response-key (e.g., ‘f’ for blue and red and ‘j’ for green and yellow). As a result of this response-mapping, standard incongruent Stroop trials like *BLUE*_{yellow} provide evidence toward two different responses (they are therefore termed *different-response trials*). Indeed, relevant color-dimension (YELLOW) prompts the correct response activity toward ‘j’ key, whilst the irrelevant word-dimension (BLUE) prompts the incorrect response activity toward ‘f’ key. There is no such (response) conflict on trials like *BLUE*_{red} since both dimensions of the Stroop stimulus provide evidence toward the same response. Consequently, according to unitary or single-stage response conflict models mentioned above (e.g., Roelofs, 2003, see also e.g. Cohen et al., 1990; Glaser & Glaser, 1989; Logan, 1980; Melara & Algom, 2003; Phaf et al., 1990), these latter trials should not produce any interference, as they do not involve any response conflict.

However, in De Houwer (2003)’s study not only were both different- and same-response trials responded to slower than baseline identity or color-congruent trials (*BLUE*_{blue}), but also, different-response trials were responded to slower than same-response trials (e.g., RTs to *BLUE*_{green} > RTs to *BLUE*_{red}). Therefore, in line with the two-conflicts model of Zhang and colleagues mentioned above (H. Zhang et al., 1999; H. Zhang & Kornblum, 1998a), these results allowed De Houwer (2003) to conclude that different-response trials generate both

² Recall that both MacLeod (1991) and Logan and Zbrodoff (1998) argued that the evidence in favor of early selection accounts anticipating this type of conflict was not sufficient. Therefore, until up to Zhang and Kornblum (1998), this issue was considered as solved such that there is no such contribution of the Stroop interference effect.

semantic and response conflicts, whereas interference produced by the same-response trials results solely from semantic conflict.

These findings have largely been replicated since (e.g., A. Chen et al., 2011, 2013; Z. Chen et al., 2013b; Hershman & Henik, 2019a, 2020b; Jiang et al., 2015; Šaban & Schmidt, 2021b; Schmidt et al., 2018; Schmidt & Cheesman, 2005a; Shichel & Tzelgov, 2018b; Van Veen & Carter, 2005). Despite these convergent empirical results, the contribution of semantic conflict to the overall color-word Stroop effect still remains an open empirical issue (see Parris et al., 2021 for ample discussion).

Indeed, all studies employing this measure of semantic conflict – including De Houwer (2003) – have used color-congruent (or identity) trials as the baseline against which interference induced by same-response trials is measured. Therefore, unitary or single-stage response conflict models of the Stroop effect (Cohen et al., 1990a; Glaser & Glaser, 1989; Logan, 1980; Melara & Algom, 2003; Phaf et al., 1990; Roelofs, 2003a) can still easily account for this difference as resulting from facilitation on color-congruent trials (as opposed to interference on same-response trials). In line with this possible interpretation, Hasshim and Parris (2014, 2015) consistently reported significantly longer RTs for same-response trials than for color-congruent trials, but no difference between same-response trials and color-neutral trials (i.e., trials that are free of facilitation, see e.g., Brown, 2011; MacLeod, 1991a for discussion). The absence of semantic conflict was supported further by Bayesian evidence for the null-hypothesis and by the unchanged magnitude of associated pre-response pupillometric measures of effort (i.e., a reliable and more sensitive measure of the potential differences between conditions; Hasshim & Parris, 2015; Parris et al., 2021).

While in terms of reaction times, the study of Hershman and Henik (2020b) reported no evidence for a difference between same-response and *non-word* color-neutral trials (e.g., XXXX_{blue}, hence replicating the pattern of results reported by Hasshim and Parris, including

conclusions based on Bayes factors), they did report a difference between these two conditions in pupillometric measures (see also Hershman & Henik, 2020b). Moreover, in perfect contrast to Hasshim and Parris (2014, 2015), Burca et al. (2021, see Chapter 1) reported a significant difference between same-response and color-neutral word trials (e.g., *DEAL*_{blue}). They, therefore, argued that the difference between same-response and color-congruent trials (i.e., when no color-neutral baseline is included) simply confounds the *semantic (stimulus) conflict* produced by same-response trials and *facilitation* produced by color-congruent trials (MacLeod, 1991a). However, since their study did not include color-congruent trials, their reasoning remained tentative until the actual inclusion of this type of trial in a subsequent study. This latter addition made it possible for Burca et al. (2022, see Chapter 2) to isolate a robust contribution of semantic conflict to the overall Stroop effect that was clearly independent of both response conflict and Stroop facilitation. Indeed, overall, the Bayesian evidence in favor of semantic conflict was extreme (so was the Bayesian evidence in favor of response conflict, while that in favor of Stroop facilitation was moderate). Therefore, beyond confirming the composite nature of the overall Stroop effect (i.e., as comprising both interference and facilitation phenomena; see MacLeod, 1991a), this latter study provided the first unambiguous empirical demonstration that Stroop *interference* itself is a composite phenomenon. Indeed, as originally claimed by Zhang and colleagues' multi-stage model of Stroop interference (H. Zhang et al., 1999; H. Zhang & Kornblum, 1998a), it results from at least two distinct types of conflict (semantic vs. response conflict). Nevertheless, this latter demonstration comes from a single experiment – as opposed to several experiments carried out by Hasshim and Parris (2014, 2015). Therefore, the actual contribution of semantic conflict in the two-to-one Stroop paradigm, and more generally the composite (as opposed to unitary) nature of color-word Stroop interference – suggesting that

selection in the Stroop task can occur independently of response competition – remains an open empirical issue.

The Present study

In line with the idea of robust and reproducible science, the first goal of the present study was to replicate the contribution of semantic conflict in the two-to-one Stroop paradigm administered with both color-congruent and color-neutral baselines (Hasshim & Parris, 2014, Exp. 2; Burca et al., 2022). As in all studies using the original two-to-one Stroop paradigm (De Houwer, 2003), RTs for same – and different-response trials were both expected to be longer than those observed for color-congruent trials, and RTs for different-response trials were expected to be longer than those for same-response trials. If, in agreement with unitary or single-stage response competition models (Cohen et al., 1990a; Glaser & Glaser, 1989; Logan, 1980; Melara & Algom, 2003; Phaf et al., 1990; Roelofs, 2003a), there is no contribution of semantic conflict to the overall Stroop effect, the difference between same-response and color-congruent trials will be entirely driven by facilitation. In that case, the Bayes factor will show evidence against the difference in RTs between same-response and color-neutral trials – as previously reported by Hasshim and Parris (2014, 2015). In contrast, if in line with Zhang and colleagues (H. Zhang et al., 1999; H. Zhang & Kornblum, 1998a), the Stroop effect is indeed a composite phenomenon (and so is Stroop interference), then same-response trials will produce significant semantic conflict (i.e., the difference between same-response and color-neutral trials will be significant). Also, and importantly, its contribution will be independent of both response conflict (i.e., the significant difference between different and same-response trials) and facilitation (i.e., the significant difference between congruent and color-neutral trials) – as previously reported by Burca et al. (2022).

Since this latter prediction corresponds to the replication of our own findings, it was the one that the present study a priori favored.

Also, and importantly, the contribution of all these latter constituents (facilitation and interference – itself composed of both semantic and response conflicts) is the *sine qua non* condition for examining the level of processing at which the benefit of a short (as opposed to long) RSI takes its effect. This examination corresponded, indeed, to the second goal of the present study. If the ability to inhibit the irrelevant word dimension of Stroop words is being more efficiently deployed when individuals' attention is fully focused on the color dimension (as it is with short RSI, see e.g., De Jong et al., 1999, Jackson & Balota, 2013), then both semantic and response conflicts in tandem or semantic conflict alone should be reduced in short (as opposed to long) RSI condition. Conversely, if short (as opposed to long) RSI leaves the ability to inhibit the irrelevant word dimension of Stroop words unaffected and influences only the ability to inhibit the (irrelevant) response that this latter dimension prompts, then it should only reduce response conflict (Augustinova et al., 2018). Since the former (unlike the latter) pattern of results is likely to extend those of Burca and colleagues (2021) to the effect of RSI, it corresponded to our a priori prediction.

It should be noted that neither of these competing predictions are actually contradictory with the one of Parris (2014). Recall that Parris (2014) attributed reduced Stroop interference in short (as opposed to long) RSI to lowered task conflict. Yet, the reduction of this latter conflict can occur with that of other types of conflict (see e.g., Shichel & Tzelgov, 2018). In the latter study both task conflict and semantic conflict (although derived as a difference between same-response and color-congruent items) were reduced by the variations in the proportion of non color-neutral trials but the response conflict was not. If task conflict is indeed reduced along with semantic and response conflict (or semantic conflict alone, see above) in the present study, then reduced interference in short RSI should occur in

concert with boosted facilitation – the latter being a *prima facie* evidence that the short RSI (also) influence the task conflict in the Stroop task. To examine these different hypotheses linked to the benefit of short (as opposed to long) RSI, slow responding was induced as in the initial study of De Jong and colleagues (1999) by employing an RSI of 2000 ms (henceforth long RSI) for half of the participants, and faster responding was induced by employing an RSI of 200 ms (henceforth short RSI) for the other half of the participants.

2. Method

Participants and Design

Fifty-two younger adults (32 females and 20 males, $M_{\text{age}} = 24.75$, $SD_{\text{age}} = 3.53$) reporting normal or corrected to-normal color vision volunteered to take part in this experiment. Its data collection was function of 4 (Stimulus-Type: different-response vs. same-response vs. neutral vs. congruent) \times 2 (RSI: short vs. long) experimental design with the RSI being a between subjects' factor.

Procedure, apparatus, stimuli and design

Once they have signed an informed consent, all the participants underwent a short interview confirming that six months prior to inclusion in the study, none of them suffered from psychiatric and/or neurological disorders. It further confirmed that 24 hours prior to inclusion in the study, none of them consumed substances known to influence the response. Finally, this interview further confirmed that they indeed presented normal color vision. They then completed a computerized version of the Stroop Task presented via Eprime 2.0 software (Schneider et al., 2002). To this end, the participants were seated at a distance of 70cm in front of the 13'' portable computer, and instructed to identify the color of the stimulus presented on the screen, as quickly and accurately as possible by pressing the appropriate

color-button, while ignoring everything else on the display. They were specifically instructed to concentrate on the fixation cross (“+”) that appeared in the center of the screen at the beginning of each trial, with a duration of 200msec in the short RSI condition vs. a duration of 2000msec in the long RSI condition. The participants were assigned randomly to the short or long RSI condition such that 29 participants completed the short RSI condition, and 23 completed the long RSI condition³. The stimulus remained on the screen until the participants gave an answer or until 3500msec had elapsed. All stimuli were presented in lowercase Courier font, size 18, on a black background and subtended an average visual angle of 0.9° high × 3.0° wide. The used color-words were: *rouge* [red], *jaune* [yellow], *bleu* [blue], and *vert* [green], and the neutral non-color words were: *plomb* [lead], *liste* [list], *page* [page], *cave* [basement]. The color and non-color words were matched in length and frequency via *Lexique* 3.38 (New et al., 2004).

The 2-to-1 paradigm was used in this study, thus the participants responded manually by pressing the “k” key – which was labeled with the colors red and blue, or the “d” key – which was labeled with the colors green and yellow on an AZERTY keyboard. Before the experimental phase, all participants underwent a training session, which consisted of 60 practice trials, composed of colored asterisks. Both color-words and neutral words were presented in all colors mentioned above. The experimental block consisted of 120 stimuli, and contained 48 stimuli different-response stimuli, 24 same-response stimuli, 24 neutral stimuli, and 24 congruent stimuli. This block was repeated 9 times with a self-paced break in-between blocks. Thus, a total of 1080 stimuli was presented to each participant.

³ The uneven number of participants in each condition is a result of the impossibility of completing the recruitment, during the COVID-19 sanitary crisis. Note: however, that effect sizes are comparable with those of Jackson and Balota (2013) that examined the data of 26 participants in the short-RSI condition and of 25 participants in the long-RSI condition.

3. Results and discussion

The data from all 52 participants were used for statistical analysis. As in Balota et al. (2013), the first 100 trials were considered as practice, and thus were not taken into account during the statistical analysis (i.e., 980 trials per participant were analyzed). RTs greater than 3SDs above or below each participant's mean latency for each condition were excluded from the analysis (i.e., 1.99% of the total data).

Both RTs and errors were first analyzed in an omnibus 4 (Stimulus-Type: different-response vs. same-response vs. neutral vs. congruent) \times 2 (RSI: short vs. long) was conducted in both classical and Bayesian ANOVAs, with the RSI being a between-subjects factor. On RTs, it revealed a main effect of the Stimulus-Type [$F(3,150) = 20.57$; $p < .001$, $\eta_p^2 = 0.291$; $BF_{10} = 1.121e+7$] and that of RSI [$F(1,50) = 8.202$; $p = .006$, $\eta_p^2 = 0.141$; $BF_{10} = 4.015$]. As expected, these main effects were also included in the significant Stimulus-Type \times RSI interaction [$F(3,150) = 4.700$; $p = .004$, $\eta_p^2 = 0.086$] that was supported further by a moderate Bayesian evidence for the interaction only⁴.

This same omnibus ANOVA conducted on error rates revealed a main effect of Stimulus-Type [$F(3,150) = 177.93$; $p < .001$, $\eta_p^2 = 0.781$; $BF_{10} = 1.121e+7$]. It further revealed that the main effect of RSI remained non-significant [$F(1,50) = 1.14$; $p = .290$, $\eta_p^2 = 0.022$; $BF_{10} = 0.352$] and Bayes factor ($BF_{01} = 2.842$) provided at least anecdotal evidence against this effect. Although Stimulus-Type \times RSI interaction was marginally significant [$F(3,150) = 2.472$; $p = .064$, $\eta_p^2 = 0.047$] with standard frequentist statistics, again, Bayes factors $BF_{10} = 0.457/BF_{01} = 1.340$ provided at least evidence against this interaction.

⁴ Since the different models that a Bayesian ANOVA compares against a null model never include interaction alone, the BF values reported for all the interactions correspond to values obtained by dividing the BF value of the model containing the two main effects and their interaction by the BF value of the model with the two main effects only.

Is there a semantic conflict in the 2-to-1 paradigm?

To answer this question – constituting a necessary pre-requisite to examining the type of conflict that is influenced by RSI-manipulation in the Stroop task –, the significant main effect of the Stimulus-Type observed in the ANOVA conducted on both RTs and errors was further decomposed. On errors, this decomposition revealed, that more errors were committed for different-response (2.36%) as compared to same-response stimuli (1.87%; $p = .031$, $BF_{10} = 18.17^5$) but the error rate was comparable for different-response (2.36%) and congruent stimuli (1.93%; $p = .399$; $BF_{10} = 1.200$). The neutral stimuli, surprisingly, had the highest error rate (5.97%), which was statistically higher than that observed for all other stimuli (all $p_s < .001$, and all $BFs > 100$).

On RTs, this decomposition revealed that the different-response trials were longer than those observed to all the other stimuli (all $p's < .001$). This means that the participants displayed a significant overall Stroop effect (significant difference between different-response and congruent stimuli, $p < .001$; $BF_{10} = 94901.71$) of a magnitude of 36ms. It was further composed by a marginally significant ($p = .080$) facilitation effect (difference between neutral and congruent stimuli, 9ms). However, Bayes factor had actually provided moderate evidence in favor of this difference ($BF_{10} = 4.95$). Further decompositions revealed a significant ($p < .001$; $BF_{10} = 94901.71$) Stroop interference (difference between different-response and neutral stimuli, 27ms). While this latter Stroop phenomenon resulted from a significant ($p = .003$; $BF_{10} = 94.39$) contribution of response conflict (difference between different-response and same-response stimuli, 14 ms), that of semantic conflict (difference between same-response and neutral stimuli, 12ms) failed to reach significance ($p = .129$; $BF_{10} = 2.29$).

⁵ Reported BFs were computed by conducting Bayesian Paired Samples t-Test, where the alternative hypothesis specified that measure 1 was greater than measure 2 (i.e., RTs or errors to different-response > RTs or errors to congruent trials).

In light of significant Stimulus-Type \times RSI interaction, might possibly be due the fact that short RSI significantly reduced semantic conflict. To address this possibility further this latter interaction was decomposed by examining the simple main effect of the Stimulus-Type at level of RSI. It was significant at both long [$F(3,48) = 12.50$; $p < .001$, $\eta_p^2 = 0.439$; $BF_{10} = 14295.04^6$] and short RSI [$F(3,48) = 3.09$; $p = .036$, $\eta_p^2 = 0.162$; $BF_{10} = 36.33$]. Its further decomposition revealed that in long RSI condition, RTs to different-response trials were indeed significantly longer than those observed to and congruent stimuli ($p < .001$; $BF_{10} = 677.65$) such that the participants assigned to this condition displayed overall Stroop effect of 51ms (see Table 1 for Means and SDs). Again, while the facilitation effect (12ms) failed to reach significance in the standard frequentist ANOVA ($p = .171$), Bayes factor actually moderates evidence in favor of this effect ($BF_{10} = 3.403$). The aforementioned overall Stroop effect also resulted from a significant ($p < .001$; $BF_{10} = 41.93$) Stroop interference of 39ms. This latter Stroop phenomenon resulted in turn from a significant ($p = .032$; $BF_{10} = 8.596$) response conflict of 17 ms and a significant ($p = .037$; $BF_{10} = 3.959$) semantic conflict of 22 ms (see Table 1 for Means and SDs).

Therefore, these results replicate those of Burca and colleagues (2022) showing that when the Stroop task is administered under normal pace (as induced by RSI of 2000ms) both significant facilitation and interference effect contribute to the overall Stroop effect. Also, and importantly, exactly like in this study of Burca and colleagues both a semantic and response conflict contributed to the overall Stroop interference that participants displayed in the present study. The direct implication of this finding is that there is room for the magnitude of these different constituents of the Stroop effect in the long RSI to be significantly reduced by short RSI.

⁶ Because, there is no test for simple main effect in Bayesian ANOVA, corresponding BF values were obtained by conducting Bayesian Paired-Samples T-Tests at each level of RSI.

Table 1. Color-Identification Performance (Means, Standard Deviation, 95%-Confidence Intervals and Error Percentages) observed as a Function of Stimulus-Interference- or Conflict Type and Response-Stimulus-Interval

	Response-Stimulus-Interval					
	Long RSI			Short RSI		
	M (SD)	CI	%ER	M (SD)	CI	RSI effect (ER)
Different Response (DR) Color-Incongruent items	700 (204)	[620, 774]	2.65	558 (151)	[492, 624]	2.07 - .5 ^{ns/ns}
Same Response (SR) Color-Incongruent items	683 (196)	[612, 754]	2.02	546 (143)	[483, 609]	1.73 - .3 ^{ns/ns*}
Color-Neutral items	661 (159)	[560, 722]	5.97	543 (156)	[489, 598]	5.97 - .1 ^{ns/ns*}
Color-Congruent items	649 (166)	[588, 711]	2.47	537 (131)	[483, 593]	1.38 - .008 ^{ns/ns}
	M (SD)	CI		M (SD)	CI	
Overall Stroop Interference (DR – Congruent)	51 ^{*/††} (50)	[34, 67]		21 ^{*/††} (27)	[6, 35]	
Facilitation Effect (Neutral – Congruent)	12 ^{ns*} (25)	[1, 22]		6 ^{ns/ns} (25)	[-3, 15]	
Standard Stroop Interference (DR – Neutral)	39 ^{†/†} (8)	[23, 55]		14 ^{ns**} (43)	[-321, 29]	
Response Conflict (DR – SR)	17 ^{*/*} (26)	[5, 28]		12 ^{ns*} (27)	[2, 22]	
Semantic Stroop Interference (SR – Neutral)	22 ^{*/*} (8)	[7, 38]		3 ^{ns/ns} (56)	[-11, 16]	

Note. Inferential approach: ^{ns}non-significant; *significant at $p < .05$; **significant at $p < .01$; †significant at $p < .001$; Bayesian approach (presented after the slash): ^{ns}no evidence of an effect, with both BF_{10} and BF_{01} values between 0–1; ^{ns*}anecdotal evidence in favor of H1, but moderate evidence in favor of H0, with BF_{10} value between 0–1, and BF_{01} value between 3–10; ^{*}moderate evidence in favor of H1, with BF_{10} value between 3–10; ^{**}strong evidence in favor of H1, with BF_{10} value between 10–30; [†]very strong evidence in favor of H1, with BF_{10} between 30–100; ^{††}extreme evidence in favor of H1/of an effect, with $BF_{10} < 100$.

At what level of processing the benefit of a short (as opposed to long) RSI takes its effect?

To address this question, the benefit of a short (as opposed to long) RSI was first established by decomposing the simple main effect of RSI at each level of Stimulus-Type. As can be seen in Table 1, this effect was indeed significant at each level of Stimulus-Type such that short RSI of 200ms shortened responses to all four types of Stroop words (including color-neutral and color-congruent ones). Thus, in line with De Jong et al.'s reasoning, shortening RSI indeed induces faster pace that produces faster responding. This faster responding was in turn beneficial for participants' performance in the Stroop as further indicated by the decomposition of the simple main effect of Stimulus-Type at the level of short RSI [$F(3,48) = 3.09$; $p = .036$, $\eta_p^2 = 0.162$; $BF_{10} = 36.33$].

