

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). Copyright © [2022] (Experimental Psychology Society). DOI: [DOI 0.1177/17470218211030863].

**Prepotent task-irrelevant semantic information is dampened by domain-specific control mechanisms during visual word recognition.**

Simone Sulpizio<sup>1,2</sup>, Remo Job<sup>3</sup>, Paolo Leoni<sup>3</sup>, & Michele Scaltritti<sup>3</sup>

<sup>1</sup>Department of Psychology, University of Milano-Bicocca, Italy

<sup>2</sup>Milan Center for Neuroscience (NeuroMI), University of Milano-Bicocca

<sup>3</sup>Department of Psychology and Cognitive Science, University of Trento, Italy

Running Head: SEMANTIC CONTROL IN WORD RECOGNITION

**Author Note**

Data are available on the Open Science Framework: <https://osf.io/nk3yt/>

Correspondence concerning the article should be addressed to: Michele Scaltritti,

Dipartimento di Psicologia e Scienze Cognitive, Corso Bettini, 31 – 38068, Rovereto (TN), Italy,

email: [michele.scaltritti@unitn.it](mailto:michele.scaltritti@unitn.it)

## Abstract

We investigated whether semantic interference occurring during visual word recognition is resolved using domain-general control mechanism or using more specific mechanisms related to semantic processing. We asked participants to perform a lexical decision task with taboo stimuli, which induce semantic interference, as well as well as a semantic Stroop task and a Simon task, intended as benchmarks of linguistic-semantic and non-linguistic interference, respectively. Using a correlational approach, we investigated potential similarities between effects produced in the three tasks, both at the level of overall means and as a function of response speed (delta-plot analysis). Correlations selectively surfaced between the lexical decision and the semantic Stroop task. These findings suggest that, during visual word recognition, semantic interference is controlled by semantic-specific mechanisms, which intervene to face prepotent but task-irrelevant semantic information interfering with the accomplishment of the task's goal.

*Keywords:* visual word recognition; taboo words; lexical decision; cognitive control; semantics; semantic Stroop; Simon

Reading, on par with most of human behaviors, can be considered as a goal-directed activity in which the orthographic input is decoded and processed in order to extract meaning. On the one side, visual word recognition thus hinges upon specific processes related to the orthographic decoding and to the mapping of orthographic information onto meaning and phonology. On the other side, visual word recognition should also be accessible to control mechanisms to ensure a reading performance that meets the relevant contextual goals. Consistently, research has shown the importance of executive attention during visual word recognition in, for example, controlling the balance between lexical vs sublexical processing (e.g., Reynolds & Besner, 2005; Zevin & Balota, 2000; but see also Kinoshita & Lupker, 2002) or regulating the information flow through different representational stages (e.g., O'Malley & Besner, 2008; Scaltritti et al., 2013) in order to meet contextual demands (for reviews, Balota et al., 1999; Besner et al., 2016).

The role of cognitive control remains comparatively under-investigated with respect to semantic processing, possibly because the retrieval of semantic information is assumed to occur automatically, after the “magic moment” (Balota & Yap, 2006) of word recognition. However, recent proposals clearly point out that, other than being retrieved from long-term memory, conceptual representations are also dynamically processed via control mechanisms that align semantic information with contextual goals and constraints (Lambon Ralph et al., 2017).

Whether this semantic control is performed or not by the same general mechanisms and resources recruited in other domains is a matter of ongoing debate (e.g., Belke & Stielow, 2013; Chiou et al., 2018; Montefinese et al., 2020; Piai et al., 2013).

Domain-general mechanisms may be more involved in biasing semantic activation to meet explicit tasks goals when a competition surfaces between secondary but task-relevant representations and prepotent but task-irrelevant ones. Differently, control mechanisms specific

for semantic processing would be more involved in the controlled retrieval of task-relevant but non-dominant semantic features (e.g., Davey et al., 2016; Hoffman et al., 2018; Lambon Ralph et al., 2017). This distinction between domain-general vs specific control mechanisms becomes cogent in visual word recognition in order to a) provide a framework to interpret the role of semantics beyond the assumption of an epiphenomenal consequence of word recognition, and b) understand the mechanisms linking word processing and executive control. The present study aims to contribute to these issues by exploring whether semantic interference phenomena occurring during visual word recognition are resolved via domain-general control mechanisms or, instead, may also require more specific mechanisms related to linguistically-driven semantic processing. To reach this aim, we adopted a cross-task correlational approach and tested correlations between semantic interference effects arising during visual word recognition with both a) a non-semantic conflict effect (the Simon effect), and b) a semantic interference effect (the semantic Stroop effect). The cross-task correlational approach is widely used to investigate both similarities and differences across different functions in the executive domain (e.g., Fan et al., 2003; Spagna et al., 2015; Stins et al., 2005), as well as the involvement of executive control in language processing (e.g., Boned et al., 2021; Crowther & Martin, 2014; Korko et al., 2021). The use of this approach thus makes our results directly comparable with the extant literature.

For visual word recognition, we used a lexical decision task exploiting the taboo connotation of the stimuli to trigger a semantic interference. Compared to neutral stimuli, taboo (i.e., socially inappropriate) words have been repeatedly showed to slow down responses in visual lexical decision (Carretié et al., 2008; Geer & Bellard, 1996; Madan et al., 2017; Sulpizio et al., 2019; Sulpizio & Navarrete, 2020). This detrimental influence has been mostly related to the attentional-capture phenomena triggered by the salience of the taboo connotation (e.g.,

MacKay et al., 2004), which would divert resources away from the main task. Interestingly, taboo words have been found to slow down response latencies even when used as mere distractors within word production tasks such as the picture word interference (e.g., Dhooge & Hartsuiker, 2011). This effect has been hypothesized to reflect (post-lexical) monitoring mechanisms that would operate to prevent any erroneous lexicalization of embarrassing and inappropriate taboo distractors. Recent investigations of the two accounts within the lexical decision paradigm have shown evidence which more directly supports an early locus of the effect, at least in the context on visual word recognition (Scaltritti et al., 2021; Sulpizio et al., 2019; 2020).