RTs to different-response trials were significantly longer than those observed to and congruent stimuli ($p = .041$; $BF_{10} = 173.23$) such that the participants assigned to this condition displayed the overall Stroop effect of 21ms (see Table 1 for Means and SDs). Contrary to Parris (2014), short RSI failed to boost facilitation. Indeed, this effect (6ms) failed to reach significance ($p = 1.000$; $BF_{10} = 0.779$) and $BF_{01} = 1.283$ provided at least anecdotal evidence against the contribution of facilitation to the overall Stroop effect. Although the contribution of Stroop interference (14ms) failed to reach significance in the standard frequentist ANOVA ($p = .330$), it was supported by strong Bayesian evidence ($BF_{10} = 26.23$) in favor of this effect. Although this latter Stroop phenomenon resulted from a significant contribution of response conflict (12 ms, $p = .152$; $BF_{10} = 4.624$ at least with Bayesian approach), that of semantic conflict (3ms) failed to reach significance with both approaches ($p = 1.000$; $BF_{10} = 0.292/BF_{01} = 0.25$). These different results suggest in sum that short (as opposed to long) RSI influences all types of Stroop phenomena.

To inspect this latter idea directly, a 2 (Type of Stroop phenomenon: Stroop facilitation vs. interference) \times 2 (RSI: long vs. short) was conducted. It revealed a significant main effect of the Type of Stroop phenomenon [$F(1,50) = 23.70$; $p < .001$, $\eta_p^2 = 0.322$; $BF_{10} = 1349.90$]. Unsurprisingly, the magnitude of facilitation (9ms) was significantly smaller ($p < .001$; $BF_{10} = 949.15$) than the one of (overall) Stroop interference (36ms) overall. The main effect of RSI – suggesting that the reduction of both Stroop phenomena – was again significant [$F(1,50) = 6.06$; $p = .017$, $\eta_p^2 = 0.108$], although supported only by anecdotal Bayesian evidence ($BF_{10} = 2.183$). Stroop phenomenon \times RSI interaction was also significant [$F(1,50) = 5.06$; $p = .029$, $\eta_p^2 = 0.092$] but again it was supported only by anecdotal Bayesian evidence ($BF_{10} = 2.220$) in favor of this interactive effect. Further decomposition of this interaction by testing the simple main effect of RSI at each Type of Stroop phenomenon revealed that this simple main effect was significant for the overall Stroop interference [$F(1,50) = 7.75$; $p = .008$, $\eta_p^2 = 0.134$; $BF_{10} = 11.927$]. Indeed, its magnitude was significantly reduced (by 30ms) in the short- as compared to long-RSI condition. The simple main effect of RSI remained non-significant for Stroop facilitation [$F(1,50) = .651$; $p = .424$, $\eta_p^2 = 0.013$; $BF_{10} = 0.560$] and Bayes factor provided at least anecdotal evidence against this effect for facilitation ($B_{01} = 1.785$). Therefore, comparable levels of Stroop facilitation in both conditions of RSI suggest that short RSI fails to influence (i.e., reduce) substantially the task conflict in the Stroop task.

To further inspect the way, the overall Stroop interference is actually reduced (i.e., by the reduction of semantic vs. response vs. both conflicts), an additional 2 (Conflict-Type: response vs. semantic conflict) \times 2 (RSI: short vs. long) ANOVA was conducted. The main effect of Conflict-type failed to reach significance [$F(1,50) = .068$; $p = .795$, $\eta_p^2 = 0.001$; $BF_{10} = 0.225$] and $BF_{01} = 4.452$ actually provided moderate evidence against this effect. In line with previous analyses, the main effect of RSI was significant [$F(1,50) = 5.06$; $p = .029$, $\eta_p^2 = 0.092$]. However, Bayesian evidence remained rather inconclusive ($BF_{10} = 0.832/BF_{01} =$

1.202). Conflict-Type \times RSI interaction also remained non-significant [$F(1,50) = 1.01$; $p = .313$, $\eta_p^2 = 0.020$; $BF_{10} = 0.530$] and $BF_{01} = 1.897$ provided at least anecdotal evidence against this interaction. These results suggest in sum that the aforementioned reduction of Stroop interference in short (as compared to long) RSI is due to the fact that both semantic and response conflicts are reduced in tandem by this experimental manipulation.

4. General Discussion and Conclusion

The first goal of the present study was to replicate the contribution of semantic conflict in the two-to-one Stroop paradigm administered with both color-congruent and color-neutral baselines (Hasshim & Parris, 2014, Exp. 2; Burca et al., 2022). In line with both of these studies, the overall Stroop (of congruency) effect resulted from the contribution of both facilitation and interference (see MacLeod, 1991; Parris et al., 2021) – suggesting again that this overall measure is not suitable for investigating interference (see e.g., Jackson & Balota, 2013) or conflict (e.g., Egner, 2010) only.

However, contrary to Burca and colleagues (2022) and in line with Hasshim and Parris (2014, see also Hasshim & Parris, 2015), overall, the contribution of semantic conflict (difference between same-response and neutral stimuli, 12ms) to Stroop interference failed to reach significance ($p = .129$) in the present study. But in contrast to Hasshim and Parris (2014), the associated Bayes factor failed to provide evidence against this contribution. Indeed, this contribution was actually supported by at least anecdotal Bayesian evidence while that of response conflict (difference between different-response and same-response stimuli, 14 ms) was supported by strong Bayesian evidence.

However, several studies recently suggested that the increased response speed is likely to reduce semantic conflict induced by semantic associates (Scaltritti et al., 2022; Sulpizio et al., 2022) or out-of-the set items (Hasshim et al., 2019). Although the extent to which these

items actually induce semantic (as opposed to response) conflict remains unclear (see Introduction), these studies imply that semantic conflict contributes to the overall Stroop interference only in long RSI condition. It should be remembered at this point that short RSI of 200 ms clearly speeded up responses in the present study. Both the significant Stimulus-Type \times RSI interaction and the significant simple main effect of the Stimulus-Type at long RSI were consistent with this latter assumption. Indeed, the results reported above revealed that the overall Stroop interference is composite in its nature when the Stroop task is administered under conventional RSI of 2000 ms. In line with this idea, same-response trials produced significant semantic conflict that was indeed independent of both response conflict and facilitation – as previously reported by Burca et al. (2022) and originally assumed by De Houwer (2003)⁷. It is also important to note that while same-response trials certainly involve at least some response facilitation – since their color- and word-dimension converge toward the same color-response (see Hasshim & Parris, 2014; Parris et al., 2021 for ample discussion of this issue) – this opposing (i.e., facilitatory) effect can only reduce and not inflate the magnitude of semantic interference produced by these trials.

Since unitary or single-stage response competition models (e.g., Cohen et al., 1990a; Glaser & Glaser, 1989; Logan, 1980; Melara & Algom, 2003; Phaf et al., 1990; Roelofs, 2003a) can account for facilitation but not interference produced by same-response trials (see Introduction above), they are clearly unable to account for the present findings showing that semantic conflict along with response conflict contributes to overall Stroop interference. Therefore, the present study provides additional (Burca et al., 2021, 2022) and converging evidence that can only be accounted for via the multi-stage model of Zhang and colleagues (H. Zhang et al., 1999; H. Zhang & Kornblum, 1998a). Taken together, currently existing evidence is likely to suggest that the overall color-word Stroop interference has at least two

⁷ It should however be noted that the issue of facilitation is eluded in this study.

distinct loci. Indeed, while the contribution of task conflict to Stroop interference and facilitation cannot be excluded, future studies need to address its direct contribution to the overall Stroop effect.

If some of the past studies failed to uncover the aforementioned composite nature of the Stroop effect in the two-to-one Stroop paradigm (Hasshim & Parris, 2014, 2015) under virtually the same conditions as in the present study and that of Burca et al. (2022), it is potentially because the contribution of semantic conflict is subject to some boundary conditions. Indeed, the additional and perhaps the most important finding of the present study is that response speed might constitute such a condition – as semantic conflict was statistically eliminated under a short RSI. Therefore, the present study replicates and extends the findings of past studies (see respectively Hasshim et al., 2019; Scaltritti et al., 2022; Sulpizio et al., 2022) in the two-to-one Stroop paradigm providing an unambiguous measure of semantic conflict.

The most important implication of this latter result is that it allows to reconcile contrasting findings of Hasshim and Parris (2014, 2015) and those that of Burca et al. (2022). Indeed, in Experiment 2a of Hasshim and Parris (2014), response times were indeed significantly shorter than even in younger participants of Burca et al. (2022, see Supplementary online materials of this study for non z-scored RTs) – although it is still unclear why participants responded faster in Hasshim and Parris (2014). Therefore, future studies need to address the role of further response speed further since the present findings suggest that any factor that influences response speed should modify the nature of the Stroop effect and that this modification should primarily concern the contribution of semantic conflict to the overall Stroop interference.

The second goal of the present study was to examine closer the benefit of a short (as opposed to long) RSI. In agreement with past studies on the influence of RSI in the color-

word Stroop task, and as expected, RSI of 200 ms (compared to RSI of 2000 ms) induced faster responses in the present study. It also significantly reduced the magnitude of the overall Stroop effect as compared to long RSI (Augustinova et al., 2018; Jackson & Balota, 2013; Parris, 2014; but see also Sharma et al., 2010 for the absence of this benefit and Galer et al., 2014 for the inverse benefit of (very) long RSI). Therefore, the present study extended the benefit of short RSI to the two-to-one Stroop paradigm (De Houwer, 2003).

However, while in line with De Jong et al.'s original reasoning, many subsequent studies agree on the fact that the benefit of a short RSI results from heightened attentional selectivity, they still seem to disagree on the level of processing at which this benefit is actually achieved (see Augustinova et al., 2018 for this type of argument). Therefore, the present study was specifically aimed at examining the level of processing at which the benefit of a short (as opposed to long) RSI takes its effect.

It should be remembered that De Jong and colleagues (1999) originally concluded that individuals are always able to inhibit the influence of a word meaning but that this ability is efficiently deployed only when their attention is fully focused on the color dimension – as it is with short RSI. We reasoned above (see Introduction) that if this conclusion is correct, then both semantic and response conflicts in tandem or semantic conflict alone should be reduced in short RSI condition. In line with the former prediction, the significant reduction of Stroop interference – observed in the present study – was likely to result from the fact that both conflicts were reduced. Therefore, the results reported in the present study are likely to mirror those of Burca and colleagues (2021). Indeed, their results suggest that when only a single letter (as opposed to all letters) of a target display is colored in an incongruent color from the response set and spatially pre-cued (hereafter SLCC vs. ALCC [all-letters colored and cued]), Stroop interference observed in the two-to-one Stroop paradigm is substantially reduced. Also, and importantly, exactly like in this present study, this reduction is due to the joint

reduction of both semantic and response conflicts. The benefit of SLCC is attributed to the fact that (unlike ALCC), it is likely to keep two informational sources (i.e., color and word) separate (i.e., *odd-one-out effect* of SLCC, Besner et al., 1997). Similarly, by enhancing the pace at which the Stroop task is completed, a short (as opposed to long) RSI increases attentional selectivity. Said differently, the present study shows that a short RSI of 200 ms (exactly like SLCC) constitutes another powerful attention-driving moderator that influences people's performance in the two-to-one Stroop paradigm. This finding also implies that the performance of older adults might also benefit from this intervention – the idea that future studies need to address it directly.

The fact that in present study, a short RSI of 200 ms reduced both semantic and response conflicts in tandem runs clearly against Augustinova et al. (2018)' claim that RSI leaves semantic conflict unaffected. Although the differences in the induction of semantic conflict might explain – at least in part – these contrasting findings, the most plausible explanation is that the magnitude of semantic conflict induced via semantic associates is too small to be affected by RSI (or another moderator). Additionally, unlike in Parris (2014a), shortening the RSI failed to boost the magnitude of Stroop facilitation in the present study (and that in concert with reduced interference). Also, and importantly, the magnitudes of facilitation remained comparable in both conditions of RSI. Therefore, the present findings are rather incompatible with the idea that RSI reduces task conflict (as opposed to other types of conflict).

Although this latter conclusions contradict previous stands of our group on the effect of RSI (Augustinova al., 2018; Augustinova & Ferrand, 2014; Parris, 2014), the present study provides clear evidence that the Stroop effect and Stroop interference, it entails, decrease as responses are faster and that the short RSI leads to faster responses. Thus, the present data concur with the account initially proffered by De Jong et al. (1999) that a fast-paced block

(induced by short RSI of 200 ms) optimizes participant's focus on the relevant color-dimension. However, it also remains possible that the changes in goal maintenance – anticipated by De Jong et al. (1999)'s account are not necessary and any factor that – exactly like RSI – changes response speed modifies the nature of the Stroop effect. To illustrate, Stroop studies – including the present one – do not usually control for associative-priming confounds (e.g., Henson et al., 2014; Hommel, 2004; Schmidt & Weissman, 2016) that can substantially impact response speed and therefore – as shown in the present study – impact the type of conflict that is subsequently experienced on a given trial (response conflict alone or response conflict followed by semantic one). In sum, any factor that modifies response speed should modify how much semantic conflict is observed. This would show that the composite nature of Stroop interference (semantic vs. response conflict) and more generally of both Stroop facilitation and interference – originally anticipated by the single multi-stage model of Zhang and colleagues (H. Zhang et al., 1999; H. Zhang & Kornblum, 1998a) – is likely a function of (slow) response speed, and thus of poorer performance. In sum, more empirical studies are needed to address this important issue directly.

5. References

- Augustinova, M., & Ferrand, L. (2014). Automaticity of Word Reading. *Current Directions in Psychological Science*, 23(5), 343–348. <https://doi.org/10.1177/0963721414540169>
- Augustinova, M., Silvert, L., Spatola, N., & Ferrand, L. (2018). Further investigation of distinct components of Stroop interference and of their reduction by short response-stimulus intervals. *Acta Psychologica*, 189, 54–62. <https://doi.org/10.1016/J.ACTPSY.2017.03.009>
- Bench, C. J., Frith, C. D., Grasby, P. M., Friston, K. J., Paulesu, E., Frackowiak, R. S. J., & Dolan, R. J. (1993). Investigations of the functional anatomy of attention using the stroop test. *Neuropsychologia*, 31(9), 907–922. [https://doi.org/10.1016/0028-3932\(93\)90147-R](https://doi.org/10.1016/0028-3932(93)90147-R)
- Brown, T. L. (2011). The relationship between stroop interference and facilitation effects: Statistical artifacts, baselines, and a reassessment. *Journal of Experimental Psychology: Human Perception and*

- Performance*, 37(1), 85–99. <https://doi.org/10.1037/a0019252>
- Burca, M., Beaucousin, V., Chausse, P., Ferrand, L., Parris, B. A., & Augustinova, M. (2021). Is There Semantic Conflict in the Stroop Task? *Experimental Psychology*, 68(5), 274–283. <https://doi.org/10.1027/1618-3169/a000530>
- Burca, M., Chausse, P., Ferrand, L., Parris, B. A., & Augustinova, M. (2022). Some further clarifications on age-related differences in the Stroop task: New evidence from the two-to-one Stroop paradigm. *Psychonomic Bulletin & Review*, 29, 492–500. <https://doi.org/10.3758/s13423-021-02011-x>
- Chen, A., Bailey, K., Tiernan, B. N., & West, R. (2011). Neural correlates of stimulus and response interference in a 2–1 mapping stroop task. *International Journal of Psychophysiology*, 80(2), 129–138. <https://doi.org/10.1016/j.ijpsycho.2011.02.012>
- Chen, A., Tang, D., & Chen, X. (2013). Training Reveals the Sources of Stroop and Flanker Interference Effects. *PLoS ONE*, 8(10), e76580. <https://doi.org/10.1371/journal.pone.0076580>
- Chen, Z., Lei, X., Ding, C., Li, H., & Chen, A. (2013). The neural mechanisms of semantic and response conflicts: An fMRI study of practice-related effects in the Stroop task. *NeuroImage*, 66, 577–584. <https://doi.org/10.1016/j.neuroimage.2012.10.028>
- 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>
- De Houwer, J. (2003). On the role of stimulus-response and stimulus-stimulus compatibility in the Stroop effect. *Memory & Cognition*, 31(3), 353–359. <https://doi.org/10.3758/BF03194393>
- De Jong, R., Berendsen, E., & Cools, R. (1999). Goal neglect and inhibitory limitations: Dissociable causes of interference effects in conflict situations. *Acta Psychologica*, 101(2–3), 379–394. [https://doi.org/10.1016/S0001-6918\(99\)00012-8](https://doi.org/10.1016/S0001-6918(99)00012-8)
- Galer, S., Schmitz, R., Leproult, R., De Tiège, X., Van Bogaert, P., & Peigneux, P. (2014). Response-stimulus interval duration modulates interference effects in the Stroop task. *Psychologica Belgica*, 54(1), 97–110. <https://doi.org/10.5334/pb.ad>
- Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General*, 118(1), 13–42.
- Goldfarb, L., & Henik, A. (2007). Evidence for task conflict in the Stroop effect. *Journal of Experimental Psychology: Human Perception and Performance*, 33(5), 1170–1176. <https://doi.org/10.1037/0096->

- Hasshim, N., Downes, M., Bate, S., & Parris, B. A. (2019). Response Time Distribution Analysis of Semantic and Response Interference in a Manual Response Stroop Task. *Experimental Psychology*, 66(3), 231–238. <https://doi.org/10.1027/1618-3169/a000445>
- Hasshim, N., & Parris, B. A. (2014). Two-to-one color-response mapping and the presence of semantic conflict in the Stroop task. *Frontiers in Psychology*, 5, 1157. <https://doi.org/10.3389/fpsyg.2014.01157>
- Hasshim, N., & Parris, B. A. (2015). Assessing stimulus–stimulus (semantic) conflict in the Stroop task using saccadic two-to-one color response mapping and prereponse pupillary measures. *Attention, Perception, & Psychophysics*, 77(8), 2601–2610. <https://doi.org/10.3758/s13414-015-0971-9>
- Henson, R. N., Eckstein, D., Waszak, F., Frings, C., & Horner, A. J. (2014). Stimulus–response bindings in priming. *Trends in Cognitive Sciences*, 18(7), 376–384. <https://doi.org/10.1016/j.tics.2014.03.004>
- Hershman, R., & Henik, A. (2019). Dissociation between reaction time and pupil dilation in the Stroop task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(10), 1899–1909. <https://doi.org/10.1037/xlm0000690>
- Hershman, R., & Henik, A. (2020). Pupillometric contributions to deciphering Stroop conflicts. *Memory & Cognition*, 48(2), 325–333. <https://doi.org/10.3758/s13421-019-00971-z>
- Hommel, B. (2004). Coloring an action: Intending to produce color events eliminates the Stroop effect. *Psychological Research*, 68(2–3), 74–90. <https://doi.org/10.1007/s00426-003-0146-5>
- Jackson, J. D., & Balota, D. A. (2013). Age-Related Changes in Attentional Selection: Quality of Task Set or Degradation of Task Set Across Time? *Psychology and Aging*, 28(3), 744. <https://doi.org/10.1037/A0033159>
- Jiang, J., Zhang, Q., & Van Gaal, S. (2015). EEG neural oscillatory dynamics reveal semantic and response conflict at difference levels of conflict awareness. *Scientific Reports*, 5(1), 12008. <https://doi.org/10.1038/srep12008>
- Kalanthroff, E., Davelaar, E. J., Henik, A., Goldfarb, L., & Usher, M. (2018). Task conflict and proactive control: A computational theory of the Stroop task. *Psychological Review*, 125(1), 59–82. <https://doi.org/10.1037/rev0000083>
- Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology*, 12(4), 523–553. [https://doi.org/10.1016/0010-0285\(80\)90019-5](https://doi.org/10.1016/0010-0285(80)90019-5)
- 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>
- Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review*, 110(3), 422–471. <https://doi.org/10.1037/0033-295X.110.3.422>
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, 36(3), 516–524. <https://doi.org/10.3758/BF03195598>
- Parris, B. A. (2014). Task conflict in the Stroop task: When Stroop interference decreases as Stroop facilitation increases in a low task conflict context. *Frontiers in Psychology*, 5, 1182. <https://doi.org/10.3389/fpsyg.2014.01182>
- Parris, B. A., Hasshim, N., Wadsley, M., Augustinova, M., & Ferrand, L. (2021). The loci of Stroop effects: A critical review of methods and evidence for levels of processing contributing to color-word Stroop effects and the implications for the loci of attentional selection. *Psychological Research*, 0123456789. <https://doi.org/10.1007/s00426-021-01554-x>
- Phaf, R. H., Van der Heijden, A. H. C., & Hudson, P. T. W. (1990). SLAM: A connectionist model for attention in visual selection tasks. *Cognitive Psychology*, 22(3), 273–341. [https://doi.org/10.1016/0010-0285\(90\)90006-P](https://doi.org/10.1016/0010-0285(90)90006-P)
- Roelofs, A. (2003). Goal-Referenced Selection of Verbal Action: Modeling Attentional Control in the Stroop Task. *Psychological Review*, 110(1), 88–125. <https://doi.org/10.1037/0033-295X.110.1.88>
- Šaban, I., & Schmidt, J. R. (2021). Stimulus and response conflict from a second language: Stroop interference in weakly-bilingual and recently-trained languages. *Acta Psychologica*, 218, 103360. <https://doi.org/10.1016/j.actpsy.2021.103360>
- Scaltritti, M., Job, R., & Sulpizio, S. (2022). Different types of semantic interference, same lapses of attention: Evidence from Stroop tasks. *Memory & Cognition* 2021, 1–13. <https://doi.org/10.3758/S13421-021-01256-0>
- Schmidt, J. R., & Cheesman, J. (2005). Dissociating stimulus-stimulus and response-response effects in the stroop task. *Canadian Journal of Experimental Psychology*, 59(2), 132–138. <https://doi.org/10.1037/H0087468>
- Schmidt, J. R., Hartsuiker, R. J., & De Houwer, J. (2018). Interference in Dutch–French Bilinguals. *Experimental Psychology*, 65(1), 13–22. <https://doi.org/10.1027/1618-3169/a000384>
- Schmidt, J. R., & Weissman, D. H. (2016). Congruency sequence effects and previous response times: Conflict adaptation or temporal learning? *Psychological Research*, 80(4), 590–607.