Importantly, although the taboo interference elicited by printed words may influence different levels of processing, all accounts converge in highlighting its semantic nature and origin. It is in fact only by accessing to the word semantics that any taboo connotation can become available and thus influence the performance. Within the scope of the present experiment, the taboo connotation was used as a proxy for a prepotent semantic content potentially requiring control, due to its detrimental influence on the achievement of task goals: According to Lambon Ralph et al. (2017), dealing with an emphatic uncharacteristic feature of the stimulus would increase the reliance on the control processes. Further, in order to focus on a purely semantic level, we ensured that taboo connotation was not associated with a specific response, i.e., it was not predictive of lexical status. The taboo connotation was thus manipulated within both words and pseudowords (which, following Sulpizio et al., 2020, were created by changing one letter in taboo and non-taboo words).

To measure semantic and non-semantic interference we used the (semantic) Stroop and the Simon task, respectively. At first sight, the use of these tasks to measure control mechanisms

might seem problematic, and not just because of the multiple stark differences between these paradigms and the lexical decision task. In fact, executive control literature has shown no correlations between the interference effects triggered by the Simon and the Stroop tasks, suggesting that different forms of interference may map onto different control networks and that they may not be suited as measures of individual differences (e.g., Fan et al., 2003; Stins et al., 2005; Shilling et al., 2002). Nonetheless, as we argue below, there are some elements of the Simon and the Stroop paradigm that we deem potentially useful for our purposes.

In the Simon task (Simon & Small, 1969), a response conflict is induced by the mismatch between task relevant and task-irrelevant prepotent information. In the classic task configuration, participants are required to choose between a left- and right-hand manual response according to the color of a visual stimulus. Critically, the visual stimulus can be either to the left or to the right of a fixation point. Reaction times are slower in incompatible trials, that is when there is a mismatch between stimulus location and response-hand. A line of investigations has highlighted the importance of selective suppression mechanisms recruited to dampen the conflict triggered by the prepotent but task irrelevant location of the stimulus. Supporting evidence for this mechanism is provided by the distributional profile of the Simon effect, as the slowdown for incompatible trials is reduced (and actually shows a reversal) in slower responses, where the selective suppression process has sufficient time to fully accrue (e.g., Burle et al., 2002; Pratte et al., 2010; van den Wildenberg et al., 2010).

Notably, in a recent lexical decision study with taboo and non-taboo words and pseudowords, Scaltritti et al. (2021) have reported that taboo interference effect has a similar time course. Although the taboo interference and the Simon effect are clearly different phenomena – the former is elicited by a semantic feature of the linguistic stimuli, whereas the

latter stems from response conflict triggered by stimulus location – in both tasks participants have to face a prepotent, task-irrelevant feature of the target that hampers the performance. The similarity in the pattern may thus point toward a domain general mechanism of selective suppression that intervenes across different situations.

Differently, the semantic Stroop task induces a semantic interference and may thus be informative on the involvement of more specific control mechanisms acting on semantic information. In the semantic Stroop task (Neely & Kahan, 2001) participants are presented with colored words (e.g., STRAWBERRY) that are semantically associated with a color (i.e., red) and have to name the color while ignoring the word. Colored words are presented in incongruent colors (e.g., STRAWBERRY written in blue) and compared with color-neutral words (e.g., TABLE written in blue). This task configuration allows to isolate semantic interference from the response conflict which typically occurs in the standard Stroop task, where the color-incongruent word activates an incorrect response (Augustinova et al., 2018; Augustinova & Ferrand, 2014). In the semantic Stroop, in fact, the color words are not part of the response set. This is similar to what happens in the taboo lexical decision task, in which the alleged interference arises at the semantic level only, since taboo information is not associated to any specific response – i.e., participants cannot make a decision on the basis of taboo information. A similarity in the interference pattern of these two tasks, together with their dissociation from the pattern reported in the Simon task, would speak in favor of a more specific control mechanism acting on semantic information triggered by linguistic stimuli, which would be responsible for the identification and the management of the task's relevant aspect of meaning. Note that, in the present study, a manual version of the Semantic Stroop task was used, which prevented further interference rising from the generation of a vocal response (e.g., Kinoshita et al., 2017) and made the task more

similar, and thus comparable, to the other two tasks – all requiring manual categorizations.

To systematically explore the overlap in cognitive control mechanisms across our three tasks, we implemented a series of analyses that could be informative on different aspects of executive control. As a first step, we evaluated the presence of the three effects typically reported in the three tasks – i.e., the taboo-interference, the Simon effect, and the semantic Stroop effect – and the correlation between them. Importantly, the taboo-interference effect was investigated considering the impact of habituation. It is well-known that, as the experiment progresses, the reaction to taboo stimuli is reduced due to habituation (e.g., Bertels & Kokinsky, 2016; Harris & Pashler, 2004; MacKay et al., 2004; 2015). In order to consider potential reductions of taboo interference due to habituation, trial history was modeled when analysing lexical decision data and in the correlation analyses with other tasks. With this first step, we expected to have a first rough indication of similarities among the processing at work in the three tasks. At this general level, whereas a correlation between the taboo-interference and the semantic Stroop effect is conceivable, due to the shared linguistic and semantic origin, the one between taboo interference and the Simon effect seems unlikely. A more detailed approach was thus necessary to better assess peculiar control mechanisms and, in particular, the recruitment of selective suppression for prepotent but task-irrelevant information.

As a second step, we implemented a delta-plot analysis (e.g., de Jong et al., 1994) and, for each experiment, we looked at how the different effects unfolded as a function of response speed. As mentioned, the Simon effect is known to be larger in the fastest responses and reverse in the slowest ones (for a review, see van den Wildenberg et al., 2010): Cognitive control counteracts incongruent information via selective suppression, a mechanism that needs time to be fully implemented. A similar pattern has been reported for the taboo interference in lexical



decision (Scaltritti et al., 2021). Further, previous evidence suggests that the semantic Stroop interference may have a limited impact on the slower tail of the RTs distribution (White et al., 2016). We thus focused our correlational analyses on the last segment of the delta-plots for the different effects. We reasoned that the slope of this last segment should capture the recruitment of selective suppression across tasks: A correlation between the slopes of the last segments across tasks may speak in favor of a domain-general mechanisms, whereas a more selective correlation between taboo and semantic Stroop interference would favor a domain-specific implementation of selective suppression.

## **Method**

### **Participants**

One-hundred and two participants took part in the experiments (46 females, mean age = 28.29; SD = 8.50). Eleven participants were recruited from direct contacts of the experimenters and 91 participants via the research platform Prolific Academic and rewarded with £5. Participants were all Italian native speakers and reported normal or corrected-to-normal vision and no history of learning disabilities. Twenty-two participants were removed from the sample either because their average response accuracy in one of the tasks was below the (arbitrary) threshold of .7 ( $N = 10$ ), or because too many trials were missing in the final datafile (mean number of missing trials for these participants = 214.42, SD = 135.47) due to problems in data-transfer ( $N = 12$ ). The final sample was thus composed of 80 participants (35 females, mean age: 28.16, SD = 8.21). The study was approved by the Ethical Committee of the University of Milano-Bicocca (protocol n.: RM-2020-279) in accordance with the standards of the Declaration of Helsinki.

### **Stimuli**

### *Lexical Decision task*

Ninety taboo words (selected from ITABOO, Sulpizio et al., 2020) and 90 non-taboo words (selected from Italian adaptation of the Affective Norms for English Words, Montefinese et al., 2014) were selected. Taboo words were socially-inappropriate words and belonged to the domains of sexuality, insults, and scatology/disgust. Non-taboo words were socially appropriate words and belonged to different domains (e.g., objects, animals, furniture). Taboo words and non-taboo words differed in terms of arousal and valence – with Taboo words showing significantly higher arousal and lower valence than Non-taboo words –, but were matched on several psycholinguistic variables (see Table 1).

Ninety taboo pseudowords and 90 non-taboo pseudowords were created by replacing one letter from the word stimuli (first and last letter were never changed). Taboo pseudowords were included in the experiment in order: a) to investigate the effects of taboo content independently of stimulus lexicality, b) not to make the taboo dimension predictive of the response. The two sets of pseudowords were matched on: number of letters, orthographic neighborhood size, and orthographic Levenshtein distance. Pseudowords and words were also matched on these three variables. Psycholinguistic properties of the stimuli are listed in Table 1.

Stimuli were divided in two subsets. In this way, all stimuli (either as words or as pseudowords) appeared equally often across participants and, for each participant, each stimulus appeared only once (either as word or pseudoword). Each subset consisted of 90 words – 45 taboo and 45 non-taboo words – and 90 pseudowords – 45 taboo and 45 non-taboo pseudowords. Within each subset, taboo and non-taboo stimuli, as well as words and pseudowords, were comparable in terms of the same psycholinguistic properties reported above. The same applies for the comparison between the two subsets. Half of the participants saw the first subset, and the

other half the second one.

### ***Semantic Stroop task***

Four possible response colors were selected: Green (RGB 0,155,0), red (RGB 255,0,0), blue (RGB 0,170,255), and yellow (RGB 255,255,0). Four color-associated words were selected on the basis of their perceptual association with the response colors: *Lawn*, *strawberry*, *sky*, and *lemon*. Four control words not associated to the response colors were also selected: *Chain*, *stage*, *crater*, and *house*. Color-associated and control words were comparable in terms of: Frequency, number of letters, orthographic neighborhood size, and orthographic Levenshtein distance (see Table 1). Words and colors did not share their initial phonemes.

Each color-associated word was presented in all the three unassociated colors (e.g., *strawberry* was presented in green, blue, and yellow). Similarly, each control word was presented in three colors. Each word was presented in each color 12 times for a total of 288 trials (12 repetition x 8 words x 3 colors each).

### ***Simon task***

One green (RGB 0,128,0) and one red (RGB 255,0,0) square were used. Each square was presented 144 times, half of the times in a spatially-compatible position with that of the response, and half of the time in a spatially-incompatible position (for further details see Apparatus and Procedure).

**Table 1**

*Psycholinguistic properties of the stimuli used in the lexical decision and the semantic Stroop experiment.*

Lexical Decision									
Variables	Words			Pseudowords			Words	PWs	<i>p</i>
	Taboo	Neutral	<i>p</i>	Taboo	Neutral	<i>p</i>			
Frequency (log)	5.86	6.07	>.46	-	-	-	-	-	-
N. of Letters	7.8	8.03	>.36	7.8	8.03	>.36	7.91	7.91	1
Orth. N.	2.93	2.54	>.45	2.52	3.31	>.26	2.73	2.91	>.68
OLD	2.33	2.55	>.47	2.44	2.42	>.88	2.28	2.42	>.07
Concreteness	6.31	6.33	>.89	-	-	-	-	-	-
Imageability	6.11	6.36	>.21	-	-	-	-	-	-
Valence	4.01	5.8	<.00	-	-	-	-	-	-
			1						
Arousal	5.11	4.9	.01	-	-	-	-	-	-

Stroop		
Variables	Color-associated words	Control words
Frequency (log)	4.06	4.04
N. of Letters	5.75	5.75
Orth. N.	5.25	5.75
OLD	1.71	1.64

*Note.* PWs = pseudowords; N. of Letters = number of letters; Orth. N = orthographic neighborhood size; OLD = orthographic Levenshtein distance. Frequency values (log-transformed) and Orthographic Neighborhood Size were taken from the SUBTLEX-IT database (Crepaldi, Keuleers, Manderla, & Brysbaert, 2013). Concreteness, Imageability, Valence and Arousal scores were taken from ITABOO (Sulpizio et al., 2020) and the Italian adaptation (Montefinese et al., 2014) of the Affective Norms for English Words database (ANEW; Bradley & Lang, 1999), for taboo and non-taboo words, respectively.