<https://doi.org/10.1007/s00426-015-0681-x>

Schneider W., Eschman A., Zuccolotto A. (2002). *E-Prime user's guide*. Pittsburgh, USA: Psychology Software Tools Inc.

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>

Shichel, I., & Tzelgov, J. (2018). Modulation of conflicts in the Stroop effect. *Acta Psychologica*, 189, 93–102. <https://doi.org/10.1016/j.actpsy.2017.10.007>

Sulpizio, S., Job, R., Leoni, P., & Scaltritti, M. (2022). Prepotent task-irrelevant semantic information is dampened by domain-specific control mechanisms during visual word recognition. *Quarterly Journal of Experimental Psychology*, 75(3), 390–405. <https://doi.org/10.1177/17470218211030863>

Van Veen, V., & Carter, C. S. (2005). Separating semantic conflict and response conflict in the Stroop task: A functional MRI study. *NeuroImage*, 27(3), 497–504. <https://doi.org/10.1016/J.NEUROIMAGE.2005.04.042>

Zhang, H., & Kornblum, S. (1998). The Effects of Stimulus-Response Mapping and Irrelevant Stimulus-Response and Stimulus-Stimulus Overlap in Four-Choice Stroop Tasks with Single-Carrier Stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 24(1), 3–19. <https://doi.org/10.1037/0096-1523.24.1.3>

Zhang, H., Zhang, J., & Kornblum, S. (1999). A Parallel Distributed Processing Model of Stimulus–Stimulus and Stimulus–Response Compatibility. *Cognitive Psychology*, 38(3), 386–432. <https://doi.org/10.1006/COGP.1998.0703>

CHAPTER FOUR: HOW TO EFFECTIVELY IMPROVE OLDER ADULTS' PERFORMANCE IN THE STROOP TASK?

1. Introduction

Selective attention is often defined as the ability to focus on one thing while ignoring other things (see e.g., Maquestiaux, 2017; Ruthruff & Lien, 2016). This ability to selectively attend to and process only certain things in the environment while ignoring or actively inhibiting others is indeed crucial in many cognitive activities (e.g., Jackson & Balota, 2013). That is perhaps why the famous Stroop task (Stroop, 1935) has achieved such widespread use in clinical, experimental, and neuropsychological research and practice. It addresses this ability by requiring individuals to identify, as quickly and accurately as possible, the font color of written characters (i.e., to attend to and process a relevant color dimension) without reading them (i.e., to simultaneously ignore or inhibit an irrelevant word dimension). The typical result is that a participant's color identification times are longer for *color-incongruent* Stroop words (i.e., words that are displayed in a color that is different from the one they designate such as “BLUE” displayed in green; hereafter *BLUE_{green}*), than for *color-neutral* words (e.g., the word “DEAL” displayed in green, hereafter *DEAL_{green}*).

The magnitude of this latter difference – referred to as Stroop interference – is usually thought to reflect selective attention abilities. Since these abilities are thought to decline with healthy aging (see Comalli et al., 1962, for the first empirical demonstration), numerous other studies have reported higher levels of Stroop interference in healthy older than in young adults (e.g., Andrés et al., 2008; Aschenbrenner, & Balota, 2015; Belleville, et al., 2006; Davidson et al., 2003; Dulaney & Rogers, 1994; Mayas et al., 2012, Wolf et al., 2014). Also and importantly, these difference persist even after controlling for the speed of processing

(e.g. Aschenbrenner et al., 2017; Aschenbrenner & Balota, 2015; Bugg et al., 2007; Burca et al., 2021; Jackson & Balota, 2013, 2013; Li & Bosman, 1996; Nicosia & Balota, 2020; Spieler et al., 1996). Given that selective attention is of prime importance in accomplishing daily-life activities that contribute to maintain the independence of the older adults (e.g., driving, e.g., Pollatsek et al., 2012), the present study was aimed at assessing whether and the extent to which the substantial magnitude of Stroop interference that the older adults usually display can be reduced through attention driving interventions.

2. Experiment 1: The influence of variations in response-stimulus intervals on the elderly's performance in the Stroop task

To this end, the present experiment examined the extent to which older participants (i.e., older than 60 years) are sensitive to a procedure designed to promote attentional focus on the relevant task set (i.e., color dimension of Stroop words). This idea derives directly from the past work of Jackson and Balota (2013), itself a continuation of that of De Jong and colleagues (1999).

De Jong and colleagues initially reasoned that instead of reflecting limitations in (young) people's inhibitory abilities, high levels of interference might instead reflect their limitations in deploying these inhibitory abilities consistently across time. They additionally reasoned that this deployment is more likely to be efficient when fast responding is encouraged, as it is the case when the time that elapses between the individual's response and the presentation of a new stimulus on a computer screen (i.e., response stimulus interval, hereafter RSI) is shortened. In line with this reasoning that people's attention is more consistently focused on the relevant dimension of the Stroop task when the conventional response-stimulus interval (hereafter RSI) of 2000 ms is shorten to 200 ms, Stroop(-like) interference they showed in this latter condition is clearly reduced, compared to the former.

Jackson and Balota (2013) replicated this benefit of short RSI in younger participants and extended it to older ones ($M_{\text{age}} = 77$ years). More specifically, in their Experiment 2, the Stroop effect (i.e., the difference between standard color-incongruent and color-congruent trials) was reduced at the short RSI as compared to the long RSI, but there was no evidence of an interaction between RSI and age group. Indeed, at both RSIs, the level of the Stroop effect was greater in the older (compared to the younger) participants. Jackson and Balota subsequently concluded that there is an age-related breakdown in the quality of attentional selection processes, namely the inhibitory abilities, but that both age groups are similar in their (in)ability to maintain an appropriate task set across time. Indeed, if older participants had found it more difficult to maintain a focus on the relevant color dimension across time, they would have produced greater Stroop effect than the younger participants at the long RSI compared to the short RSI (i.e., leading to an interaction between age and RSI). To summarize, the absence of this interaction led Jackson and Balota (2013) to refine the standard interpretation of age-related differences in Stroop interference in terms of attentional selectivity. Indeed, the data from their study suggest that healthy aging lowers the ability to *ignore* or *inhibit* an irrelevant word dimension of Stroop words rather than the ability to *attend to* and *process* the relevant color dimension of these words (see also e.g., Bugg et al., 2007; Wolf et al., 2014).

The present experiment was designed to address this conclusion in more details. First of all, Jackson and Balota (2013) reasoned about Stroop interference from the overall Stroop effect (i.e., a sum of facilitation and interference – each in unknown amounts, see MacLeod, 1991; Parris et al., 2021 for discussion). Therefore, the present experiment used a design allowing to partial out the contribution of each of the two aforementioned Stroop phenomena and thus the examine the extent to which interference (as opposed to facilitation) is indeed reduced by short RSI – as originally argued by Jackson and Balota (2013).

Second, if in line with their interpretation of age-related differences in Stroop interference, it is indeed the ability to ignore or inhibit an *irrelevant word dimension* of Stroop words that is lowered in older adults, then the magnitude of semantic conflict they display should be higher in the condition of long RSI (2000 ms) as compared to short RSI of 200 ms. This would then extend findings of Burca and colleagues' work (2022) showing that older individuals display significantly higher magnitude of semantic as compared to response conflict – running against the idea that the ability to inhibit of *prepotent responses* (or response inhibition, Miyake et al., 2000) is reduced in older adults.

In sum, the present experiment was designed to examine the type of conflict that short RSI of 200 ms optimizes in older adults – the issue that despite its theoretical and clinical importance remains so-far unanswered. In line with past studies, examining this issue actually requires that – exactly like in Burca et al. (2022), semantic conflict unambiguously contributes to Stroop interference displayed in older participants. Therefore, the present study was also designed to replicate this finding.

2.1 Method

Participants and Inclusion process and criteria

Sixty community dwelling healthy older adults (35 females and 25 males, $M_{\text{age}} = 69.96$, $SD_{\text{age}} = 5.11$) participated in this experiment. All participants were native French-speakers, and reported normal or corrected-to-normal vision, without impairment in color discrimination (see Table 1S in Supplementary Materials for further demographic and psychometric information). To ensure they fit the inclusion criteria (see Supplementary Materials for more information) the participants first signed an informed consent, and completed a demographic interview. In addition to collect demographic information, the interview ensured that the participants were not undergoing any medical treatment that is

known to impact the nervous system and therefore cognitive abilities, that they did not have had a head injury, and that they were not diagnosed with a psychiatric and/or neurological disorder.

Moreover, the participants completed a psychometric evaluation battery, the goal being to ensure that the older adults presented indeed with a healthy aging. To this, the participants first underwent the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005), where a cutoff score higher than 26 is considered in the norm. Then a depression/anxiety questionnaire was administered, the Hospital Anxiety and Depression Scale-HAD (Zigmond, 1983), where a score higher than 7 was considered borderline abnormal. Various cognitive abilities were then assessed: working memory (via administering the forward and backward digit span of WAIS, Wechsler, 2008); processing speed (via french equivalent, of the letter-comparison test of Salthouse, 1990; see Bugaiska et al., 2007); and finally, the vocabulary/crystallized abilities were assessed (via use of Mill-Hill, Raven, 1958; for scores and means of this demographic information see Table 1S). All the participants fit the inclusion criteria, and none presented abnormal/pathological scores at the psychometric evaluation.

Procedure, apparatus, stimuli and design.

All participants received a 20€ voucher in exchange for their participation. Following the psychometric evaluation depicted above, the participants completed a computerized version of the Stroop Task presented via Eprime 2.0 software (Schneider et al., 2002). To this end, the participants were seated at a distance of 70cm in front of the 13” portable computer, and instructed to identify the color of the stimulus presented on the screen, as quickly and accurately as possible by pressing the appropriate color-button, while ignoring everything else on the display. They were specifically instructed to concentrate on the fixation cross (“+”) that

appeared in the center of the screen at the beginning of each trial, with a duration of 200msec in the short RSI condition vs. a duration of 2000msec in the long RSI condition. The participants were assigned randomly to the short or long RSI condition. The stimulus remained on the screen until the participants gave an answer or until 3500msec had elapsed. All stimuli were presented in lowercase Courier font, size 18, on a black background and subtended an average visual angle of 0.9° high \times 3.0° wide. The used color-words were: *rouge* [red], *jaune* [yellow], *bleu* [blue], and *vert* [green], and the neutral non-color words were: *plomb* [lead], *liste* [list], *page* [page], *cave* [basement]. The color and non-color words were matched in length and frequency via *Lexique 3.38* (New et al., 2004).

The 2-to-1 paradigm was used in this study, thus the participants responded manually by pressing the “k” key – which was labeled with the colors red and blue, or the “d” key – which was labeled with the colors green and yellow on an AZERTY keyboard. Before the experimental phase, all participants underwent a training session, which consisted of 60 practice trials, composed of colored asterisks. Both color-words and neutral words were presented in all colors presented above. The experimental block consisted of 120 stimuli, and contained 48 stimuli different-response stimuli, 24 same-response stimuli, 24 neutral stimuli, and 24 congruent stimuli. This block was repeated 9 times with a self-paced break between blocks. Thus, a total of 1080 stimuli were presented to each participant. As a result, present study was function of a 4 (Stimulus-Type: different-response vs. same-response vs. neutral vs. congruent) \times 2 (RSI: short vs. long) design with the RSI as a between subject factor.

2.2 Results and Discussion

Five older adults were excluded from the analysis: four due to a high number of errors (>25%) and one due to a malfunction of the data recording. Thus, the data of fifty-five older adults (34 females and 21 males, $M_{age} = 69.67$, $SD_{age} = 5.18$) was used for the statistical

analysis: 25 older adults assigned to short RSI condition and 30 older adults assigned to long RSI condition¹. As in Balota et al. (2013), the first 100 trials were considered as practice, and thus were not taken into account during the statistical analysis (i.e., 1080 trials in total were presented, and 980 trials per participant were analyzed). RTs greater than 3SDs above or below each participant's mean latency for each condition were excluded from the analysis (i.e., 1.85% of the total data).

Both RTs and errors were first analyzed in an omnibus 4 (Stimulus-Type: different-response vs. same-response vs. neutral vs. congruent) \times 2 (RSI: short vs. long) was conducted in both classical and Bayesian ANOVAs, with the RSI being a between-subjects factor. On RTs, this analysis revealed a significant main effect of Stimulus-Type [$F(3,159) = 53.51$; $p < .001$, $\eta_p^2 = 0.502$; $BF_{10} = 1.706e+21$]. The main effect of the RSI condition was not significant [$F(1,53) = .062$; $p = .804$, $\eta_p^2 = 0.001$; $BF_{10} = 0.568$] and $BF_{01} = 1.761$ actually provided at least anecdotal evidence against the effect of RSI. Stimulus-Type \times RSI interaction was not significant either [$F(1,53) = .062$; $p = .804$, $\eta_p^2 = 0.001$; $BF_{10} = 0.050$] and $BF_{01} = 19.744$ provided with strong Bayesian evidence against this interaction (see Table 2S in Supplementary materials for a full descriptive statistics).

The same pattern of results was observed on the error rates. The main effect of Stimulus-Type was significant [$F(3,159) = 35.02$; $p < .001$, $\eta_p^2 = 0.398$; $BF_{10} = 1.898e+14$]. The main effect of RSI was only marginally [$F(1,53) = 3.09$; $p = .085$, $\eta_p^2 = 0.055$; $BF_{10} = 0.846$] but $BF_{01} = 1.182$, actually showed at least anecdotal evidence against this effect in errors. The Stimulus-Type \times RSI interaction did not reach significance either [$F(3,159) = 2.20$; $p = .090$, $\eta_p^2 = 0.040$; $BF_{10} = 0.947$] with again $BF_{01} = 1.815$ providing anecdotal

¹The uneven number of participants is due to the impossibility of completing the inclusion of participants, due to COVID-19 sanitary crisis. Note however, that the number of participants was comparable to that of Jackson and Balota (2013) that tested 27 older adults at each level of RSI.

evidence against this interaction by Bayesian evidence (for a full description of the analysis on error rates see Table 2S in Supplementary Materials).

Is there semantic conflict in the Stroop task?

To answer this question – constituting a necessary pre-requisite to examining the type of conflict that is influenced by RSI-manipulation in the Stroop task –, the significant main effect of the Stimulus-Type observed in the ANOVA conducted on both RTs and errors was further decomposed. On errors, it revealed that more errors were committed for different-response stimuli (4.46%) as compared to same-response (1.95%; $p = .002$; $BF_{10} = 92.77$) and congruent (2.15%) stimuli ($p < .05$; $BF_{10} = 38.39$). The neutral stimuli, surprisingly, had the highest error rate (6.14%), however as compared to different-response stimuli, this difference was not statistically significant ($p = .061$; $BF_{10} = 9.42$), however this effect was supported by moderate Bayesian evidence. However, the error rate committed for neutral trials was higher as compared to same-response and congruent trials (both p 's $< .001$; both BF s > 100 , providing extreme evidence in favor of this effect).

On RTs, this decomposition revealed that of the different-response trials were longer than those observed to all the other stimuli (all p 's $< .001$; BF s > 100 , see Table 3S for complete descriptive statistics). This means that the participants (i.e., older adults) displayed a significant overall Stroop effect (significant difference between different-response and congruent stimuli) of an important magnitude of 112ms ($p < .001$; $BF_{10} = 1.147e+10$). It was further composed by a significant facilitation effect (difference between neutral and congruent stimuli) 24ms ($p < .001$; $BF_{10} = 327.00$) and a significant Stroop interference (difference between different-response and neutral stimuli) 88ms ($p < .001$; $BF_{10} = 4.278e+6$). This latter Stroop phenomenon was in turn composed by both a significant response conflict (difference between different-response and same-response stimuli) with a magnitude of 39ms ($p < .001$;

$BF_{10} = 1.1915.70$), and a semantic conflict (difference between same-response and neutral stimuli) with a magnitude of 50ms ($p < .001$; $BF_{10} = 123497.08$). These results therefore replicate those of Burca and colleagues (2022, see Chapter 2) showing that both significant facilitation and interference effect contribute to the overall Stroop effect observed in older adults. Also, and importantly, exactly like in this study both a semantic and response conflict contributed to the overall Stroop interference these participants displayed. Finally, as in these past studies and contrary to those conducted with the Semantic Stroop paradigm (Augustinova et al., 2018a), older participants displayed substantially greater magnitude of semantic as compared to response conflict. The direct implication of this finding is that there is therefore a room for this magnitude to be reduced.

Does RSI manipulation improve older adults' performance in the Stroop task?

As mentioned above, neither the main effect of RSI, nor the Stimulus-Type \times RSI interaction found by Balota et al. (2013) was replicated in the present experiment. In order to explore this lack of effect of the RSI, additional analyses were conducted. First, Bayesian sequential analysis were conducted for all stimuli types, with RSI as grouping variable (see Figure 1S for these graphs in Supplementary Materials). This analysis revealed that the RSI did not seem to gain in power if more participants would be added. If anything, this analysis suggests that the RSI is most likely to have an effect when the sample consists of 20-30 participants. Therefore, the uneven number of participants in the short vs. long RSI condition might have potentially influenced this lack of effect with this sample size.

Second, to address the possibility that only the slow (vs. fast) participants and/or slow (vs. fast) trials benefited from short RSI, the extent to which 4 (Stimulus-Type: different-response vs. same-response vs. neutral vs. congruent) \times 2 (RSI: slow vs. fast) was moderated respectively by participants' overall (fast vs. slow participants) and relative speed (fast vs.

slow trials). Results of these analyses (see Supplementary Materials) suggest that regardless of both, the variations in RSI failed to affect participants' response. Said differently, neither the fast nor the slow participants or fast or slow trials benefited from short RSI.

In light of this lack RSI effect, the present experiment failed to provide any support for the idea that shortening habitual RSI from 2000 ms to 200 ms improves participants' performance in the Stroop task. While this seems to be at odd with Jackson and Balota (2013) observing this effect in older adults and with other studies (Augustinova et al., 2018a; Jackson & Balota, 2013b; Parris, 2014, see also Chapter 3) observing this effect in younger adults, it should be noted that other studies failed to observed this benefit (e.g., Martinon et al., submitted; Sharma et al., 2010 for the absence of this benefit and Galer et al., 2014 for the inverse benefit of (very) long RSI). It however does not mean that selective attention in older adults cannot be improved. Therefore, the next experiment addressed this issue with Single Letter Coloring and Cueing (SLCC).

3.1 Experiment 2: The influence of Single letter coloring and cueing

Indeed, since Kahneman and Henik's (1981) pioneering work, the presentation of Stroop stimuli in which only a single letter (as opposed to all letters) of a target display is colored in an incongruent color from the response set and spatially pre-cued (hereafter SLCC vs. ALCC [all-letters colored and cued] is known to promote attentional focus on the relevant color dimension of Stroop words. This is because SLCC (unlike ALCC) is likely to keep two informational sources (i.e., color and word) separate (i.e., *odd-one-out effect* of SLCC, Besner et al., 1997). Consistent with reasoning is the fact that ERPs are sensitive to SLCC procedure early in processing of Stroop words (i.e., before being actually sensitive to their color-incongruency) such that SLCC elicits significantly greater negativity (from 185 to 245 ms) at occipito-parietal sites compared to ALCC (see Augustinova et al., 2018a for discussion of this

unpublished data). Consequently, an information provided by the irrelevant word-dimension of compound Stroop items (e.g., “blue” for *BLUE*_{yellow}) is thought to interfere less with that induced by the relevant color-dimension (e.g., “yellow” for *BLUE*_{yellow}) in SLCC than in ALCC (but see e.g., Küper & Heil, 2012 for a review another process potentially driving this effect).

Results of a substantial number of past studies are consistent with these ideas. Indeed, they have reported SLCC to substantially reduce or even eliminate Stroop interference (e.g., Augustinova & Ferrand, 2007; Augustinova et al., 2010; Augustinova et al., 2015; Augustinova, Clarys, Spatola, & Ferrand, 2018; Besner & Stolz, 1999ab; Brown, Joneleit, Robinson, & Brown, 2002; Catena, Fuentes, & Tudela, 2002; Küper & Heil, 2012; Manwell, Roberts, & Besner, 2004; Manwell et al., 2004; Mari-Beffa et al., 2000; Marmurek, 2003; Monahan, 2001; Parris, Sharma, & Weekes, 2007). Therefore, unsurprisingly, Augustinova and colleagues (2018b) set out to examine the extent to which it SLCC (as compared to ALCC) reduces Stroop interference in older as compared to younger adults with a specific goal to address the type of conflict that SLCC actually reduces.

Their results replicated those reported in studies mentioned above. Indeed, SLCC (as opposed to ALCC) significantly reduced standard Stroop interference. But it failed to influence significantly semantic-associative interference. Therefore, in line with their past studies (Augustinova & Ferrand, 2007; Augustinova et al., 2010; Augustinova et al., 2015); Augustinova and colleagues (2018b) have attributed this reduction to response (as opposed to semantic conflict).

Additionally, Age-group also interacted in this latter study with the Type of interference such that older adults displayed significant higher levels standard Stroop interference (as compared to their younger counterparts) but both age-groups showed comparable level of semantic-associative interference and no 3-way interaction was observed

between Interference-Type (standard vs. semantic) \times Age-group (younger vs. older participants) and Coloring-Type: SLCC vs. ALCC). Augustinova and colleagues had consequently argued that both younger and older adults are equally (in)-efficient in suppressing the irrelevant word dimension of Stroop words (e.g., reading the word). Given that controlled processes are likely to decline with healthy ageing whereas those that are automatic are likely to be preserved (and start to decline only with pathological ageing, see e.g. Lemaire & Bherer, 2005 for a review), the hypothesis of semantic conflict being automatic was advanced. In this view, the suppression of this latter type of conflict is indeed unlikely (Augustinova & Ferrand, 2014).

Also, and importantly for the issue under consideration here, both young and older adults seemed to be highly responsive to this SLCC manipulation in this study and given the higher baseline levels of Stroop interference in older adults (compared to young one), older adults seem to benefit even more from this SLCC manipulation. However, this latter benefit was clearly observed at the level of response conflict, as opposed to semantic one. This latter reasoning seems to reinforce that of Jackson and Balota (2013; see also e.g., Bugg et al., 2007; Wolf et al., 2014) that the ability to *attend to* and *process* the relevant color dimension of Stroop words remains preserved in healthy ageing. Otherwise, both Coloring \times Age-Group and the overall Interference-Type \times Coloring \times Age-Group interactions would have been significant.

However, the fact regardless of age, suppression of semantic conflict seemed unlikely in this study, simultaneously challenged Jackson and Balota's reasoning that healthy aging lowers the ability to *ignore* or *inhibit* an irrelevant word dimension of Stroop words. In line with more recent meta-analytic findings (Rey-Mermet & Gade, 2018), higher levels of response conflict in older participants are consistent with the idea that healthy ageing lowers

the ability to inhibit irrelevant responses (i.e., response primed by the irrelevant word-dimension of Stroop words).

Nonetheless, the use of semantic associative interference as a proxy for measuring of semantic conflict has been criticized (Hashim & Parris, 2014, 2015; Parris et al., 2021; see also Burca et al., 2021, 2022). Also, and importantly, Burca et al. (2022) have recently argued that the absence of an age effect on semantic associative interference in Augustinova et al.' (2018, see also Li & Bosman, 1996) is likely due to proportional transformation applied in this study to control for age-related general slowing. Somewhat counterintuitively, this type of transformation (unlike z-scores) creates an advantage for older adults in the presence of slower RTs (see e.g., Faust et al., 1999; Hedge et al., 2018 for arguments making z-scores more suitable transformation). Indeed, the re-analysis of the z-scored data from Augustinova et al. (2018) by Burca and colleagues (2022) showed that the originally significant Interference-Type \times Aging interaction become non-significant, and the additional Bayesian analyses actually provided moderate evidence against this interaction. This suggests that the magnitudes of both semantic and response conflict in Augustinova et al.'s z-scored data tended to be greater in older adults than in their younger counterparts.

While the results regarding semantic conflict are in line with those reported in 2-to-1 Stroop paradigm (De However, 2003) with an additional color-neutral baseline (Hashim & Parris, 2014) by Burca et al. (2022), discrepancies currently remain regarding the effect of healthy aging on response conflict. Indeed, this latter study had actually found that healthy ageing leaves response conflict unaffected. Therefore, the goal of the present experiment was to address this remaining issue along with the one of SLCC-benefit on semantic vs. response conflict. Indeed, recent study of Burca and colleagues (2021) seems rather inconclusive with regard to the effect of SLCC in the 2-to-1 Stroop paradigm. Therefore, the experiment reported below was aimed at shedding an additional light on this issue.

3.2 Method

Participants and Inclusion process and criteria

Fifty-one community dwelling older adults (37 females and 14 males; $M_{\text{age}} = 68.54$, $SD = 5.70$) volunteered to participate in this study. All participants were native French-speakers, reported normal or corrected-to-normal vision, and presented no impairment in color discrimination (see Table 3S in Supplementary Materials for further demographic and psychometric information). To ensure they fit the inclusion criteria, as in Experiment 1, after they had participants signed an informed consent, the participants first completed a demographic and anamnestic interview, which was identical to the one reported in Experiment 1 (see Supplementary materials for more information) and then underwent a psychometric evaluation battery. Results of showed that all the participants fit the inclusion criteria meaning that none presented abnormal/pathological scores at the psychometric evaluation.

Procedure, Apparatus, Stimuli and Design

After they had completed the aforementioned demographic and anamnestic interview and psychometric evaluation, participants completed a computerized version of the Stroop Task similar to Burca et al. (2021, see also Chapter 1), with one minor difference that instead of seeing 240 stimuli in total (i.e., 120 stimuli in each of the Coloring-condition), the present study, the participants have seen 192 trials in total (i.e., 96 stimuli in each of the Coloring-condition, see also Supplementary Materials of this chapter for details of Apparatus and Stimuli). Therefore, as in Burca et al. (2021), the data was collected using a 4 (Stimulus-Type: different-response vs. same response vs. neutral different response vs. neutral same response) \times 2 (Coloring: All Letters Colored and Cued - ALCC vs. Single Letter Colored and Cued - SLCC) fully within-participants experimental design.

3.3 Results and Discussion

No participants were excluded from the present study. The error rate was extremely low in the present dataset (1.01%), and results on such a small data sample cannot be meaningfully interpreted, thus no analysis on the error rates were conducted. RTs greater than 3SDs above or below each participant's mean latency for each condition were excluded from the analysis (resulting in the exclusion of 0.88% of total data). Remaining RTs from 51 older adults were subsequently analyzed in an omnibus 4 (Stimulus-Type: DR-incongruent vs. SR-incongruent vs. DR-neutral response vs. SR-neutral) \times 2 (Coloring: All Letters Colored and Cued - ALCC vs. Single Letter Colored and Cued - SLCC), with both standard frequentist and Bayesian ANOVA.

It revealed the significant main effect of Stimulus-Type [$F(3,150) = 17.529$; $p < .001$, $\eta_p^2 = 0.260$; $BF_{10} = 2.576e+7$] and marginally significant of Coloring [$F(1,50) = 3.477$; $p = .068$, $\eta_p^2 = 0.065$], with $BF_{10} = 1.217$ providing anecdotal evidence in favor of such this effect. Finally, Stimulus-Type \times Coloring interaction revealed to be significant [$F(3, 150) = 6.437$; $p < .001$, $\eta_p^2 = 0.114$] with $BF_{10} = 11.108$ providing moderate Bayesian evidence in favor of this interaction.

Is there semantic conflict in the Stroop task?

As above (see section 2.2), to answer this question – constituting a necessary pre-requisite to examining the type of conflict that is influenced by SLCC-manipulation in the Stroop task –, the aforementioned significant main effect of Stimulus type was decomposed. This decomposition showed that RTs to DR-incongruent stimuli were significantly longer than those observed for their DR-neutral counterparts ($p < .001$; $BF_{10} = 7.472e+7$) – revealing at hand significant standard Stroop interference of 59msec. This decomposition also showed that RTs to SR-incongruent stimuli were significantly longer than those observed for their SR-

neutral counterparts ($p = .033$; $BF_{10} = 12.609$) – revealing that semantic Stroop interference of 49 ms was also significant. However, exactly like in Burca et al. (2021)’ study, the difference between RTs to DR-incongruent and SR-incongruent remained not significant ($p = 1.000$; $BF_{10} = 0.509$) and $BF_{01} = 1.966$ actually provided anecdotal evidence against this effect. In sum, while these results again attest that the contribution of semantic conflict to the overall (i.e., standard Stroop interference) undeniably occurs, this contribution does not seem mandatory for response conflict.

To explore this issue of semantic conflict further, we then decomposed the aforementioned Stimulus-Type \times Coloring interaction by testing the simple main-effect of Stimulus-Type at each level of Coloring (see Table 1 for descriptive statistics). This simple main-effect was significant in both ALCC [$F(3,48) = 14.28$; $p < .001$, $\eta_p^2 = 0.472$; $BF_{10} = 6.779e+7^2$] and SLCC [$F(3,48) = 10.00$; $p < .001$, $\eta_p^2 = 0.385$; $BF_{10} = 2.591/BF_{01} = 0.386$] and revealed significant standard Stroop interference (DR-incongruent – DR-neutral trials) in both conditions of Coloring: ALCC- ($p < .001$; $BF_{10} = 890838.28$) and SLCC-condition respectively ($p < .001$; $BF_{10} = 2674.350$, see Table 1 for descriptive statistics). Further pairwise comparisons, again unequivocally revealed the presence of significant semantic Stroop interference (SR-incongruent – SR-neutral trials) in ALCC-condition ($p = .003$) that was again supported by strong Bayesian evidence ($BF_{10} = 94.51$). However, in SLCC-condition, the contribution of semantic conflict was no longer significant ($p = 1.000$; $BF_{10} = 1.000$) and $BF_{01} = 4.615$ had actually provided moderate evidence against this effect. Interestingly enough this elimination of semantic Stroop interference is actually driving Stimulus-Type \times Coloring interaction as the contrasts between the two incongruent stimuli (DR-incongruent

²A Bayesian repeated measures ANOVA compares a series of different models against a null model. It therefore cannot be decomposed by testing simple main effects per se. Therefore, BFs provided test main effect of Stimulus-Type in ALCC and SLCC respectively – which explains why a Bayesian approach that is generally more conservative, actually yields larger effects.

SR-incongruent, i.e., response conflict) remained non-significant in both ALCC- ($p = 1.000$; $BF_{10} = 0.669/BF_{01} = 1.496$) and SLCC-condition ($p = 1.000$; $BF_{10} = 0.238/BF_{01} = 4.206$).

Furthermore, unlike in Burca and colleagues (2021)' study, there was no additional evidence for response conflict in slow (as opposed to fast) participants (see Supplementary Materials of their study for these additional analyses) and adding more participants would have not solved this issue (see Supplementary Materials here below for these sequential analyses). However, given that the overall (i.e., standard) Stroop interference remained significant in SLCC-condition despite the complete elimination of semantic Stroop interference (i.e., semantic conflict) suggests that at least some conflict (probably the sum of semantic and response conflict – each in unknown and separately in non-significant amount) undeniably occurs.

The presence of significant semantic interference overall and in ALCC-condition is thus consistent with the idea that a semantic conflict indeed genuinely contributes to the overall Stroop interference effect even when color-neutral *words* are used as baseline (see also Hershman & Henik, 2020' pupillometric measures, but see Hashim & Parris, 2014, 2015; Hershman & Henik, 2020' RTs). And this contribution occurs without any presence of color-congruent items (see also Burca et al., 2021) that are known to boost interference (Roelofs, 2010) and therefore possibly enlarge magnitudes of different types of conflict. However, again, as Burca et al. (2021), the contribution of response conflict to the overall (i.e., standard) Stroop interference seems optional.

Does SLCC manipulation improve older adults' performance in the Stroop task?

The aforementioned elimination of semantic interference in SLCC-conditions clearly indicates that SLCC improves older adults' performance in the Stroop task to the point of complete elimination of semantic interference. Consistent with this latter idea, SLCC (as

compared to ALCC) indeed significantly shortened RTs to SR-incongruent trials [$F(1,50) = 6.37$; $p = .015$, $\eta_p^2 = 0.113$; $BF_{10} = 5.281$]. However, SLCC also significantly shortened RTs to DR-incongruent trials [$F(1,50) = 11.85$; $p = .001$, $\eta_p^2 = 0.192$; $BF_{10} = 48.94$] and again, as in Burca and colleagues (2021)' study, it failed to reduce RTs to both color-neutral trials. Indeed, the lack of simple main effect of Coloring on both DR-neutral [$F(1,50) = 0.05$; $p = .824$, $\eta_p^2 = 0.001$; $BF_{10} = 0.129/BF_{01} = 7.736$] and SR-neutral trials [$F(1,50) = 1.10$; $p = .299$, $\eta_p^2 = 0.022$; $BF_{10} = 0.080/BF_{01} = 12.557$] revealed that color-neutral trials remained unaffected by Coloring manipulation (but see e.g., Küper & Heil, 2012 for this result). These results therefore suggest SLCC influences both semantic and standard Stroop interference *per se*, such that it reduces RTs to incongruent trials.

Table 1. *Color-Identification Performance (Means, Standard Deviation, 95%-Confidence Intervals) observed as a Function of Stimulus- or Interference- Type and Coloring*

	All Letters Colored and Cued (ALCC)		Coloring Single Letter Colored and Cued (SLCC)		SLCC effect (RT)
	M (SD)	CI	M (SD)	CI	
Color-Incongruent Different Response (DR)	1121 (265)	[1047,1196]	1077 (263)	[1003,1150]	- 44 ^{††}
Color-Neutral Different Response (Neutral DR)	1039 (269)	[963,1114]	1041 (248)	[971,1110]	+2 ^{ns/ns*}
Color-Incongruent Same Response (SR)	1107 (270)	[1031,1182]	1071 (271)	[994,1147]	-36 ^{**}
Color-Neutral Same Response (Neutral SR)	1050 (280)	[988,1143]	1065 (276)	[651, 690]	+15 ^{ns/ns*}
	M (SD)	CI	M (SD)	CI	
Standard Stroop Interference (DR – Neutral DR)	83 ^{††/††} (90)	[57,108]	36 ^{†/††} (53)	[21,51]	
Semantic Stroop Interference (SR – Neutral SR)	57 ^{*/††} (110)	[26,87]	5 ^{ns/ns*} (89)	[-19,30]	

Note. Inferential approach: ^{ns}non-significant; ^{*}significant at $p < .05$; ^{**}significant at $p < .01$; [†]significant at $p < .001$; Bayesian approach (presented after the slash): ^{ns}no evidence of an effect, with both BF_{10} and BF_{01} values between 0–1; ^{ns*}anecdotal evidence in favor of H1, but moderate evidence in favor of H0, with BF_{10} value between 0–1, and BF_{01} value between 3–10; ^{*}moderate evidence in favor of H1, with BF_{10} value between 3–10; ^{**}strong evidence in favor of H1, with BF_{10} value between 10–30; [†] very strong evidence in favor of H1, with BF_{10} between 30 –100; ^{††}extreme evidence in favor of H1/of an effect, with $BF_{10} < 100$.

To examine this idea further, magnitudes of both types of interference were further analyzed in a 2 (Interference-Type: Standard vs. Semantic interference) \times 2 (Coloring: ALCC vs. SLCC) ANOVA (see Table 1). This analysis revealed main effects of Interference-Type [$F(1,87) = 7.285$; $p = .008$, $\eta_p^2 = 0.077$; $BF_{10} = 2.207$], and of Coloring [$F(1,87) = 6.066$; $p = .016$, $\eta_p^2 = 0.065$; $BF_{10} = 5.243$]. However Interference-Type \times Coloring interaction remained non-significant [$F(1,87) = 1.925$; $p = .169$, $\eta_p^2 = 0.0322$, $BF_{10} = 0.401$] and $BF_{01} = 2.491$ actually provided anecdotal evidence against this interaction – the lack of which suggests that SLCC reduces both semantic and standard Stroop interference in tandem. Although this latter result show that the performance of older adults in the Stroop task clearly benefits from SLCC manipulation, as we will discuss in further details, the type of conflict that is driving this benefit is still relatively unclear.

4. General Discussion and Conclusion

The primary goal of this study was to investigate the extent to which the ability of selective attention in older adults' benefits from attention driving interventions. As we argued above, the implication of selective attention in daily activities such as driving is paramount. Since driving is closely related to the maintenance of older adults' independence (see e.g., McCloskey & Webb, 2021), improving selective attention in older adults substantially increases their quality of life (see e.g., Owsley, 2002). To address this issue of applied importance, we examined the effects of variations in Response-Stimulus Intervals (RSI) and variations in Coloring and Cueing of Stroop stimuli (SLCC).

Both of these moderators have indeed substantial attention in prior research. However, the type of conflict (semantic vs. response vs. both) that actually benefits from these interventions is still unclear in general population and clearly unknown in older adults. Therefore, the additional goal of this study was to address just this issue of a more theoretical

importance. Because this type of investigation has a necessary pre-requisite that semantic conflict indeed contributes to the overall Stroop interference (i.e., contribution that is not anticipated by currently dominant models of Stroop interference; see e.g. Cohen et al., 1990; Glaser & Glaser, 1989; Logan, 1980; Melara & Algom, 2003; Phaf et al., 1990; Roelofs, 2003), the present study also addressed this more fundamental issue.

In both experiments reported above, semantic conflict contributed significantly to the overall (i.e., standard Stroop interference). In line with Burca and colleagues (2022), in Experiment 1, this contribution was independent of both response conflict and of Stroop facilitation that were both significant in this study. These results also showed that both interference and facilitation contribute to the overall Stroop or *congruency* effect³. However, since color-congruent items are known to boost the overall Stroop effect (Roelofs, 2010) and therefore magnitude of its different constituents, these trials were not included in Experiment 2 (as in Burca et al., 2021). In line with this latter study, and Experiment 1, the contribution of semantic conflict was again significant overall. However, this contribution interacted with the effect of Coloring such that the contribution of semantic conflict was only significant under standard presentation of Stroop stimuli (i.e., in ALCC-condition).

Also, and importantly, as in Burca et al. (2021), the contribution of response conflict remained non-significant in Experiment 2 (see also Šaban, 2021) and taking into account the speed of processing or adding more participants would not solve this issue. Therefore, exactly like the contribution of semantic conflict to the overall Stroop interference challenges the dominant models of Stroop interference (e.g. Cohen et al., 1990a; Glaser & Glaser, 1989; Logan, 1980; Melara & Algom, 2003; Phaf et al., 1990; Roelofs, 2003a), it is also the case of the absence of response conflict in the 2-to-1 Stroop paradigm. Indeed, this pattern of results

³ Therefore, congruency effect observed without any color-neutral baseline does not constitute a suitable proxy for studying cognitive conflict and its control (see e.g., Egner, 2010).

reinforces the idea that a selection in the Stroop task can occur without any response competition. It should be noted however that the only two-stage model of Stroop interference (Zhang et al., 1999; Zhang & Kornblum, 1998) is equally unable to account for this pattern of finding. Therefore again (see also Burca et al., 2021, 2022), the most obvious implication of these results is that more theoretical effort is still needed to account for this polymorphic (i.e., with or without response conflict) nature of the Stroop effect (see Martinon et al., submitted; for discussion).

Concerning the extent to which older adults benefit from attention-driving interventions, Experiment 1 failed to provide any evidence that shortening Response-Stimulus Intervals (RSI) from habitual 2000 ms to 200 ms is beneficial for older adults' performance in the Stroop task – as previously shown by Jackson and Balota (2013). Perhaps the most parsimonious explanation is that shorted RSI failed to increase participants' response speed. Yet, this seems to be the key factor behind the participants' consistent deployment of their inhibitory abilities across time (see De Jong et al., 1999). Due to this lack of effect on response speed, attentional lapses remain even in short RSI – as suggested by the lack of significant Stimulus-Type \times RSI interaction. However, as already mentioned, this lack of RSI-effect is not completely unknown in the literature (see e.g., Martinon et al., Experiment 1 submitted; Sharma et al., 2010 for the absence of this benefit and Galer et al., 2014 for the inverse benefit of (very) long RSI). As a result, this experimental induction might not be as potent as the numerous citations of De Jong and colleagues' study might suggest. Since despite some replications (see e.g., Augustinova et al., 2018a; Martinon et al., Exp. 2; Parris, 2014), our own group also repeatedly failed to replicate the effect of RSI by the part, it is important that journal editors continue to encourage complete reporting of experimental outcomes, including null effects, such that the actual efficiency of this induction can be efficiently assessed.