## **Apparatus and Procedure**

All experiments were programmed with Open Sesame software, version 3.2.8 (Psychology Software Tools, Pittsburgh, PA) and on-line data collection managed with JATOS, version 3.5.3 (Lange, Kühn, & Filevich, 2015). Each participant received a single link that could be accessed only once. At the beginning of the experiment, participants were asked to visualize the experiment in full screen and close all the other windows. Then, they were presented with the informed consent and the study description and were asked whether they wanted proceed, giving their consent, or abandon the study. After acceptance, participants completed two questions collecting age and gender. Then, the first experimental procedure started. Each participant performed all the three tasks. Task order was counterbalanced across participants.

### ***Lexical Decision task***

Participants were instructed to categorize letter strings as words or pseudowords by pressing two buttons of their keyboard (A, L) with their left and right index fingers. Response mapping was counterbalanced across participants. Stimuli were presented in a random order.

Each trial started with a fixation cross, that was presented at the center of the screen and whose duration was randomly selected among 3 alternatives (450, 500, 550 ms). Then, after a short blank screen (150 ms), the stimulus appeared and remained on the screen until the participant's response or for a maximum of 1500 ms. In case the participant was too slow and the stimulus disappeared before her/his response, a feedback message (TOO SLOW!) was displayed for 300 ms. There was a 500 ms inter-stimulus interval. A short practice (8 trials, half words and half pseudowords, none included in the experimental list) preceded the experiment. During the practice a feedback message was displayed in case of either too slow or wrong responses

(ERROR). All stimuli appeared in white Sans font (45 px) on a black background. Feedbacks appeared in white Serif font (45 px).

### *Semantic Stroop task*

Participants were presented with colored words and had to categorize them for the color in which they were written by pressing one of 4 buttons (red: Z; yellow: X; green: N; blue: M). Participants were instructed to use their right and left index and middle fingers, so that they used one finger for each button/color.

The trial structure was identical to that used in the Lexical Decision task, with the only difference that the inter-stimulus interval lasted 800 ms. Trials were administered in two blocks, with a self-terminated break between them. Trial order was randomized. The experiment was preceded by two practice sessions. In the first one, participants were instructed to associate each response button with the corresponding color: Participants were presented with a string of hash-marks and were instructed to classify each string as a function of its color. Each color was presented 4 times, for a total of 16 trials. In the second practice session participants were instructed on the task: They were presented with colored words and were asked to categorize each word for its color. Four words were used, each presented in 3 colors. There was a total of 36 practice trials (4 words x 3 colors x 3 repetitions each). In order to facilitate the color-response association, along the entire duration of the two practice sessions, 4 small colored squares were displayed in the bottom part of the screen: A red and a yellow square were displayed on the left in correspondence with their associated response buttons (Z,X), whereas a green and a blue square were displayed on the right in correspondence with their associated response buttons (N,M) (for the same procedure, see Kinoshita et al., 2018). In both practice sessions, a feedback message was displayed for 300 ms in case of a too slow (TOO SLOW!) or a wrong response

(ERROR).

All stimuli appeared in Sans font (45 px) on a black background. Feedbacks appeared in white Serif font (45 px).

### ***Simon task***

Participants were instructed that, during the experiment, they would see colored squares presented at the right and left side of the screen and had to categorize them on the basis of their color by pressing the corresponding button (A or L) on the keyboard, using their left and right index fingers, respectively. Response mapping was counterbalanced across participants.

The trial structure was identical to that used in the Lexical Decision task. Trials were administered in two blocks, with a self-terminated break between them. Each block contained 144 trials: In half trials the square appeared on the left side of the screen and in half trials on the right side of the screen (the distance was set to 400 px from the center). Moreover, for each position, half of the trials were compatible (i.e., the response was delivered with the button located on the same side of the stimulus; e.g., stimulus displayed on the left and response button A) and half were incompatible (i.e., response delivered with the button located on the opposite side of the stimulus). A short practice (24 trials) preceded the experiment, during which, as in previous procedures, a feedback message was displayed for 300 ms in case of a too slow (TOO SLOW!) or a wrong response (ERROR).

### **Statistical Analyses**

All the analyses involving RTs were conducted using linear mixed-effects models. Response accuracy was analyzed using generalized mixed-effects models. Analyses were conducted using the lme4 library (version 4\_1.1–21; Bates et al., 2015) in R (R Core Team, 2015). The models included random intercepts for participants and items, whereas fixed effects

(and their interactions) were retained only in those cases in which the exclusion of the term would determine a significant decrease in goodness-of-fit. This was assessed by comparing a model in which the fixed term under examination was present vs. a model in which the fixed term was absent, via likelihood ratio tests. In case any interaction resulted significant, all the lower-order terms involved were retained.

**Analysis 1.** We assessed the effects produced by the experimental manipulations, and tested for potential correlations of the interference effects across tasks, using Spearman's rho. These correlation tests were one-tailed, our a-priori hypothesis concerned positive correlations between different interference effects. To compensate for multiple comparisons, a false-discovery rate correction was applied on the resulting p-values.

**Analysis 2.** We assessed the modulations of interference effects as a function of response-speed. Specifically, RTs within each participant and experimental conditions were partitioned in 5 quantiles. The first quantile would thus include the fastest 20% of the responses from a given participant in a specific condition, the second quantile the next fastest 20% of the responses, and so on until the last quantile, the fifth one, which would include the slowest 20% of the RTs. The variable Quantile was then considered as a fixed effect in the statistical models. Non-linearities were considered by using orthogonal quadratic polynomials in fitting the Quantile variable. The model featuring the quadratic polynomial was assessed against its purely linear homologue, using likelihood ratio tests. Non-linear relationships were retained only if they increased goodness-of-fit.