Contrary to the effect of RSI, performance of older adults in the Stroop task clearly benefited from SLCC (as compared to standard presentation of Stroop stimuli as in ALCC). However, as mentioned above, the type of conflict that is driving this benefit is still relatively unclear. Recall indeed that SLCC completely eliminated the contribution of semantic conflict to Stroop interference – as indicated by the non-significant semantic interference in SLCC-condition. This finding seems therefore perfectly line with Manwell et al.’s Account 2 of the SLCC-effect (see also Labuschagne & Besner, 2015) suggesting that SLCC *eliminates* semantic processing in the Stroop task altogether. Consistent with this latter idea, SLCC (as compared to ALCC) indeed significantly shortened RTs to same-response trials but also those different-response trails. Given that response conflict in this Experiment 2 remained non-significant in both ALCC- and SLCC-condition, it seems reasonable to assume that this latter shortening can reflect the reduction of semantic conflict that different-response trials also involve (in addition to response conflict). Yet, the fact that standard Stroop interference remained significant in SLCC despite a complete elimination of semantic conflict (and non-significant response conflict) runs counter the idea semantic processing (along with response processing) is curtailed altogether. For instance, Heil et al.’ (2004) indeed showed the lack of semantic priming effect in RTs while a significant negative-going ERP around 400 ms (used as an indicator of semantic activation) was detected at the same time. Thus, as already argued by Neely and Kahan (2001; see also Augustinova et al., 2010, Augustinova & Ferrand, 2014b), an apparent elimination of semantic Stroop interference does not necessarily mean that semantic activation is curtailed.

It should be noted however that this remaining standard Stroop interference in SLCC-condition can still be driven by a more general conflict (i.e., task conflict) that – for all readable Stroop items including color-neutral ones – derives from the simultaneous preparation of two task sets: word-reading vs. color-naming, e.g., Goldfarb & Henik, 2007;

Hershman & Henik, 2019). This idea is still contradicted by the lack of SLCC-effect on both types of color-neutral trials (but see Burca et al., 2021 for discussion of this finding that the present study further replicates). In sum, in line with Manwell et al.'s reasoning (and clearly counter to Augustinova et al.'s work), the effect of SLCC is largely driven by the reduction of semantic conflict, even though the reduction of response and task conflicts by this intervention cannot be ruled out in the present study. Therefore, future studies need to address this issue while isolation the contribution of all these conflicts to Stroop interference. If anything, standard Stroop interference that remained significant in the SLCC-condition (unlike that semantic Stroop interference) is likely to suggest that the Stroop effect is indeed composite phenomenon – resulting from the joint contribution of all of these conflicts along with their perhaps complex interactions. As above, we are likely to attribute this remaining interference to the sum of these different conflicts – each in unknown and separately non-significant amounts. Said differently, we are tempted to conclude that a task, semantic and response processing still occur in SLCC, simply not in amounts that would allow to capture their specific contributions in RTs. This reasoning is indeed reinforced by the idea that ERPs are sensitive to SLCC procedure early (from 185 to 245 ms) in processing of Stroop words and at this stage independently of color-incongruency that is processed later (see also Navarette et al., 2015). In sum, is it possible that better attentional focus on the relevant dimension that SLCC induces (as compared to ALCC), reduces significantly the early (Hershman & Henik, 2019, 2020) task conflict. As a result of this reduction of the irrelevant (i.e., word-reading) task-set (see also Kinoshita et al., 2017), word-information of color-incongruent Stroop words interferes less with the relevant information (i.e., color to be named). Said differently, it interferes less not because it is better inhibited but because it activated less strongly in SLCC (as compared to ALCC). Again, future research deploying time-course sensitive indicators such as ERPs (e.g., Augustinova et al., 2015; Navarette et al.,

2015) or mouse-trajectories (see e.g., Bunt et al., 2015; Quétard et al., submitted) is needed to address these competing possibilities thoroughly. Indeed, the most obvious limitation of the present study is that this type of processing – taking place before the actual output – can only be inferred from RTs. In sum, more research is still needed to account for both loci of the Stroop effect, but also changes in these loci as a function of attention-driving moderators.

To conclude, Experiment 2 successfully replicated results reported by Burca et al. (2021) in older adults. Also, and importantly, the results of this replication are encouraging since they suggest that older adults actually benefit from SLCC intervention more than do younger adults of Burca and colleagues' study. Indeed, in this latter study, semantic Stroop interference was still supported by strong Bayesian evidence in SLCC-condition, whereas in the present study (i.e., study conducted in older adults) showed moderate Bayesian evidence against this semantic interference. Also, and importantly, this latter reduction needs to be seen in light of higher baseline levels of this type of interference in older adults (see Burca et al., 2022)⁴. Therefore, the results of the present study clearly indicate that the cognitive functioning of the elderly can be substantially improved through more effect design of different types of devices used for both work and leisure purposes.

5. References

- Andrés, P., Guerrini, C., Phillips, L. H., & Perfect, T. J. (2008). Differential Effects of Aging on Executive and Automatic Inhibition. *Developmental Neuropsychology*, 33(2), 101-123.
<https://doi.org/10.1080/87565640701884212>
- Aschenbrenner, A. J., & Balota, D. A. (2015). Interactive effects of working memory and trial history on Stroop interference in cognitively healthy aging. *Psychology and Aging*, 30(1), 1–8.
<https://doi.org/10.1037/pag0000012>

⁴Still future studies need to control for age orthogonally to avoid comparisons between independent studies.

- Aschenbrenner, A. J., Balota, D. A., Weigand, A. J., Scaltritti, M., & Besner, D. (2017). The first letter position effect in visual word recognition: The role of spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, 43(4), 700–718. <https://doi.org/10.1037/xhp0000342>
- Augustinova, M., & Ferrand, L. (2007). Influence de la présentation bicolore des mots sur l'effet Stroop. *L'Année Psychologique*, 107(02), 163. <https://doi.org/10.4074/S0003503307002011>
- Augustinova, M., & Ferrand, L. (2014a). Social priming of dyslexia and reduction of the Stroop effect: What component of the Stroop effect is actually reduced? *Cognition*, 130, 442–454. <https://doi.org/10.1016/j.cognition.2013.11.014>
- Augustinova, M., & Ferrand, L. (2014b). 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>
- Augustinova, M., Flaudias, V., & Ferrand, L. (2010). Single-letter coloring and spatial cuing do not eliminate or reduce a semantic contribution to the Stroop effect. *Psychonomic Bulletin & Review*, 17(6), 827–833. <https://doi.org/10.3758/PBR.17.6.827>
- Augustinova, M., Silvert, L., Ferrand, L., Llorca, P. M., & Flaudias, V. (2015). Behavioral and electrophysiological investigation of semantic and response conflict in the Stroop task. *Psychonomic Bulletin & Review*, 22(2), 543–549. <https://doi.org/10.3758/s13423-014-0697-z>
- Augustinova, M., Silvert, L., Spatola, N., & Ferrand, L. (2018a). Further investigation of distinct components of Stroop interference and of their reduction by short response-stimulus intervals. *Acta Psychologica*, 189, 54–62. <https://doi.org/10.1016/J.ACTPSY.2017.03.009>
- Augustinova, M., Clarys, D., Spatola, N., & Ferrand, L. (2018b). Some further clarifications on age-related differences in Stroop interference. *Psychonomic Bulletin & Review*, 25(2), 767–774. <https://doi.org/10.3758/s13423-017-1427-0>
- Belleville, S., Rouleau, N., & Van der Linden, M. (2006). Use of the Hayling task to measure inhibition of prepotent responses in normal aging and Alzheimer's disease. *Brain and cognition*, 62(2), 113–119. <https://doi.org/10.1016/j.bandc.2006.04.006>
- Besner, D., & Stolz, J. A. (1999a). 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. (1999b). Unconsciously controlled processing: The Stroop effect reconsidered. *Psychonomic Bulletin & Review*, 6(3), 449–455. <https://doi.org/10.3758/BF03210834>

- Besner, D., Stolz, J. A., & Boutilier, C. (1997). The Stroop effect and the myth of automaticity. *Psychonomic bulletin & review*, 4(2), 221-225. <https://doi.org/10.3758/BF03209396>
- Brown, T. L., Joneleit, K., Robinson, C. S., & Brown, C. R. (2002). Automaticity in reading and the Stroop task: testing the limits of involuntary word processing. *The American Journal of Psychology*. <https://doi.org/10.2307/1423526>
- Bugaiska, A., Clarys, D., Jarry, C., Taconnat, L., Tapia, G., Vanneste, S., & Isingrini, M. (2007). The effect of aging in recollective experience: The processing speed and executive functioning hypothesis. *Consciousness and Cognition*, 16(4), 797-808. doi:10.1016/j.concog.2006.11.007
- Bugg, J. M., DeLosh, E. L., Davalos, D. B., & Davis, H. P. (2007). Age Differences in Stroop Interference: Contributions of General Slowing and Task-Specific Deficits. *Aging, Neuropsychology, and Cognition*, 14(2), 155–167. <https://doi.org/10.1080/138255891007065>
- Burca, M., Beaucousin, V., Chausse, P., Ferrand, L., Parris, B. A., & Augustinova, M. (2021). Is There Semantic Conflict in the Stroop Task? *Experimental Psychology*, 68(5), 274–283. <https://doi.org/10.1027/1618-3169/a000530>
- Burca, M., Chausse, P., Ferrand, L., Parris, B. A., & Augustinova, M. (2022). Some further clarifications on age-related differences in the Stroop task: New evidence from the two-to-one Stroop paradigm. *Psychonomic Bulletin & Review*, 29,. <https://doi.org/10.3758/s13423-021-02011-x>
- Catena, A., Fuentes, L. J., & Tudela, P. (2002). Priming and interference effects can be dissociated in the Stroop task: New evidence in favor of the automaticity of word recognition. *Psychonomic Bulletin & Review*, 9(1), 113-118. <https://doi.org/10.3758/BF03196265>
- 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>
- Comalli, P. E., Wapner, S., & Werner, H. (1962). Interference Effects of Stroop Color-Word Test in Childhood, Adulthood, and Aging. *The Journal of Genetic Psychology*, 100(1), 47–53. <https://doi.org/10.1080/00221325.1962.10533572>
- Davidson, D. J., Zacks, R. T., & Williams, C. C. (2003). Stroop interference, practice, and aging. *Aging, Neuropsychology, and Cognition*, 10(2), 85-98.

- De Jong, R., Berendsen, E., & Cools, R. (1999). Goal neglect and inhibitory limitations: Dissociable causes of interference effects in conflict situations. *Acta Psychologica*, 101(2–3), 379–394.
[https://doi.org/10.1016/S0001-6918\(99\)00012-8](https://doi.org/10.1016/S0001-6918(99)00012-8)
- Dulaney, C. L., & Rogers, W. A. (1994). Mechanisms underlying reduction in Stroop interference with practice for young and old adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(2), 470. <https://doi.org/10.1037/0278-7393.20.2.470>
- Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information-processing rate and amount: implications for group differences in response latency. *Psychological bulletin*, 125(6), 777. <https://doi.org/10.1037/0033-2909.125.6.777>
- Galer, S., Schmitz, R., Leproult, R., De Tiège, X., Van Bogaert, P., & Peigneux, P. (2014). Response-stimulus interval duration modulates interference effects in the Stroop task. *Psychologica Belgica*, 54(1).
<http://doi.org/10.5334/pb.ad>
- Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General*, 118(1), 13–42.
- Goldfarb, L., & Henik, A. (2007). Evidence for task conflict in the Stroop effect. *Journal of Experimental Psychology: Human Perception and Performance*, 33(5), 1170–1176. <https://doi.org/10.1037/0096-1523.33.5.1170>
- Hasshim, N., & Parris, B. A. (2014). Two-to-one color-response mapping and the presence of semantic conflict in the Stroop task. *Frontiers in Psychology*, 5, 1157. <https://doi.org/10.3389/fpsyg.2014.01157>
- Hasshim, N., & Parris, B. A. (2015). Assessing stimulus–stimulus (semantic) conflict in the Stroop task using saccadic two-to-one color response mapping and prereponse pupillary measures. *Attention, Perception, & Psychophysics*, 77(8), 2601–2610. <https://doi.org/10.3758/s13414-015-0971-9>
- Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior research methods*, 50(3), 1166–1186.
<https://doi.org/10.3758/s13428-017-0935-1>
- Heil, M., Rolke, B., & Pecchinenda, A. (2004). Automatic semantic activation is no myth: Semantic context effects on the N400 in the letter-search task in the absence of response time effects. *Psychological Science*, 15(12), 852–857.

- Hershman, R., & Henik, A. (2019). Dissociation between reaction time and pupil dilation in the Stroop task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(10), 1899–1909.
<https://doi.org/10.1037/xlm0000690>
- Hershman, R., & Henik, A. (2020). Pupillometric contributions to deciphering Stroop conflicts. *Memory & Cognition*, 48(2), 325–333. <https://doi.org/10.3758/s13421-019-00971-z>
- Jackson, J. D., & Balota, D. A. (2013). Age-related changes in attentional selection: Quality of task set or degradation of task set across time? *Psychology and Aging*, 28(3), 744–753.
<https://doi.org/10.1037/a0033159>
- Kahneman, D. & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy, J. R Pomerantz (Eds.), *Perceptual organization* (pp. 181-211). Hillsdale, NJ: Erlbaum.
- Kinoshita, S., De Wit, B., & Norris, D. (2017). The magic of words reconsidered: Investigating the automaticity of reading color-neutral words in the Stroop task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(3), 369–384. <https://doi.org/10.1037/xlm0000311>
- Küper, K., & Heil, M. (2012). Attentional Focus Manipulations Affect Naming Latencies of Neutral But Not of Incongruent Stroop Trials. *Swiss Journal of Psychology*, 71(2), 93-100. <https://doi.org/10.1024/1421-0185/a000075>
- Labuschagne, E. M., & Besner, D. (2015). Automaticity revisited: When print doesn't activate semantics. *Frontiers in Psychology*, 6, 117. <https://doi.org/10.3389/fpsyg.2015.00117>
- Lemaire, P., & Bherer, L. (2005). *Psychologie du vieillissement: Une perspective cognitive*. De Boeck Supérieur: Bruxelles.
- Li, K. Z. H., & Bosman, E. A. (1996). Age differences in stroop-like interference as a function of semantic relatedness. *Aging, Neuropsychology, and Cognition*, 3(4), 272–284.
<https://doi.org/10.1080/13825589608256630>
- Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology*, 12(4), 523–553. [https://doi.org/10.1016/0010-0285\(80\)90019-5](https://doi.org/10.1016/0010-0285(80)90019-5)
- 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>
- Maquestiaux, F. (2017). *Psychologie de l'attention*. De Boeck Supérieur: Bruxelles.

- Mari-Beffa, P., Fuentes, L. J., Catena, A., & Houghton, G. (2000). Semantic priming in the prime task effect: Evidence of automatic semantic processing of distractors. *Memory & Cognition*, 28(4), 635-647.
<https://doi.org/10.3758/BF03201253>
- Marmurek, H. H. C. (2003). Coloring only a single letter does not eliminate color-word interference in a vocal-response Stroop task: Automaticity revealed. *Journal of General Psychology*, 130, 207-224.
- Monahan, J. S. (2001). Coloring single Stroop elements: Reducing automaticity or slowing color processing? *Journal of General Psychology*, 128, 98-112. <https://doi.org/10.1080/00221300109598901>
- Martinon L., Ferrand L., Burca M., Hasshim N., Lakhzoum D., Parris B.A, Silvert L., Augustinova M. (submitted). Variations in Response Speed reveal the Polymorphic Nature of the Stroop Effect. *Journal of Experimental Psychologie: Learning, Memory and Cognition*.
- Mayas, J., Fuentes, L. J., & Ballesteros, S. (2012). Stroop interference and negative priming suppression in normal aging. *Archives of Gerontology and Geriatrics*, 54, 333-338.
<https://doi.org/10.1016/j.archger.2010.12.012>
- McCloskey, K., & Webb, M. (2021). Driving in older adults: A literature review of interventions to promote safety and independence. *New Zealand Journal of Occupational Therapy*, 68(2), 4-10.
- Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review*, 110(3), 422-471. <https://doi.org/10.1037/0033-295X.110.3.422>
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49-100. <https://doi.org/10.1006/cogp.1999.0734>
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695-699.
<https://doi.org/10.1111/j.1532-5415.2005.53221.x>
- Navarrete, E., Sessa, P., Peressotti, F., & Dell’Acqua, R. (2015). The distractor frequency effect in the colour-naming Stroop task: An overt naming event-related potential study. *Journal of Cognitive Psychology*, 27(3), 277-289. <https://doi.org/10.1080/20445911.2014.1002786>
- Neely, J. H., & 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>
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, 36(3), 516–524. <https://doi.org/10.3758/BF03195598>
- Nicosia, J., & Balota, D. (2020). The consequences of processing goal-irrelevant information during the Stroop task. *Psychology and Aging*, 35(5), 663–675. <https://doi.org/10.1037/pag0000371>
- Owsley, C., Sloane, M., McGwin Jr, G., & Ball, K. (2002). Timed instrumental activities of daily living tasks: relationship to cognitive function and everyday performance assessments in older adults. *Gerontology*, 48(4), 254–265. <https://doi.org/10.1159/000058360>
- Parris, B. A. (2014). Task conflict in the Stroop task: When Stroop interference decreases as Stroop facilitation increases in a low task conflict context. *Frontiers in Psychology*, 5, 1182. <https://doi.org/10.3389/fpsyg.2014.01182>
- Parris, B. A., Sharma, D., & Weekes, B. (2007). An optimal viewing position effect in the Stroop task when only one letter is the color carrier. *Experimental Psychology*, 54, 273–280. <https://doi.org/10.1027/1618-3169.54.4.273>
- Parris, B. A., Hasshim, N., Wadsley, M., Augustinova, M., & Ferrand, L. (2021, online first). The loci of Stroop effects: A critical review of methods and evidence for levels of processing contributing to color-word Stroop effects and the implications for the loci of attentional selection. *Psychological Research*. <https://doi.org/10.1007/s00426-021-01554-x>
- Phaf, R. H., Van der Heijden, A. H. C., & Hudson, P. T. W. (1990). SLAM: A connectionist model for attention in visual selection tasks. *Cognitive Psychology*, 22(3), 273–341. [https://doi.org/10.1016/0010-0285\(90\)90006-P](https://doi.org/10.1016/0010-0285(90)90006-P)
- Pollatsek, A., Romoser, M. R., & Fisher, D. L. (2012). Identifying and remediating failures of selective attention in older drivers. *Current Directions in Psychological science*, 21(1), 3–7. 10.1177/0963721411429459
- Quétard, B., Spatola, N., Parris, B. A., Ferrand, L., & Augustinova, M. (submitted). A mouse-tracking study of the composite nature of the Stroop effect at the level of response execution. *Quarterly Journal of Experimental Psychology*.
- Raven, J. C. (1958). Guide to using the Mill Hill Vocabulary Scale with the Progressive Matrices Scales.
- Rey-Mermet, A., & Gade, M. (2018). Inhibition in aging: What is preserved? What declines? A meta-analysis. *Psychonomic Bulletin & Review*, 25(5), 1695–1716. <https://doi.org/10.3758/s13423-017-1384-7>

- Roelofs, A. (2003). Goal-Referenced Selection of Verbal Action: Modeling Attentional Control in the Stroop Task. *Psychological Review*, 110(1), 88–125. <https://doi.org/10.1037/0033-295X.110.1.88>
- Roelofs, A. (2010). Attention and facilitation: converging information versus inadvertent reading in Stroop task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(2), 411. <https://doi.org/10.1037/0033-295X.110.1.88>
- Ruthruff, E., & Lien, M. C. (2016). Aging and attention. *Encyclopedia of Geropsychology*, 1-7.
- Šaban, I. (2021). Basic processes in interference paradigms. Defended PhD dissertation at Université de Bourgogne Franche-Comté.
- Šaban, I., & Schmidt, J. R. (2021). Stimulus and response conflict from a second language: Stroop interference in weakly-bilingual and recently-trained languages. *Acta Psychologica*, 218, 103360. <https://doi.org/10.1016/j.actpsy.2021.103360>
- 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>
- Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human perception and performance*, 22(2), 461-479. <https://doi.org/10.1037/0096-1523.22.2.461>
- 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>
- Wechsler, D., Psychological Corporation., & PsychCorp (Firm). (2008). WAIS-IV technical and interpretive manual. San Antonio, Texas: Pearson.
- Wolf, D., Zschuschke, L., Scheurich, A., Schmitz, F., Lieb, K., Tüscher, O., & Fellgiebel, A. (2014). Age-related increases in stroop interference: Delineation of general slowing based on behavioral and white matter analyses. *Human brain mapping*, 35(5), 2448-2458. <https://doi.org/10.1002/hbm.22340>
- Schneider W., Eschman A., Zuccolotto A. (2002). E-Prime user's guide. Pittsburgh, USA: Psychology Software Tools Inc.
- Zhang, H., & Kornblum, S. (1998). The Effects of Stimulus-Response Mapping and Irrelevant Stimulus-Response and Stimulus-Stimulus Overlap in Four-Choice Stroop Tasks with Single-Carrier Stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 24(1), 3–19. <https://doi.org/10.1037/0096-1523.24.1.3>

Zhang, H., Zhang, J., & Kornblum, S. (1999). A Parallel Distributed Processing Model of Stimulus–Stimulus and Stimulus–Response Compatibility. *Cognitive Psychology*, 38(3), 386–432.

<https://doi.org/10.1006/COGP.1998.0703>

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, 67(6), 361-370. <https://doi.org/10.1111/j.1600-0447.1983.tb09716.x>

6. Supplementary Materials

Table 1S. *Demographic Variables and Mean Psychometric Test Scores of Participants taking part in Experiment 1*

	N=50		
	M	SD	CI
Age	70.01	5.60	[68.53, 71.49]
Education (years)	10.96	3.46	[10.64, 11.27]
MOCA	27.09	1.85	[26.92, 27.25]
MILL-HILL	26.01	4.31	[25.62, 26.41]
HADS	11.98	4.94	[11.53, 12.43]
Processing speed	23.70	5.84	[23.17, 24.24]

Table 2S. Color-Identification Performance (Means, Standard Deviation, 95%-Confidence Intervals and Error Percentages) observed as a Function of Stimulus- Interference- or Conflict Type and Response-Stimulus-Interval of Participants taking part in Experiment 1

	Response-Stimulus-Interval							
	Long RSI			Short RSI				
	M (SD)	CI	%ER	M (SD)	CI	%ER	RSI effect (RT)	RSI effect (ER)
Different Response (DR) Color-Incongruent items	917 (340)	[813, 1021]	5.81	930 (195)	[816, 1043]	3.10	+13 ^{ns/ns}	- 2.71 ^{ns/ns}
Same Response (SR) Color-Incongruent items	873 (310)	[779, 968]	2.50	896 (173)	[793, 1000]	1.40	+23 ^{ns/ns}	- 1.09 ^{ns/ns}
Color-Neutral items	827 (266)	[747, 909]	6.48	842 (148)	[754, 931]	5.80	+15 ^{ns/ns}	- .679 ^{ns/ns}
Color-Congruent items	804 (268)	[723, 885]	2.41	818 (144)	[730, 908]	1.88	+14 ^{ns/ns}	- .525 ^{ns/ns}

Note. Inferential approach: ^{ns}non-significant; *significant at $p < .05$; **significant at $p < .01$; †significant at $p < .001$; Bayesian approach (presented after the slash): ^{ns}no evidence of an effect, with both BF_{10} and BF_{01} values between 0–1; ^manecdotal evidence in favor of H1, but moderate evidence in favor of H0, with BF_{10} value between 0–1, and BF_{01} value between 3–10; *moderate evidence in favor of H1, with BF_{10} value between 3–10; **strong evidence in favor of H1, with BF_{10} value between 10–30; † very strong evidence in favor of H1, with BF_{10} between 30–100; ††extreme evidence in favor of H1/of an effect, with $BF_{10} < 100$.

Table 3S. *Color-Identification Performance (Means, Standard Error, 95%-Confidence Intervals) observed as a Function of Stimulus-Type and of its decompositions (Experiment 1)*

	M (SE)	CI
Overall Stroop Interference (DR – Congruent)	112 ^{†/††} (8.47)	[15, 62]
Facilitation Effect (Neutral - Congruent)	24 ^{**/††} (5.83)	[8, 40]
Standard Stroop Interference (DR – Neutral)	88 ^{†/††} (13)	[53, 124]
Response Conflict (DR – SR)	39 ^{†/††} (8)	[15, 62]
Semantic Stroop Interference (SR – Neutral)	50 ^{*/††} (8)	[27, 73]

Note. Inferential approach: ^{ns}non-significant; *significant at $p < .05$; **significant at $p < .01$; [†]significant at $p < .001$; Bayesian approach (presented after the slash): ^{ns}no evidence of an effect, with both BF_{10} and BF_{01} values between 0–1; ^{ns*}anecdotal evidence in favor of H1, but moderate evidence in favor of H0, with BF_{10} value between 0–1, and BF_{01} value between 3–10; *moderate evidence in favor of H1, with BF_{10} value between 3–10; **strong evidence in favor of H1, with BF_{10} value between 10–30; [†] very strong evidence in favor of H1, with BF_{10} between 30 -100; ^{††}extreme evidence in favor of H1/of an effect, with $BF_{10} < 100$.

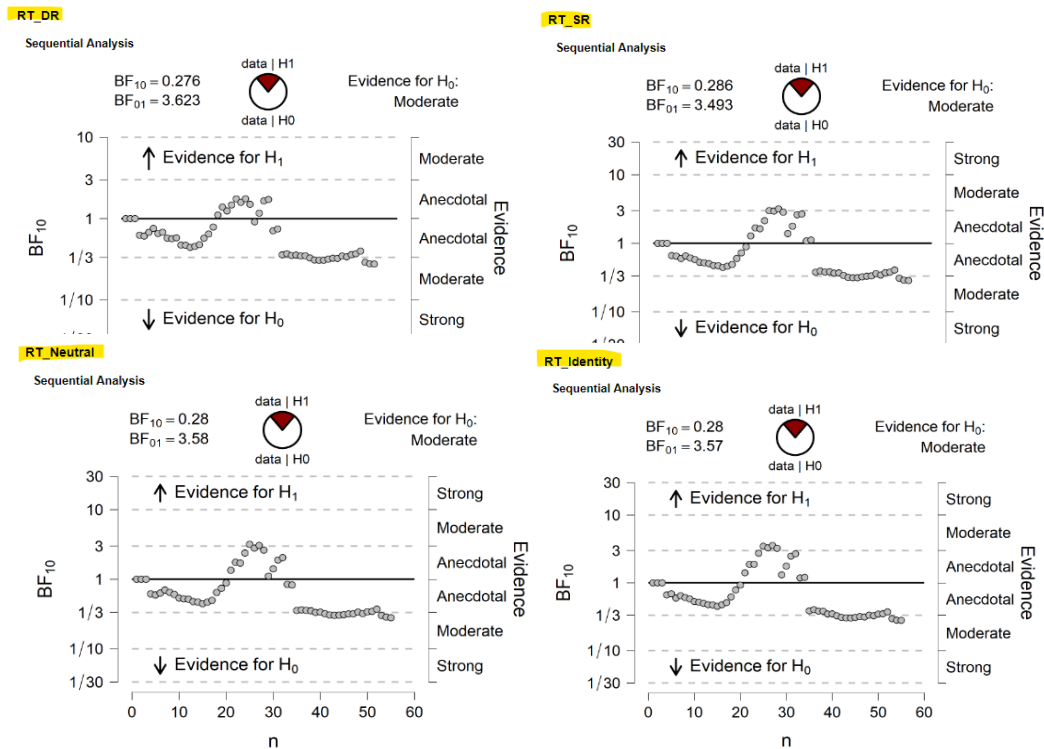


Figure 1S. *Sequential analysis for all Stimuli-Types with RSI as a grouping variable (probability of finding an effect of RSI if more participants are added in Experiment 1)*

Analysis on slow vs. fast participants (Experiment 1)

The possibility that only the slow (vs. fast) participants benefited from the RSI moderator was examined. To this end, an omnibus 4 (Stimulus-Type: different-response vs. same-response vs. neutral vs. congruent) \times 2 (RSI: slow vs. fast) \times 2 (Speed: fast vs. slow participants) was conducted. The main effect of Stimulus-Type was significant [$F(3,153) = 57.17$; $p < .001$, $\eta_p^2 = 0.529$; $BF_{10} = 1.737e+21$]. The main effect of RSI was not significant in this analysis [$F(1,51) = .749$; $p = .391$, $\eta_p^2 = 0.014$; $BF_{10} = 0.573/BF_{01} = 1.744$]. However, the main effect of speed revealed to be significant [$F(1,51) = 59.12$; $p < .001$, $\eta_p^2 = 0.537$; $BF_{10} = 1.653e+7$], showing that indeed the fast participants had shorter RT's overall. However, the interaction RSI \times Speed did not reach significance [$F(1,51) = 2.53$; $p = .118$, $\eta_p^2 = 0.047$; $BF_{10} = 1.103/BF_{01} = 0.906$], and neither did the 3-way interaction Stimulus-Type \times RSI \times Speed [$F(3,153) = .144$; $p = .934$, $\eta_p^2 = 0.003$; $BF_{10} = 1.106/BF_{01} = 0.903$].

Only the Stimulus \times Speed of participants was significant [$F(3,153) = 9.19$; $p < .001$, $\eta_p^2 = 0.153$; $BF_{10} = 1467.88$]. Its decomposition (see Table 4S bellow) showed, that all the fast participants had significantly shorter response times as compared to the slow participants (all p 's $< .001$). Moreover, this decomposition revealed, that both fast and slow participants displayed a significant overall Stroop effect (both p 's $< .001$). However, in fast participants, this latter effect did not involve facilitation ($p = .288$), whereas it was the case if slow participants ($p = .008$). The overall Stroop interference was significant for both the fast ($p = .047$) and slow participants ($p < .001$). Yet in fast participants neither the response conflict nor the semantic conflict reached significant contribution to this effect (both p 's $> .05$) and this despite quite substantial magnitudes (see here bellow Table 4S), while in slow participants, both a significant response and semantic conflicts (both p 's $< .001$) contributed significantly (see Table 4S for descriptive statistics). These results suggest, that as in Chapter 3, the speed

of processing (or responding) does influence the components of the Stroop effect and their magnitudes.

Table 4S. *Color-Identification Performance (Means, Standard Error, 95%-Confidence Intervals) observed as a Function of Stimulus \times Speed interaction (Experiment 1)*

	Fast participants		Slow participants		Speed effect
	M (SE)	CI	M (SE)	CI	
Different Response (DR) Color-Incongruent items	724 (40)	[644,804]	1121 (38)	[1045,1197]	-397**
Same Response (SR) Color-Incongruent items	703 (35)	[632,773]	1067 (34)	[1000,1135]	-364**
Color-Neutral items	675 (28)	[616,733]	996 (28)	[941,1051]	-321**
Color-Congruent items	657 (38)	[597,717]	968 (28)	[911,1025]	-311**
	M (SD)	CI	M (SD)	CI	
Overall Stroop Interference (DR – Congruent)	67** (16)	[22,112]	153** (12)	[110,196]	
Facilitation Effect (Neutral – Congruent)	18 ^{ns} (9)	[-288, -6]	28** (8)	[5,51]	
Standard Stroop interference (DR – Neutral)	50* (18)	[-12,55]	125** (17)	[78,171]	
Semantic Stroop Interference (SR – Neutral)	28 ^{ns} (12)	[-4,61]	71** (11)	[40,102]	
Response Conflict (DR-SR)	21 ^{ns} (12)	[-12,55]	53** (12)	[21,85]	

Note. Inferential approach: ^{ns}non-significant; * significant at $p < .05$; ** significant at $p < .001$;

Analysis on fast vs. slow trials (Experiment 1)

An omnibus 4 (Stimulus-Type: different-response vs. same-response vs. neutral vs. congruent) \times 2 (RSI: slow vs. fast) \times 2 (Bins: fast vs. slow trials) was conducted. Again, as expected, the main effect of RSI did not reach significance [$F(1,53) = .382$; $p = .539$, $\eta_p^2 = 0.007$; $BF_{10} = 0.305$] and this this analysis actually provided moderated evidence against this effect, $BF_{01} =$

3.281. Both the main effect of Stimulus-Type [$F(3,159) = 59.74$ $p < .001$, $\eta_p^2 = 0.530$; $BF_{10} = 4.709/BF_{01} = 0.212$], and that of Bin was significant [$F(1,53) = 227$; $p < .001$, $\eta_p^2 = 0.811$; $BF_{10} = 4.492e+110$]. The Bin \times RSI interaction remained not significant [$F(1,53) = .434$; $p = .513$, $\eta_p^2 = 0.008$; $BF_{10} = 0.421/BF_{01} = 2.372$] along with the 3-way interaction Stimulus-Type \times RSI \times Bin [$F(3,159) = 2.34$; $p = .076$, $\eta_p^2 = 0.042$; $BF_{10} = 0.099/BF_{01} = 10.05$].

However, similar to the analysis reported above on fast vs. slow participants the Stimulus \times Bin interaction was significant [$F(3,159) = 47.43$; $p < .001$, $\eta_p^2 = 0.472$; $BF_{10} = 7.832$]. The decomposition of this interaction (see Table 5S for descriptive statistics) showed that in both fast and slow trials all Stroop phenomena were significant. Indeed, in both fast and slow trials a significant Stroop effect results from both significant facilitation and interference. Moreover, in both fast and slow trials, this latter phenomenon resulted from the significant contribution of both a semantic and response conflicts (all p 's $< .05$). These results suggest that variations in speed does not affect the actual components of the Stroop effect, it influences however their magnitudes.

Table 5S. *Color-Identification Performance (Means, Standard Error, 95%-Confidence Intervals) observed as a Function of Stimulus \times Bin interaction (Experiment 1)*

	Fast Bin		Slow Bin		Speed effect
	M (SE)	CI	M (SE)	CI	
Different Response (DR) Color-Incongruent items	703 (25)	[653, 753]	1173 (52)	[1069,1278]	- 470**
Same Response (SR) Color-Incongruent items	682 (24)	[635, 730]	1119 (48)	[1022,1215]	- 437**
Color-Neutral items	658 (20)	[619,697]	1040 (42)	[955,1124]	-382**
Color-Congruent items	641 (20)	[601,682]	1004 (40)	[922,1084]	-363**
	M (SD)	CI	M (SD)	CI	

Overall Stroop Interference (DR – Congruent)	62** (5)	[7,34]	170** (17)	[125,215]
Facilitation Effect (Neutral – Congruent)	17** (3)	[10,24]	36** (11)	[7,65]
Standard Stroop interference (DR – Neutral)	45** (9)	[20,70]	134** (16)	[90,178]
Semantic Stroop Interference (SR – Neutral)	41** (6)	[25,57]	115** (14)	[78,153]
Response Conflict (DR - SR)	21** (5)	[7,34]	55** (12)	[22,88]

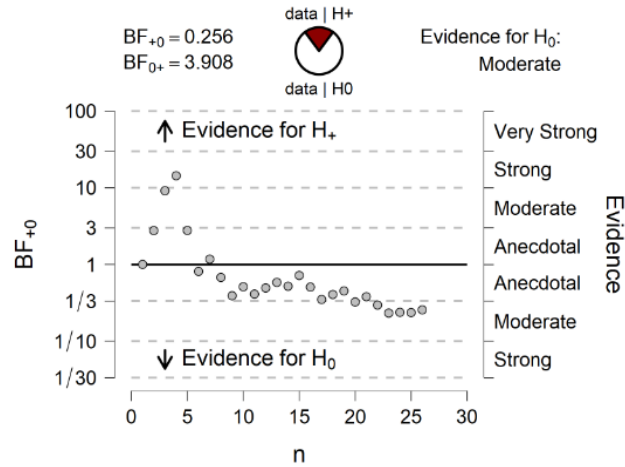
Note. Inferential approach: ^{ns}non-significant; * significant at $p < .05$; ** significant at $p < .001$;

Table 6S. *Demographic Variables and Mean Psychometric Test Scores of Participants taking part in Experiment 2*

	N=50		
	M	SD	CI
Age	68.54	5.75	[68.01, 69.09]
Education (years)	10.92	3.50	[10.59, 11.25]
MOCA	29.01	1.06	[28.91, 29.11]
MILL-HILL	25.25	5.10	[24.77, 25.73]
HADS	11.92	4.46	[11.50, 12.34]
Processing speed	24.19	5.24	[23.70, 24.69]

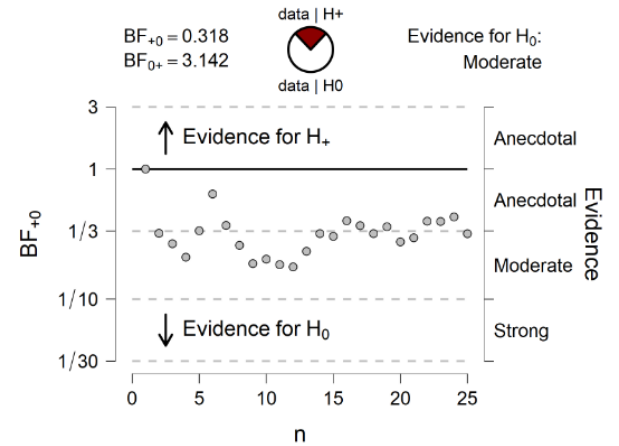
slow_SLCC_DR - slow_SLCC_SR

Sequential Analysis



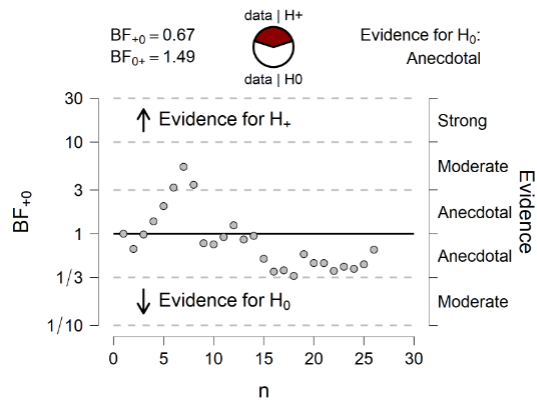
fast_SLCC_DR - fast_SLCC_SR

Sequential Analysis



slow_ALCC_DR - slow_ALCC_SR

Sequential Analysis



fast_ALCC_DR - fast_ALCC_SR

Sequential Analysis

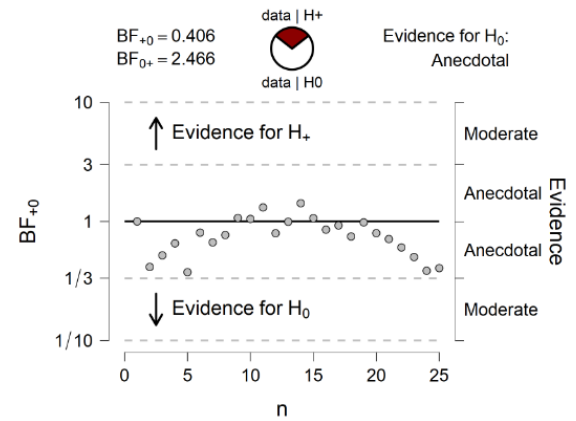


Figure 2S. Sequential analysis contrasting DR incongruent trials and SR incongruent trials for both slow and fast participants (probability of finding a response conflict if more participants are added in Experiment 2). if more participants are added in Experiment 2).

GENERAL DISCUSSION

In order to contribute to the ongoing debate about the locus as opposed to loci of the Stroop interference effect, the present dissertation was aimed at assessing whether semantic conflict (1) contributes to the Stroop interference effect observed in the two-to-one Stroop paradigm (De Houwer, 2003); (2) is spared by various experimental interventions (Augustinova & Ferrand, 2014b) and age-related differences (Augustinova, Clarys et al., 2018; Li & Bosman, 1996).

To this end, research depicted in **Chapter 1** addressed current controversies concerning semantic conflict and its reduction by single-letter coloring and cueing-procedure (SLCC). On the first issue Chapter 1 provided, for the first time, unambiguous evidence for a contribution of semantic conflict to the (overall) Stroop interference effect. The reported data remained inconclusive on the second issue, despite being collected in a considerable sample and analyzed with both classical (frequentist) and Bayesian inferential approaches. Given that in all past Stroop studies, *semantic* conflict was possibly confounded with either *response conflict* (e.g., when semantic-associative items [*SKY*_{blue}] are used to induce semantic conflict) or with *facilitation* (when color-congruent items [*BLUE*_{blue}] are used as baseline to derive a magnitude for semantic conflict), its genuine contribution to the Stroop interference effect is the most critical result of the study reported in this chapter. Indeed, it leaves no doubt – in complete contrast to dominant single-stage response competition models (e.g., Roelofs, 2003) – that selection occurs at the semantic level in the Stroop task. The immediate implications for the composite (as opposed to unitary) nature of the Stroop interference effect and other still unresolved issues in the Stroop literature were also outlined in this chapter (published as Burca et al., 2021).

Previous studies (Augustinova et al., 2018b; Li & Bosman, 1996) have shown that the larger Stroop effects reported in older adults is specifically due to age-related differences in

the magnitude of response, and not semantic conflict, both of which are thought to contribute to overall Stroop interference. However, the most recent contribution to the issue of the unitary vs. composite nature of the Stroop effect (Parris et al., 2021) argues that semantic conflict has not been clearly dissociated from response conflict in these or any other past Stroop studies, meaning that the very existence of semantic conflict is at present uncertain. To distinguish clearly between the two types of conflicts, the study conveyed in **Chapter 2** employed the two-to-one Stroop paradigm (De Houwer, 2003) with a color-neutral word baseline (Hasshim & Parris, 2014). This addition made it possible to isolate a contribution of semantic conflict that was independent of both response conflict and Stroop facilitation. Therefore, this study provides the first unambiguous empirical demonstration of the composite nature of Stroop interference – as originally claimed by multi-stage models of Stroop interference (Zhang & Kornblum, 1998; Zhang et al., 1999). This permitted the further observation of significantly higher levels of semantic conflict in older adults, whereas the level of response conflict in the present study remained unaffected by healthy aging – a finding that directly contrasts with previous studies employing alternative measures of response and semantic conflict. Qualitatively different explanations of this apparent divergence across studies were also discussed in this chapter (published as Burca et al., 2022), such as the response modality, the methods used to control for age-related slowing, and as well the employed paradigms.