For the quantile-analyses, we assessed the correlations across tasks in the difference between the interference effects detected in the fifth vs the fourth quantile. This measure essentially captures the slope of the last segment of the delta-plots representing the effect of



interest (i.e., the difference between conditions) as a function of quantile. Tests were unidirectional, as we expected positive correlations across different tasks. A false-discovery rate correction was applied to the p-values.

Across all tasks, responses below 150 ms were considered anticipations and removed from all the analyses (N = 15, 0.02% of the total trials). In the lexical decision task, for 3 stimuli the proportion of correct responses across participants was below the estimated chance level (i.e.,  $< .56$ ). These items were removed from all the analyses.

## Results

### Analysis 1. Overall interference effects and their correlations

#### *Lexical Decision*

The analysis of the RTs highlighted only a significant effect of stimulus lexicality ( $\chi^2 [1] = 160.09, p < .001$ ), with words yielding faster responses compared to pseudowords ( $b = -82.81, SE = 7.31, t = -11.32$ ). The effect of taboo connotation (taboo vs neutral) was not significant ( $\chi^2 [1] = 1.96, p = .16$ ), and there was no significant interaction between taboo connotation and lexicality ( $\chi^2 [1] = 1.14, p = .29$ ). Data are summarized in Figure 1A. Results were however markedly different when considering variations in the effect of taboo connotation as a function of the unfolding across experimental trials. Particularly, there was a 3-way interaction between lexically, stimulus connotation, and trial number ( $\chi^2 [1] = 11.13, p < .001$ ). As shown in Figure 1B, the interaction captures a peculiar pattern in which the difference between taboo and neutral words is mostly present in the first half of the experiment, and reduced in subsequent trials. For pseudowords, instead, the difference between taboo and neutral stimuli seems negligible, even in the first trials of the experiment. Consistently, follow-up analyses revealed a significant Taboo interference for words in the first half of the experimental trials (taboo interference effect:  $M =$

31.88 ms,  $SD = 56.76$  ms,  $\chi^2 [1] = 8.01$ ,  $p = .005$ ), but not in the second half ( $\chi^2 [1] = 0.11$ ,  $p = .74$ ). For pseudowords, the taboo interference was never significant ( $\chi^2$ s  $< .06$ ,  $ps > .8$ ). Also note that the influence of trial number was better captured when the corresponding fixed effect in the model was fitted with a second order polynomial, rather than in purely linear terms ( $\chi^2 [4] = 97.06$ ,  $p < .001$ ). Parameters of this model are listed in Table 2.

In terms of response accuracy, the only significant effect was the one of lexicality ( $\chi^2 [1] = 3.97$ ,  $p = .046$ ), with words displaying a higher chance of correct responses compared to pseudowords ( $b = -0.24$ ,  $SE = .12$ ,  $t = 2.02$ ). Neither the effect of taboo connotation ( $\chi^2 [1] = 1.25$ ,  $p = .26$ ), nor the interaction between taboo connotation and stimulus lexicality were significant ( $\chi^2 [1] < 1$ ,  $p = .998$ ; Figure 1A).

### ***Semantic Stroop***

For RTs, the semantic Stroop effect was significant ( $\chi^2 [1] = 6.98$ ,  $p = .008$ ), with slower responses for color-associated words compared to control ones ( $b = 11.95$ ,  $SE = 3.72$ ,  $t = 3.21$ , semantic Stroop effect:  $M = 11.84$  ms,  $SD = 24.81$  ms). There was no significant interaction between the semantic Stroop effect and the trial number ( $\chi^2 [1] = 0.00$ ,  $p = .95$ ). The semantic Stroop effect failed instead to reach conventional significance in accuracy analyses ( $\chi^2 [1] = 2.74$ ,  $p = .098$ ). The effects are summarized in Figure 1A.

### ***Simon***

The classic Simon effect was replicated in RTs ( $\chi^2 [1] = 406.34$ ,  $p < .001$ ), with slower responses for incompatible compared to compatible trials ( $b = 31.31$ ,  $SE = 1.55$ ,  $t = 20.25$ , Simon effect:  $M = 31.38$ ,  $SD = 19.72$ ). There was no interaction between compatibility and trial number ( $\chi^2 [1] = 0.45$ ,  $p = .50$ ). Responses were significantly less accurate for incompatible compared to the compatible ones ( $\chi^2 [1] = 160.09$ ,  $p < .001$ ;  $b = -0.81$ ,  $SE = 0.06$ ,  $z = -12.35$ ). These effects

are represented in Figure 1A.