Research depicted in **Chapter 3** addressed current controversies concerning the contribution of semantic conflict to the overall Stroop effect further by combining the two-to-one Stroop paradigm (De Houwer, 2003) – administered with an additional color-neutral word baseline (Hasshim & Parris, 2014) – with variations in response-stimulus intervals (RSI, De Jong et al., 1999). The study depicted in this chapter provided additional evidence in favor of multi-stage models of the Stroop effect (Zhang & Kornblum, 1998; Zhang et al., 1999).

Indeed, in the conventional (i.e., long) RSI of 2000ms, and as in Burca et al. (2022, see Chapter 2), the contribution of semantic conflict to the overall Stroop effect was independent of both Stroop facilitation and of response conflict. While facilitation remained of comparable magnitude in both RSI-conditions, as in Burca et al. (2021, see Chapter 1), a short RSI of 200 ms reduced both semantic and conflicts in tandem. Therefore, this study provided more direct evidence for De Jong et al.'s original reasoning that performing the Stroop task under much faster pace (as induced by a short RSI) reduces attentional lapses and boosts a more consistent deployment of inhibitory abilities such that both the meaning of Stroop words and an incorrect response it generates interfere less with naming of their color – as indicated respectively by the reduction of semantic and response conflicts. Still, perhaps more parsimonious explanations in terms of changes in response speed were also considered in this chapter.

In light of the significant reduction of the Stroop interference effect via SLCC (as opposed to ALCC, Chapter 1) and short (as opposed to long) RSI (Chapter 3) observed in younger adults, the first goal of experiments depicted in **Chapter 4** was to investigate whether older adults' performance in the Stroop task can also be improved by these attention driving interventions. By investigating the type of conflict (semantic vs response) these interventions reduce, this study was also aimed at addressing further current controversies concerning the contribution of semantic conflict to the overall Stroop effect. In Experiment 1, semantic conflict again contributed significantly to the overall Stroop effect and this contribution was independent of both Stroop facilitation and of response conflict. However, a short (as opposed to long) RSI failed to reduce this conflict or any other aforementioned Stroop phenomena. In sum, this experiment failed to extend results conveyed in Chapter 3 to the population of older adults. SLCC (as opposed ALCC) on the other hand successfully optimized the older adults' Stroop performance and this via the significant reduction of semantic conflict. In this

Experiment 2, semantic conflict again contributed significantly to the overall Stroop interference effect but the contribution of response conflict remained non-significant (as in Burca et al., 2021, see also Chapter 1). Indeed, perhaps the most important finding reported in this experiment is that – in complete contrast to dominant single-stage response competition models (e.g., Roelofs, 2003) – not only interference occurs in the absence of response conflict (i.e., interference generated by same-response items) but no additional interference occurs for items that should in principle generate this type of conflict (i.e., different-response items).

Taken together, the findings conveyed in Chapters 1-4 have important implications for several pending issues in the Stroop literature.

The overall Stroop effect is a composite phenomenon

The present dissertation departed from the idea that the two-to-one Stroop paradigm is currently the most popular way of distinguishing semantic from response conflict (see e.g., A. Chen et al., 2011, 2013; Z. Chen et al., 2013; Hershman & Henik, 2019, 2020b; Jiang et al., 2015; Šaban & Schmidt, 2021; Schmidt et al., 2018; Schmidt & Cheesman, 2005; Shichel & Tzelgov, 2018; Van Veen & Carter, 2005). This is not surprising given that on critical same-response trials (e.g., *BLUE*_{red}) both word and color dimensions provide evidence toward the same response. Recall indeed that the two possible responses (i.e., red and blue here) are mapped on the same (e.g., “j”) response key. Consequently, according to unitary or single-stage response conflict models (e.g., Roelofs, 2003, see also e.g., Cohen et al., 1990; Glaser & Glaser, 1989; Logan, 1980; Melara & Algom, 2003; Phaf et al., 1990), these trials should not produce any interference, as they do not involve any response conflict. Alternatively, any interference produced by these trials can only result from semantic conflict – as originally reasoned by De Houwer (2003) on the basis of the two-conflicts model of Zhang and colleagues (Zhang et al., 1999; Zhang & Kornblum, 1998).

Because we completely subscribe to this alternative, our initial idea – rooted in Hashim and Parris (2014, 2015) – was that the evidence provided in De Houwer (2003)’s initial study is not sufficient to attest the viability of two-conflicts model of Zhang and colleagues (Zhang et al., 1999; Zhang & Kornblum, 1998) over single-stage response conflict models (e.g., Cohen et al., 1990; Glaser & Glaser, 1989; Logan, 1980; Melara & Algom, 2003; Phaf et al., 1990, Roelofs, 2003). Recall indeed that in De Houwer (2003) – and in multiple subsequent studies (see above) – both different- and same-response trials were responded to slower than baseline identity or color-congruent trials (*BLUE*_{blue}), and different-response trials (i.e., trials involving both semantic and response conflicts) were responded to slower than same-response trials. Still, all of these studies used color-congruent (or identity) trials as the baseline against which interference induced by same-response trials is measured. As a consequence, unitary or single-stage response conflict models of the Stroop effect (Cohen et al., 1990; Glaser & Glaser, 1989; Logan, 1980; Melara & Algom, 2003; Phaf et al., 1990; Roelofs, 2003) can still easily account for this difference as resulting from *facilitation* on color-congruent trials (as opposed to interference on same-response trials). In line with this idea, Hashim and Parris (2014, 2015; see also Hershman & Henik, 2019, 2020 for RT results) consistently reported significantly longer RTs for same-response trials than for color-congruent trials, but no difference between same-response trials and color-neutral trials (i.e., trials that are free of facilitation, see e.g., Brown, 2011; MacLeod, 1991 for discussion).

In contrast to this interpretation, and in line with preliminary evidence from Augustinova and Ferrand’s lab (see also Hershman & Henik, 2019, 2020 for pupillometric results), we reasoned that – under standard conditions of the Stroop task (e.g., standard visual presentation of Stroop stimuli with the RSI of 2000ms) – the difference between same-response and color-congruent trials simply confounds both *facilitation* generated by color-congruent trials and (*semantic*) *interference* generated by same-response trials. Said

differently, we reasoned that both phenomena indeed contribute to this difference, each in unknown amount (e.g., MacLeod, 1991).

To test this idea directly, and as in Hasshim and Parris (2014, Exp. 2), experiments depicted in Chapter 2, Chapter 3 and Experiment 1 depicted in Chapter 4 used the two-to-one Stroop paradigm with an additional (i.e., color-neutral) baseline. All of these different experiments replicated De Houwer (2003) since both different- and same-response trials were indeed responded to slower than baseline identity or color-congruent trials (*BLUE*_{blue}), and different-response trials (i.e., trials involving both semantic and response conflicts) were responded to slower than same-response trials. Also, and importantly, we failed to replicate Hasshim and Parris (2014, 2015; see also Hershman & Henik, 2019, 2020 for RT results). Overall, same-response trials were responded to slower than color-neutral trials with Bayesian evidence for this latter difference ranging from anecdotal to strong (e.g., overall, the study depicted in Chapter 2 provided extreme evidence for semantic conflict for older adults but only anecdotal for younger adults) and at the same time, color-congruent trials were responded to faster than color-neutral trials with Bayesian evidence for this latter difference ranging from inconclusive to moderate (e.g., the study depicted in Chapter 2 provided moderate evidence for facilitation overall, but it was only anecdotal for older adults and inconclusive for younger adults)¹.

Taken together, these findings provide novel evidence that the overall Stroop or congruency effect (i.e., the difference between standard color-incongruent and color-congruent trials) is indeed a composite phenomenon resulting from both facilitation and (standard) Stroop interference. While – based on MacLeod’s seminal review (1991) – this is

¹ To illustrate further, overall, the study depicted in Chapter 3 provided moderate Bayesian evidence in favor of facilitation in concert with anecdotal evidence in favor of semantic conflict. Under standard (i.e., long) RSI of 2000 ms however, the evidence in favor of facilitation continued to be moderate, exactly like that one in favor of semantic conflict. Furthermore, Experiment 1 depicted in Chapter 4 provided extreme Bayesian evidence in favor of both facilitation and semantic conflict.

of course not a substantial scoop, it is important to underlie that until up to this dissertation this was not shown in the two-to-one Stroop paradigm. Thus, not only our findings converge with those observed with standard Stroop paradigm (see e.g., Dalrymple-Alford, 1972; Dalrymple-Alford and Budayr, 1966, Exp. 2), but they have also important implications for current studies, namely those of conflict and of its control.

Indeed, magnitudes of a general conflict – central in these control studies (e.g., Egner et al., 2010) – are currently inferred from the overall Stroop (or congruency) effect ($BLUE_{\text{yellow}} - BLUE_{\text{blue}}$). As a result, a reduction of this “general” conflict – used as evidence for boosted cognitive control – is also currently inferred from the reduced overall Stroop (or congruency) effect. This means that *interference* produced by color-incongruent items – the reduction of which constitutes in reality *prima facie* evidence for increased cognitive control – is confounded with an opposing facilitation effect produced by color-congruent items (MacLeod, 1991; Parris et al., 2021).

Moreover, facilitation can actually increase substantially when, for instance, experimental setups encourage contingent associative learning between color-stimuli and responses (Schmidt, 2013). Given the existing control experiments do not often compare performance on incongruent trials to color-neutral trials, the extent to which different interventions designed to boost cognitive control reduce interference rather than boosts facilitation and/or contingent responding in the Stroop task is at this point still unclear and this despite several precautions taken toward this latter end (see e.g., Braem et al., 2019 for discussion).

This conclusion can be extended to the most recent studies that attempted to disentangle dynamics of cognitive control at stimulus vs. response level (e.g., Gajewski et al., 2020; Hirst et al., 2019; see also Braem et al., 2019 for discussion). Indeed, as clearly shown in our studies, in the absence of a color-neutral baseline, *interference* produced by same-

response incongruent items – the reduction of which constitutes in reality *prima facie* evidence for increased cognitive control at the stimulus level – is again potentially confounded with an opposing facilitation effect produced by color-congruent items. To conclude, because color-congruent items generate facilitation, using the overall Stroop (congruency) effect as a proxy for interference and the difference between same-response and color-congruent trials as a proxy for semantic conflict is at this point clearly questionable.

The Stroop interference effect is itself a composite phenomenon

In the Stroop literature, the dominant view is that the Stroop interference effect is a unitary phenomenon. Indeed, single-stage models assume that it results from a unique conflict occurring late in processing (i.e., at the level of response output, hence it is called response conflict; see e.g. Morton & Chambers, 1973; Hommel, 1997; LaBerge & Samuels, 1974; Shiffrin & Schneider, 1977 for early developments of these *late selection* accounts and e.g., Cohen et al., 1990; Glaser & Glaser, 1989; Logan, 1980; Melara & Algom, 2003; Phaf et al., 1990; Roelofs, 2003 for their most recent developments often placing the explanatory weight on parallel processing of the irrelevant and the relevant dimensions, as opposed to a single bottleneck in attention that was central in early developments).

Findings discussed in the previous section showed that semantic conflict – induced by same-response trials (i.e., trials that are free of response conflict) – clearly contributes to the overall Stroop effect and this independently of both facilitation and response conflict. Since unitary or single-stage response competition models (Cohen et al., 1990; Glaser & Glaser, 1989; Logan, 1980; Melara & Algom, 2003; Phaf et al., 1990; Roelofs, 2003) can account for facilitation produced by color-congruent items but not interference produced by same-response trials, they are clearly unable to account for the aforementioned findings showing that semantic conflict along with response conflict contributes to overall Stroop interference

occurring in tandem with Stroop facilitation. In sum, findings discussed in the previous section provided converging evidence that can only be accounted for via the multi-stage model of Zhang and colleagues (Zhang et al., 1999; Zhang & Kornblum, 1998). This converging evidence is likely to suggest that the overall color-word Stroop *interference* has at least two distinct loci: semantic and response conflicts.

It is however important to understand at this point that color-congruent trials – trials that are systematically included in the two-to-one Stroop paradigm (De Houwer, 2003) – are also known to amplify interference (see e.g., Roelofs, 2014 for their inclusion as void trials to this end). Therefore, we also additionally reasoned that the contribution of semantic conflict to the (overall) Stroop interference effect would be the most unambiguously demonstrated without the actual inclusion color-congruent trials. To test this idea directly, the experiment depicted in Chapter 1, and Experiment 2 depicted in Chapter 4 used the two-to-one Stroop paradigm that was modified accordingly. Also, and importantly, both same-response and different response trials were paired with their own color-neutral baseline to control for contingency. Therefore, similarly to past studies using “sky-put design” (Neely and Kahan, 2001; see also Manwell et al., 2004) or semantic Stroop interference (Augustinova & Ferrand, 2014b), the magnitudes of both semantic (e.g., $BLUE_{red} - DEAL_{red}$, used as a proxy for semantic conflict) and standard Stroop interference ($BLUE_{yellow} - DEAL_{yellow}$) were derived. However, this was done without using the same color-neutral baseline twice (see Shichel & Tzelgov, 2018, for discussion).

Overall, same-response trials were responded to slower than their color-neutral counter-parts with Bayesian evidence for this latter difference ranging from moderate to extreme. To illustrate, overall, the study depicted in Chapter 1 provided extreme evidence for semantic conflict that continued to be extreme in ALCC-condition and strong in SLCC. In Experiment 2 depicted in Chapter 4, the observation of semantic conflict overall was

reinforced by moderate Bayesian evidence. However, while in ALCC-condition the evidence in favor of this effect was strong or even very strong since close to 100 (i.e., $BF_{10} = 94.51$), in SLCC-condition, the contribution of semantic conflict was no longer significant and Bayes factor had actually provided moderate evidence against this effect. Still, under standard conditions of the Stroop task (e.g., typical visual presentation of Stroop stimuli as in ALCC), these studies provided substantial evidence that same-response items (i.e., items that are free of response conflict) induce significant interference compared to their color-neutral counterparts. Again, only multi-stage model of Zhang and colleagues (Zhang et al., 1999; Zhang & Kornblum, 1998) is able to account for this converging evidence suggesting that semantic conflict clearly.

Still, somewhat in contrast to this latter model but also in contrast to all single-stage response models, neither aforementioned study (Chapter 1 and Exp. 2 depicted in Chapter 4) provided substantial evidence for the contribution of response conflict to the overall Stroop interference. To illustrate, overall, this contribution was only favored by anecdotal evidence in Chapter 1 and while Bayesian evidence in favor of this contribution was inclusive in ALCC-condition, there was a moderate Bayesian evidence against this effect in SLCC. In Experiment 2 of Chapter 4, no conclusive Bayesian evidence for this contribution was found overall and in ALCC-condition, and moderate Bayesian evidence against this contribution was found in SLCC-condition. While SLCC-manipulation – administered within participants in a single block could have been responsible for this lack of response conflict –, it is important to remember that this lack of response conflict was also observed outside of this manipulation (e.g., in standard two-to-one paradigm administered to older adults in a study depicted in Chapter 2, see also e.g., Šaban, 2021). Therefore, we are likely to conclude that while under standard conditions of the Stroop task (e.g., typical visual presentation of Stroop stimuli as in ALCC, typical RSI of 2000ms), the Stroop interference is likely to result from both semantic

and response conflicts (as predicted by Zhang and colleagues' two stage model), but under specific conditions (discussed here below in greater details) it can only result from semantic (as shown in aforementioned studies) or response conflict (as shown by Hasshim & Parris, 2014, 2015). To conclude, the finding discussed in this and the previous section clearly imply that selection in the Stroop task can occur at the level of semantics – a finding that runs counter to the dominant single-stage response competition models of Stroop interference. Therefore, as previously mentioned, these latter models need to be modified to make room for this type of conflict (see below).

Semantic conflict is permeable to moderators

As discussed in the General Introduction, several past studies indicate that semantic conflict – induced by semantic associates (e.g., *SKY_{green}*) – is spared by experimental manipulations like hypnosis-like suggestion (Augustinova & Ferrand, 2012), priming of dyslexia (Augustinova & Ferrand, 2014a), variations in Response–Stimulus Intervals (Augustinova, Silvert et al., 2018), viewing position (Ferrand & Augustinova, 2014) and Single-letter Coloring and Cueing (Augustinova & Ferrand, 2007; Augustinova et al., 2010, 2015, Augustinova, Clarys et al., 2018), while response conflict is not. Response conflict is indeed significantly reduced by these interventions in the Semantic Stroop paradigm (see Figure 3, p.15). Recall however that this paradigm has been criticized for its ambiguous induction of semantic conflict. Therefore, the second goal of this dissertation was to examine the extent to which the aforementioned dissociative pattern (i.e., preservation of semantic but not of response conflict) – used as evidence that the two conflicts (semantic vs. response) are qualitatively distinct (Augustinova & Ferrand, 2014b; see also Parris et al., 2021 for discussion) – can be extended to the two-to-one Stroop paradigm.

Results presented in all four chapter clearly converge toward the idea that, in the two-to-one Stroop paradigm, semantic conflict is largely permeable to moderators. Indeed, it is largely reduced by attention driving interventions such as SLCC (as opposed to ALCC, see Chapter 1 and Experiment 2 of Chapter 4) and short RSI (as opposed to long RSI, see chapter 3 and 4). To illustrate, SLCC shortened responses to same-response trials in both young (Chapter 1) and older adults (Exp.2, Chapter 4). However, this reduction was supported respectively by anecdotal and moderate Bayesian evidence and in both experiments, the evidence in favor of interaction between the type of interference (semantic vs. standard) was inconclusive. This latter pattern is therefore the most consistent with the idea that SLCC (vs. ALCC) is likely to reduce both semantic and response conflicts in tandem.

Similar pattern was observed with short (as opposed to long) RSI – although Experiment 1 presented in Chapter 4 failed to extend this pattern to the population of older adults. Finally, findings presented in Chapter 2 are also consistent with the idea that semantic conflict is permeable to moderators. Although the interaction between the type of conflict and age-group was only supported by anecdotal evidence, semantic conflict was significantly larger in healthy older (vs. younger) adults. Still, this interaction seems to rather reflect the lack of response conflict that was observed in older adults (see the previous section for discussion), rather than the genuine preservation of response conflict. In sum, in light of these different results, we are likely to conclude the absence of dissociative pattern in the two-to-one Stroop paradigm such that the aforementioned moderators are likely to reduce both conflicts in tandem – although, as discussed in different chapters, our data are less clear-cut on this issue. Given the differences in magnitudes of semantic conflict induced via semantic associates as compared to same-response trials, we are likely to conclude that in past studies of Augustinova and Ferrand, semantic conflict might not be affected by various experimental manipulations simply because it is too small to be affected. Said differently, it is not the type

of conflict (semantic as opposed to response conflict) but its (small) magnitude that leads to its preservation. As previously discussed in Chapter 2, this parsimonious explanation – that allows to reconcile past findings² – can actually apply to the apparent preservation of response conflict in older adults observed in the study depicted in this chapter.

This absence of dissociation might suggest that semantic and response conflict observed in the two-to-one Stroop paradigm are not qualitatively different. Still, in several studies reported in the present dissertation, response conflict failed to contribute significantly to the overall Stroop interference (Chapter 1, Chapter 2 and Exp. 2 depicted in Chapter 4). Conversely, semantic (unlike response) conflict was statistically eliminated under a short RSI (although both were reduced by this manipulation) in Chapter 3 (but see Exp.1 of Chapter 4 for the absence of this effect in older adults). This selective lack of contribution often occurred while the overall Stroop interference was still of considerable magnitude. Therefore, in line with the idea that semantic and response conflict and semantic conflict are qualitatively different, these latter results are likely to indicate that semantic and response occur independently one from another. Additionally, the fact that there is no overlap in neural activations for semantic and response conflict observed in the two-to-one Stroop paradigm (although administered with color-congruent trials as baseline) is also consistent with this latter idea (Chen et al., 2013; Milham et al., 2001, van Veen and Carter, 2005). For instance, van Veen and Carter (2005) showed the activation of dorsolateral pre-frontal cortex, posterior parietal cortex, and the ACC for semantic conflict, whereas they showed the activation of more inferior lateral prefrontal cortex, left premotor areas and of more anterior and ventral regions of the ACC for response conflict (but see Parris et al., 2019 for no involvement of

² This explanation is indeed compatible with Manwell et al.'s Account 1 suggesting that SLCC simultaneously reduces both semantic and standard interference. It is however still incompatible with their Account 2 suggesting that SLCC curtails semantic conflict altogether (Manwell et al., 2004; see also Besner et al., 2016 for discussion).

ACC). The study therefore concluded that there are separate but analogous mechanisms for dealing with these two different conflicts (see also Chen et al., 2013; Milham et al., 2001).

The fact that semantic conflict contributes to the overall Stroop interference (and this with a magnitude that is different from that of response conflict) and it is permeable to various moderators has important implication of neuropsychological practice. Indeed, current implementations of the Stroop task used in this practice (e.g., Strauss et al., 2006) – rooted in unitary models of Stroop interference – cannot account for the extent to which normal development trajectory (see e.g., Burca et al., 2022 or Chapter 2; see also Ferrand et al., 2020 for the investigation in children) along with various clinical conditions selectively impact semantic and response conflict. Indeed, as already discussed in Ferrand and colleagues (2020), and in Parris et al. (2021), the extra sensitivity of the Stroop test (stemming from the ability to detect and rate each of these components separately) would provide clinical practitioners with invaluable information since the different forms of conflict are possibly detected and resolved by different neural mechanisms operating in distinct neural regions. In sum, results of the studies included in this dissertation – including those indicating that the cognitive functioning of the elderly can be optimized (see Chapter 4) – call for changes in Stroop research practices in basic and applied clinical research (see e.g., Augustinova et al., 2018, Parris et al., 2021 for similar argument).

Note however that using same-response items to induce semantic conflict seems rather inadequate in this context since the use of these items requires manual response mode (i.e., response that is not the best suited for clinical populations)³. Yet, using semantic-associates instead does not seem a straightforward option either even though – as discussed in Chapter 2

³ As discussed in Chapter 2 for instance same-response trials have their own shortcoming. Because both dimensions of these trials provide evidence towards the same response, they cannot (unlike semantic associates) generate response conflict. However, they can still produce response facilitation. It still important to note this response facilitation actually reduces and not inflates the magnitude of interference these items generate.

– it now seems reasonable to assume that both semantic-associative and same-response trials actually induce semantic conflict (but in unknown quantities for the former). Quétard et al.’s (submitted) investigation of mouse-tracking trajectory is in line with the idea that semantic-associates are likely to involve at least some response conflict (see also Risko et al., 2006; Šaban, 2001). In this study, the magnitude of interference induced by incongruent semantic-associates (e.g., *SKY*_{green}) was smaller than the one induced by incongruent out-of-the set items (e.g., *PURPLE*_{green}) while mouse deviation was actually more important for the former than for the latter items. This pattern of results suggests that the induction of semantic conflict by out-of-the set items might be more analogous (in terms of magnitude) to the induction via same-response items. It is also consistent with the idea that semantic associates involve at least some response conflict – most likely due to their direct association with response set color (see e.g., Roelofs, 2003) that is not present for out-of-the set items. Still, future studies need to address the functional equivalence of these different types of items directly – the issue that cannot be successfully addressed without using a more time-course sensitive measure such as ERPs (or EMGs⁴). This is precisely what we intend to do in our future studies.

Concluding remarks

Results depicted in Chapter 3 suggested that the composite nature of Stroop interference (semantic vs. response conflict) and more generally of both Stroop facilitation and interference – originally anticipated by the single multi-stage model of Zhang and colleagues (Zhang et al., 1999; Zhang & Kornblum, 1998) – is likely a function of (slow)

⁴ The initial goal of the present dissertation was to address a related issue of EMG correlates of different components of the Stroop effect. However, the EMGs data, namely specific indicators of the conflict (i.e., double activations in general, and a more refined categorization where the pre-motor double activations and post-motor double activations, were examined, with the goal to account for reactive and/or pro-active cognitive control, the latencies and width and height of these activations were also examined) occurred insufficiently to be meaningfully analyzed and interpreted. Therefore, future studies need to address these issues with using much more trials per participant/conditions compared to the number of trials used in present studies.

response speed, and thus of poorer performance. To investigate this issue further, Martinon, Ferrand, Burca...and Augustinova (submitted), used again the two-to-one Stroop paradigm with an additional (i.e., color-neutral) baseline. In two experiments, Bayesian evidence in favor of semantic conflict was extreme and independent of both response conflict and facilitation; all three together comprising the overall Stroop effect. However, vincentizing reaction times further revealed that in the 25% fastest trials, the overall Stroop effect was entirely driven by facilitation. In the next 25% fastest trials, the overall Stroop effect comprised both interference and facilitation, but interference was driven solely by response conflict. It was only in the slowest 50% of trials when the overall Stroop effect comprised facilitation and both semantic and response conflict. These results suggest in sum that response speed determines the composition of the overall Stroop effect (as first suggested by results of the study depicted in Chapter 3).

As already suggested in this latter chapter, perhaps the most important implication of these results is that taking response speed into account allows to reconcile contrasting findings of Hasshim and Parris (2014) and those conveyed in the present dissertation⁵. Indeed, findings presented in Chapter 3 along with those of Martinon and colleagues (submitted) suggest that any factor that influences response speed should modify the nature of the Stroop effect and that this modification should primarily concern the contribution of semantic conflict to the overall Stroop interference. Still, as already argued by Martinon and colleagues, Zhang and colleagues' model – that the findings of the present dissertation are likely to favor – is unable to account for semantic conflict occurring after response conflict. It remains possible that semantic processing of the irrelevant word dimension of Stroop stimuli triggers the conflict at the same time as response (as Zhang et al.' model would predict), but that semantic conflict is

⁵ Recall that in Experiment 2a of Hasshim and Parris (2014), response times were indeed significantly shorter than even in younger participants of Burca et al. (2022, see Supplementary online materials of this study for non z-scored RTs).

detected only after response conflict. Again, this possibility – that we are likely to favor – that cannot be successfully addressed without using a more time-course sensitive measure such as ERPs (or EMGs) either. Even if this true, Zhang and colleagues’ model is still unable to account for task conflict that is also likely to occur in the Stroop task (the only model that can currently account for this type of conflict (i.e., Kalanthroff et al., 2018’ model) is yet unable to account for semantic conflict observed in the present dissertation that). Finally, Zhang and colleagues’ model anticipates both semantic and response conflicts are resolved at the level of response selection (i.e., before the actual motoric action is initiated). However, the aforementioned mouse-tracking study of Quétard and colleagues’ (submitted, see also Bundt et al., 2018) showed no Stroop effects in initiation times only in later mouse-tracking parameters. This pattern suggests Stroop effects can occur entirely after the response has been initiated (i.e., during the response execution phase, see also Kello et al., 2000; Exp.2 for a spillover of Stroop effect from response selection to response execution). These different shortcomings of Zhang and colleagues’ models suggest in sum that the issue of loci of the Stroop effect – that the present dissertation was aimed to address – is still far from being theoretically accounted for. Thus, more empirical and theoretical efforts are still needed.

References

- Augustinova, M., & Ferrand, L. (2007). Influence de la présentation bicolore des mots sur l’effet Stroop. *L’Année Psychologique*, 107(02), 163-179. <https://doi.org/10.4074/S0003503307002011>
- Augustinova, M., & Ferrand, L. (2012). Suggestion does not de-automatize word reading: Evidence from the semantically based Stroop task. *Psychonomic Bulletin & Review*, 19(3), 521–527. <https://doi.org/10.3758/s13423-012-0217-y>
- Augustinova, M., & Ferrand, L. (2014a). 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>

- Augustinova, M., & Ferrand, L. (2014b). Automaticity of Word Reading. *Current Directions in Psychological Science*, 23(5), 343–348. <https://doi.org/10.1177/0963721414540169>
- Augustinova, M., Flaudias, V., & Ferrand, L. (2010). Single-letter coloring and spatial cuing do not eliminate or reduce a semantic contribution to the Stroop effect. *Psychonomic Bulletin & Review*, 17(6), 827–833. <https://doi.org/10.3758/PBR.17.6.827>
- Augustinova, M., Silvert, L., Ferrand, L., Llorca, P. M., & Flaudias, V. (2015). Behavioral and electrophysiological investigation of semantic and response conflict in the Stroop task. *Psychonomic Bulletin & Review*, 22(2), 543–549. <https://doi.org/10.3758/s13423-014-0697-z>
- Augustinova, M., Silvert, L., Spatola, N., & Ferrand, L. (2018a). Further investigation of distinct components of Stroop interference and of their reduction by short response-stimulus intervals. *Acta Psychologica*, 189, 54–62. <https://doi.org/10.1016/J.ACTPSY.2017.03.009>
- Augustinova, M., Clarys, D., Spatola, N., & Ferrand, L. (2018b). Some further clarifications on age-related differences in Stroop interference. *Psychonomic Bulletin & Review*, 25(2), 767–774. <https://doi.org/10.3758/s13423-017-1427-0>
- 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>
- Braem, S., Bugg, J. M., Schmidt, J. R., Crump, M. J. C., Weissman, D. H., Notebaert, W., & Egner, T. (2019). Measuring Adaptive Control in Conflict Tasks. *Trends in Cognitive Sciences*, 23(9), 769–783. <https://doi.org/10.1016/j.tics.2019.07.002>
- Brown, T. L. (2011). The relationship between stroop interference and facilitation effects: Statistical artifacts, baselines, and a reassessment. *Journal of Experimental Psychology: Human Perception and Performance*, 37(1), 85–99. <https://doi.org/10.1037/a0019252>
- Bundt, C., Ruitenberg, M. F. L., Abrahamse, E. L., & Notebaert, W. (2018). Early and late indications of item-specific control in a Stroop mouse tracking study. *PLOS ONE*, 13(5), e0197278. <https://doi.org/10.1371/journal.pone.0197278>
- Burca, M., Beauconsin, V., Chausse, P., Ferrand, L., Parris, B. A., & Augustinova, M. (2021). Is There Semantic Conflict in the Stroop Task? *Experimental Psychology*, 68(5), 274–283. <https://doi.org/10.1027/1618-3169/a000530>

- Burca, M., Chausse, P., Ferrand, L., Parris, B. A., & Augustinova, M. (2022). Some further clarifications on age-related differences in the Stroop task: New evidence from the two-to-one Stroop paradigm. *Psychonomic Bulletin & Review*, 29, 492–500. <https://doi.org/10.3758/s13423-021-02011-x>
- Chen, A., Bailey, K., Tiernan, B. N., & West, R. (2011). Neural correlates of stimulus and response interference in a 2–1 mapping stroop task. *International Journal of Psychophysiology*, 80(2), 129–138. <https://doi.org/10.1016/j.ijpsycho.2011.02.012>
- Chen, A., Tang, D., & Chen, X. (2013). Training Reveals the Sources of Stroop and Flanker Interference Effects. *PLoS ONE*, 8(10), e76580. <https://doi.org/10.1371/journal.pone.0076580>
- Chen, Z., Lei, X., Ding, C., Li, H., & Chen, A. (2013). The neural mechanisms of semantic and response conflicts: An fMRI study of practice-related effects in the Stroop task. *NeuroImage*, 66, 577–584. <https://doi.org/10.1016/j.neuroimage.2012.10.028>
- 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>
- Dalrymple-Alford, E. C. (1972). Associative facilitation and interference in the Stroop color-word task. *Perception & Psychophysics*, 11(4), 274–276. <https://doi.org/10.3758/BF03210377>
- Dalrymple-Alford, E. C., & Budayr, B. (1966). Examination of Some Aspects of the Stroop Color-Word Test. *Perceptual and Motor Skills*, 23, 1211–1214. <https://doi.org/10.2466/pms.1966.23.3f.1211>
- De Houwer, J. (2003). On the role of stimulus-response and stimulus-stimulus compatibility in the Stroop effect. *Memory & Cognition*, 31(3), 353–359. <https://doi.org/10.3758/BF03194393>
- Egner, T. (2010). Going, going, gone: Characterizing the time-course of congruency sequence effects. *Frontiers in Psychology*, 1. <https://doi.org/10.3389/fpsyg.2010.00154>
- Ferrand, L., Ducrot, S., Chausse, P., Maionchi-Pino, N., O'Connor, R. J., Parris, B. A., Perret, P., Riggs, K. J., & Augustinova, M. (2020). Stroop interference is a composite phenomenon: Evidence from distinct developmental trajectories of its components. *Developmental Science*, 23(2). <https://doi.org/10.1111/desc.12899>
- Gajewski, P. D., Falkenstein, M., Thönes, S., & Wascher, E. (2020). Stroop task performance across the lifespan: High cognitive reserve in older age is associated with enhanced proactive and reactive interference control. *NeuroImage*, 207, 116430. <https://doi.org/10.1016/j.neuroimage.2019.116430>

- Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General*, 118(1), 13–42.
- Hasshim, N., & Parris, B. A. (2014). Two-to-one color-response mapping and the presence of semantic conflict in the Stroop task. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.01157>
- Hasshim, N., & Parris, B. A. (2015). Assessing stimulus–stimulus (semantic) conflict in the Stroop task using saccadic two-to-one color response mapping and prereponse pupillary measures. *Attention, Perception, & Psychophysics*, 77(8), 2601–2610. <https://doi.org/10.3758/s13414-015-0971-9>
- Hershman, R., & Henik, A. (2019a). Dissociation between reaction time and pupil dilation in the Stroop task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(10), 1899–1909. <https://doi.org/10.1037/xlm0000690>
- Hershman, R., & Henik, A. (2019b). Dissociation between reaction time and pupil dilation in the Stroop task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(10), 1899–1909. <https://doi.org/10.1037/xlm0000690>
- Hershman, R., & Henik, A. (2020a). Pupillometric contributions to deciphering Stroop conflicts. *Memory & Cognition*, 48(2), 325–333. <https://doi.org/10.3758/s13421-019-00971-z>
- Hershman, R., & Henik, A. (2020b). Pupillometric contributions to deciphering Stroop conflicts. *Memory & Cognition*, 48(2), 325–333. <https://doi.org/10.3758/s13421-019-00971-z>
- Hirst, R. J., Kicks, E. C., Allen, H. A., & Cragg, L. (2019). Cross-modal interference-control is reduced in childhood but maintained in aging: A cohort study of stimulus- and response-interference in cross-modal and unimodal Stroop tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 45(5), 553–572. <https://doi.org/10.1037/xhp0000608>
- Hommel, B. (1997). Interactions between stimulus-stimulus congruence and stimulus-response compatibility. *Psychological Research*, 59(4), 248–260. <https://doi.org/10.1007/BF00439302>
- Jiang, J., Zhang, Q., & Van Gaal, S. (2015). EEG neural oscillatory dynamics reveal semantic and response conflict at difference levels of conflict awareness. *Scientific Reports*, 5(1), 12008. <https://doi.org/10.1038/srep12008>
- Jong, R. D., Berendsen, E., & Cools, R. (1999). Goal neglect and inhibitory limitations: Dissociable causes of interference effects in conflict situations. *Acta Psychologica*, 16. [https://doi.org/10.1016/S0001-6918\(99\)00012-8](https://doi.org/10.1016/S0001-6918(99)00012-8)

- Kalanthroff, E., Davelaar, E. J., Henik, A., Goldfarb, L., & Usher, M. (2018). Task conflict and proactive control: A computational theory of the Stroop task. *Psychological Review*, 125(1), 59–82.
<https://doi.org/10.1037/rev0000083>
- Kello, C. T., Plaut, D. C., & MacWhinney, B. (2000). The task dependence of staged versus cascaded processing: An empirical and computational study of Stroop interference in speech perception. *Journal of Experimental Psychology: General*, 129(3), 340–360. <https://doi.org/10.1037/0096-3445.129.3.340>
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6(2), 293–323. [https://doi.org/10.1016/0010-0285\(74\)90015-2](https://doi.org/10.1016/0010-0285(74)90015-2)
- Li, K. Z. H., & Bosman, E. A. (1996). Age differences in stroop-like interference as a function of semantic relatedness. *Aging, Neuropsychology, and Cognition*, 3(4), 272–284.
<https://doi.org/10.1080/13825589608256630>
- Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology*, 12(4), 523–553. [https://doi.org/10.1016/0010-0285\(80\)90019-5](https://doi.org/10.1016/0010-0285(80)90019-5)
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109(2), 163–203. <https://doi.org/10.1037/0033-2909.109.2.163>
- Manwell, L. A., Roberts, M. A., & Besner, D. (2004). Single letter coloring and spatial cuing eliminates a semantic contribution to the Stroop effect. *Psychonomic Bulletin & Review*, 11(3), 458–462.
<https://doi.org/10.3758/BF03196595>
- Martinon L., Ferrand L., Burca M., Hasshim N., Lakhzoum D., Parris B.A, Silvert L., Augustinova M. (submitted). Variations in Response Speed reveal the Polymorphic Nature of the Stroop Effect. *Journal of Experimental Psychologie: Learning, Memory and Cognition*.
- Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review*, 110(3), 422–471. <https://doi.org/10.1037/0033-295X.110.3.422>
- Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., & Kramer, A. F. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Cognitive Brain Research*, 12(3), 467–473. [https://doi.org/10.1016/S0926-6410\(01\)00076-3](https://doi.org/10.1016/S0926-6410(01)00076-3)
- Morton, J., & Chambers, S. M. (1973). Selective Attention to Words and Colours. *Quarterly Journal of Experimental Psychology*, 25(3), 387–397. <https://doi.org/10.1080/14640747308400360>
- 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>
- Parris, B. A., Hasshim, N., Wadsley, M., Augustinova, M., & Ferrand, L. (2021). The loci of Stroop effects: A critical review of methods and evidence for levels of processing contributing to color-word Stroop effects and the implications for the loci of attentional selection. *Psychological Research*, 0123456789. <https://doi.org/10.1007/s00426-021-01554-x>
- Phaf, R. H., Van der Heijden, A. H. C., & Hudson, P. T. W. (1990). SLAM: A connectionist model for attention in visual selection tasks. *Cognitive Psychology*, 22(3), 273–341. [https://doi.org/10.1016/0010-0285\(90\)90006-P](https://doi.org/10.1016/0010-0285(90)90006-P)
- Quétaud, B., Spatola, N., Parris, B. A., Ferrand, L., & Augustinova, M. (submitted). A mouse-tracking study of the composite nature of the Stroop effect at the level of response execution. *Quarterly Journal of Experimental Psychology*.
- Risko, E. F., Schmidt, J. R., & Besner, D. (2006). Filling a gap in the semantic gradient: Color associates and response set effects in the Stroop task. *Psychonomic Bulletin & Review*, 13(2), 310–315. <https://doi.org/10.3758/BF03193849>
- Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review*, 110(1), 88–125. <https://doi.org/10.1037/0033-295X.110.1.88>
- Šaban, I. (2021). Basic processes in interference paradigms. Defended PhD dissertation at Université de Bourgogne Franche-Comté.
- Šaban, I., & Schmidt, J. R. (2021). Stimulus and response conflict from a second language: Stroop interference in weakly-bilingual and recently-trained languages. *Acta Psychologica*, 218, 103360. <https://doi.org/10.1016/j.actpsy.2021.103360>
- Schmidt, J. R., & Cheesman, J. (2005). Dissociating stimulus-stimulus and response-response effects in the Stroop task. *Canadian Journal of Experimental Psychology*, 59(2), 132–138. <https://doi.org/10.1037/H0087468>
- Schmidt, J. R., Hartsuiker, R. J., & De Houwer, J. (2018). Interference in Dutch–French Bilinguals. *Experimental Psychology*, 65(1), 13–22. <https://doi.org/10.1027/1618-3169/a000384>
- Shichel, I., & Tzelgov, J. (2018). Modulation of conflicts in the Stroop effect. *Acta Psychologica*, 189, 93–102. <https://doi.org/10.1016/j.actpsy.2017.10.007>

- 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>
- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms and Commentary*. New York: Oxford University Press.
- van Veen, V., & Carter, C. S. (2005). Separating semantic conflict and response conflict in the Stroop task: A functional MRI study. *NeuroImage*, 27(3), 497–504. <https://doi.org/10.1016/j.neuroimage.2005.04.042>
- Zhang, H. (Harry), Zhang, J., & Kornblum, S. (1999). A Parallel Distributed Processing Model of Stimulus–Stimulus and Stimulus–Response Compatibility. *Cognitive Psychology*, 38(3), 386–432.
<https://doi.org/10.1006/cogp.1998.0703>
- Zhang, H., & Kornblum, S. (1998a). The Effects of Stimulus-Response Mapping and Irrelevant Stimulus-Response and Stimulus-Stimulus Overlap in Four-Choice Stroop Tasks with Single-Carrier Stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 24(1), 3–19.
<https://doi.org/10.1037/0096-1523.24.1.3>
- Zhang, H., & Kornblum, S. (1998b). The effects of stimulus-response mapping and irrelevant stimulus-response and stimulus-stimulus overlap in four-choice Stroop tasks with single-carrier stimuli. *Journal of Experimental Psychology. Human Perception and Performance*, 24(1), 3–19.
- Zhang, H., Zhang, J., & Kornblum, S. (1999). A Parallel Distributed Processing Model of Stimulus–Stimulus and Stimulus–Response Compatibility. *Cognitive Psychology*, 38(3), 386–432.
<https://doi.org/10.1006/COGP.1998.0703>