**Table 2**

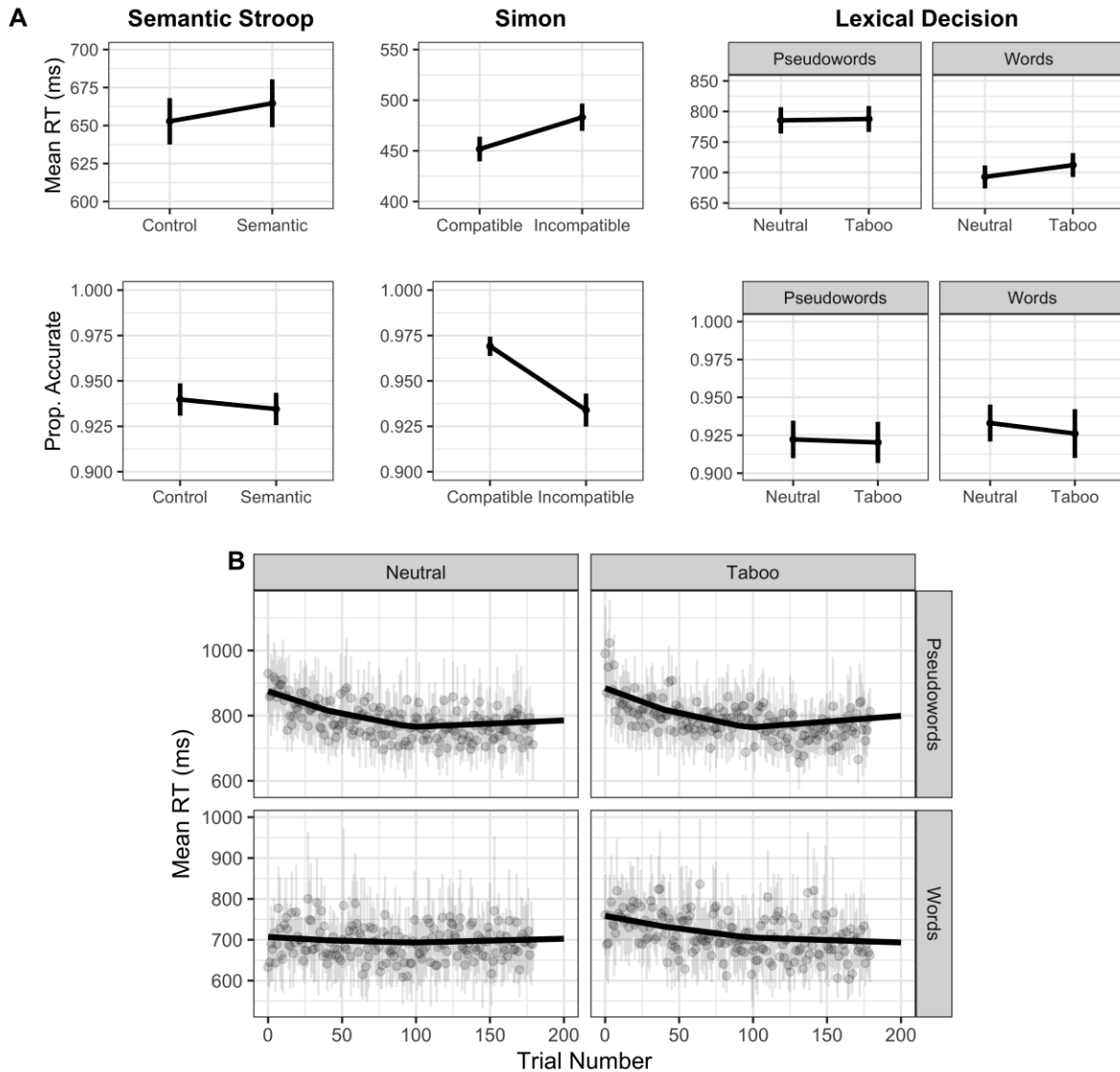
*Parameters of the model on reaction times in lexical decision, considering the fixed effect of trial number.*

Random Effects	Variance	SD	
Stimulus	4242	65.13	
Participant	7665	87.55	
Residual	19908	141.09	
Fixed Effects	b	SE	t-value
Intercept	788.41	12.23	64.457
Lex (word)	-91.64	10.37	-8.84
Connotation (taboo)	2.12	10.34	0.20
Trial N linear	-3463.36	289.68	-11.96
Trial N quadratic	1784	290.84	6.13
Lex X Connotation	15.54	14.64	1.06
Lex X Trial N linear	3194.05	405.28	7.88
Lex X Trial N quadratic	-1470.65	405.47	-3.63
Connotation X Trial N linear	-28.79	409.56	-0.07
Connotation X Trial N quadratic	357.21	411.60	0.87
Lex X Connotation X Trial N linear	-1896.25	573.50	-3.31
Lex X Connotation X Trial N quadratic	-98.22	573	-0.17

*Note.* Lex = Lexicality (pseudowords used as reference); Trial N = trial number; For Connotation, the neutral condition was used as reference.

**Figure 1**

Results of the analyses on overall effects.



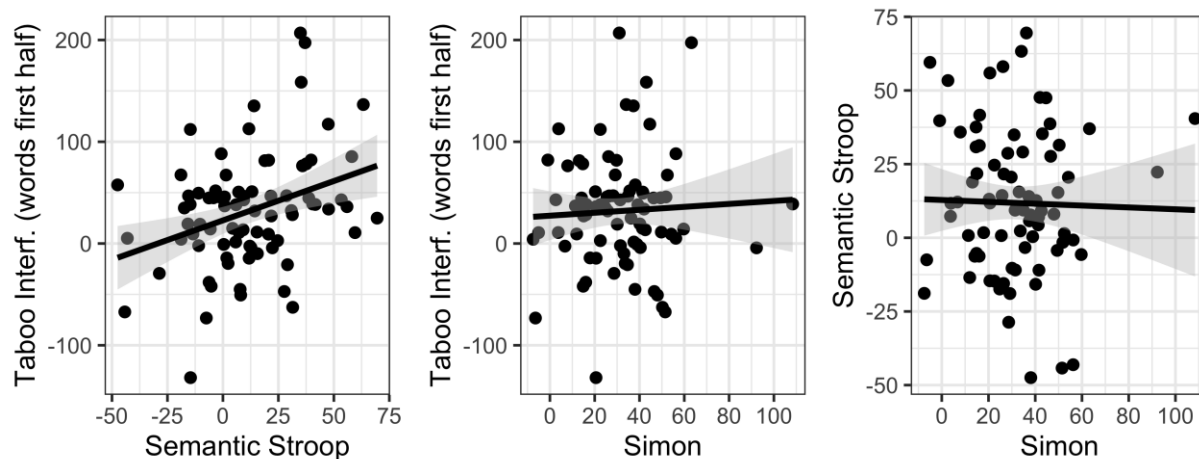
*Note.* A: mean RTs (first row) and proportion of accurate responses (second row) as a function of experimental condition within each task (columns). Prop. Accurate = mean proportion of accurate responses. B: Mean RT as a function of experimental conditions (the 4 different panels) and trial number (x axis).

### Correlations

We assessed the correlations (Spearman's rho) between the interference effects across the different tasks. For the semantic Stroop task, the effect was computed (within participants) by subtracting the average RTs in the Control condition from the average RTs in the semantic one. Similarly, for the Simon task, we computed (within participants) the difference between average RTs in the incompatible vs compatible conditions. For the lexical decision task, we focused on words presented in the first half of the experiment (where the taboo interference was actually present) and subtracted within participants the average RT for neutral words to the one of taboo words. There was a significant correlation between the semantic Stroop effect and the taboo interference ( $r_s = .25, p = .03$ ). All other correlations were not significant ( $ps > .58$ ). Correlations are represented in Figure 2.

**Figure 2**

*Correlations between interference effects of the three tasks (Semantic Stroop, Simon, and Lexical Decision).*



*Note.* The taboo interference was computed only on word stimuli from the first half of the experimental trials (see text for details).

## **Analysis 2. Quantile Analysis: Interference effects as a function of response-speed**

### ***Lexical Decision***

As the results of Analysis 1 showed a clear difference in the taboo effect as a function of trial number, quantiles were separately computed (within each participant and experimental condition) in the first vs the second half of the experiment. This choice was further justified by the presence of a significant 4-way interaction ( $\chi^2 [1] = 8.91, p = .003$ ) between experimental block (first vs second), lexicality (words vs pseudowords), stimulus connotation (taboo vs neutral), and quantiles (1 to 5). Follow-up analyses were conducted by separately fitting models within the first and the second experimental blocks.

For the first block, the analysis revealed a significant lexicality X connotation X quantile interaction ( $\chi^2 [1] = 7.60, p = .006$ ). Fitting the fixed effect of quantile with a second order polynomial further increased goodness-of-fit ( $\chi^2 [4] = 620.69, p < .001$ ). As shown in Figure 3A, for words the taboo interference tends to grow across the distribution, whereas for pseudowords it's rather negligible throughout all the quantiles. Parameters for this model are listed in Table 5. For the second block, the 3-way interaction between lexicality, stimulus connotation, and quantile was not significant when the fixed effect of quantile was fitted in purely linear terms ( $\chi^2 [1] = 2.20, p = .14$ ). It was instead significant ( $\chi^2 [2] = 7.25, p = .027$ ) when a second order polynomial was used in fitting the variable quantile. This latter model (parameters listed in Table 3) displayed a significantly better goodness-of-fit compared to the solely linear one ( $\chi^2 [4] = 647.75, p < .001$ ). In the second block of the experiment words display a reduction of the taboo interference in the slowest RTs, whereas for pseudowords the difference between taboo and neutral stimuli is negligible throughout the whole RT distribution (Figure 3).

**Table 3***Parameters of the model for the quantile analysis in lexical decision.*

Random effects	First Half		Second Half	
	Var	SD	Var	SD
Stimulus	106.8	10.33	76.72	8.76
Participant	7924.2	89.02	8514.58	92.27
Residual	7160.6	84.62	6322.43	79.514

Fixed effects	b	SE	t	b	SE	t
Intercept	808.90	10.23	79.03	761.67	10.54	72.26
Lex (word)	-113.82	3.35	-33.99	-70.00	3.07	-22.8
Connotation (taboo)	1.96	3.35	0.58	1.436	3.07	0.47
Quant lin	11190.9	174.91	63.98	10046.3	160.98	62.41
Quant quad	2043.68	173.11	11.81	2121.68	158.8	13.36
Lex X Connotation	29.37	4.72	6.22	2.88	4.36	0.66
Lex X Quant lin	-1057.5	245.43	-4.31	-244.68	228.43	-1.07
Lex X Quant quad	273.11	242.24	1.13	224.64	225.90	0.99
Connotation X Quant lin	-332.72	245.91	-1.35	-52.04	228.38	-0.23
Connotation X Quant quad	125.3	242.66	0.52	41.14	225.72	0.18
Lex X Connotation X Quant lin	1005.33	346.39	2.90	-507.24	324.9	-1.56
Lex X Connotation X Quant quad	-265.14	341.44	-0.78	-701.92	320.57	-2.19

*Note.* Lex = Lexicality (psuedowords used as reference); Quant = quantile; lin = linear; quad = quadratic. For Connotation, the neutral condition was used as reference.

### ***Semantic Stroop***

The semantic Stroop effect grows larger in slower RTs (Figure 3A). Consistently, the quantile analyses revealed a significant interaction between the semantic Stroop effect and quantiles ( $\chi^2 [1] = 12.49, p < .001$ ). Considering a second order polynomial when fitting the

variable Quantile yielded a significantly better goodness-of-fit with respect to a model considering solely a linear term ( $\chi^2 [2] = 2780.4, p < .001$ ). Parameters are listed in Table 4.

**Table 4**

*Parameters of the model for the quantile analysis of the Stroop Task.*

Random Effects	Variance	SD	
Color	13.7	3.7	
Participant	4815.3	69.39	
Residual	7131.8	84.45	
Fixed Effects	b	SE	t-value
Intercept	652.82	8.02	81.43
Condition (semantic)	11.84	1.15	10.3
Quantile linear	22862.35	120.1	190.36
Quantile quadratic	4740.35	119.5	39.67
Condition X Quantile linear	628.73	168.95	3.72
Condition X Quantile quadratic	-248.03	168.92	-1.47

*Note.* The control condition was used as reference.

### ***Simon***

The Simon effect seems to display a reduction within slowest RTs (Figure 3A), with a quantile by condition (compatible vs incompatible) interaction approaching conventional significance ( $\chi^2 [1] = 3.22, p = .07$ ). The reduction was better captured when including a second order polynomial for the quantile term of the model, which significantly increased goodness-of-fit ( $\chi^2 [2] = 2220, p < .001$ ). Parameters of the model are reported in Table 5.



**Table 5*****Parameters of the model for the quantile analysis of the Simon Task.***

Random Effects	Variance	SD	
Participant	3303	57.47	
Residual	4503	67.10	
Fixed Effects	b	SE	t-value
Intercept	451.75	6.46	69.96
Condition (incompatible)	31.3	0.91	34.52
Quantile linear	13436.20	93.94	143.02
Quantile quadratic	3650.14	94.01	38.83
Condition X Quantile linear	-263.73	134.23	-1.96
Condition X Quantile quadratic	-891.92	134.23	-6.64

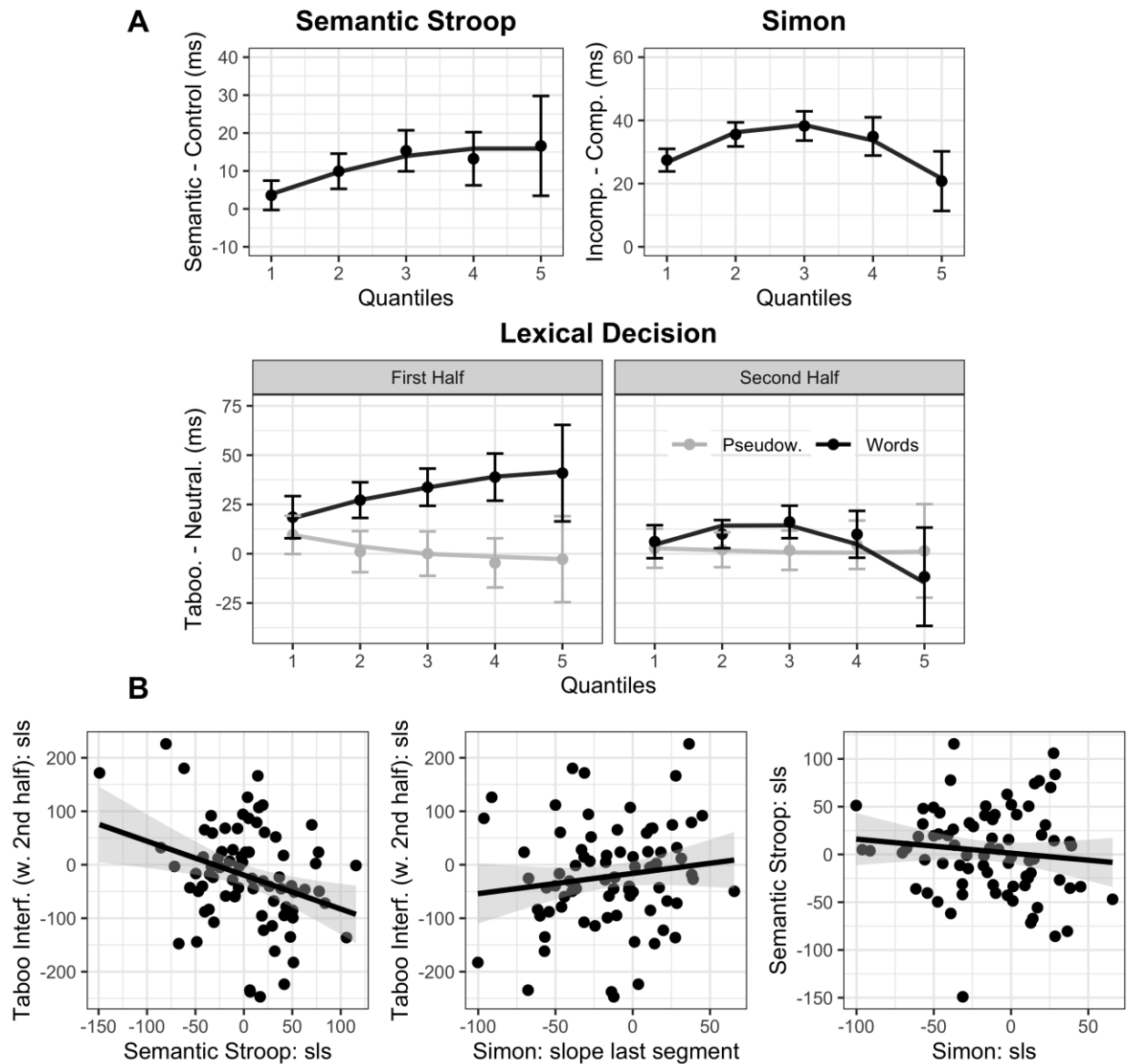
*Note.* The compatible condition was used as reference.

***Correlations***

For each task, we computed, within participants, the difference in the effect of interest (Stroop, Simon, taboo interference) between the fifth and the fourth quantile of the RTs distributions. A positive difference (i.e., a positive slope in the delta-plot of Figure 3A) would thus indicate an increase of the difference between conditions in the slowest RTs, whereas a negative score would signal a reduction of the effect in the slowest portion of the distribution. For the lexical decision task, we selectively considered the slopes of the last segment of the delta-plot computed in case of word stimuli within the second half of the experiment, i.e., where the quantile analysis provided evidence of a suppression mechanism. There was no significant correlation (all  $r_s < |.28|$ ; all  $p_s > .27$ ). Correlations are represented in Figure 3B.

**Figure 3**

Results from the quantile analysis.



*Note.* A: delta-plot of the interference effects detected in the three tasks (Semantic Stroop, Simon, and Lexical Decision). B: Correlations between slopes of the last segment (sls) of the delta-plots computed for the interference effects of the three tasks. Only the taboo interference computed on word stimuli from the second half of the experimental trials was considered (see text for details). Incomp. – Comp. = Incompatible – Compatible; Taboo Interf. = Taboo Interference; w. 2nd half = words, second half of the experiment.

A-posteriori, however, we noted a potentially relevant, albeit unpredicted, inverse correlation between the semantic Stroop and the taboo interference for words in the second half of the experiment ( $r_s = -.28$ ). When all correlations were re-computed using bidirectional tests, this correlation was significant ( $p = .04$ , corrected using false-discovery rate).

### **Discussion**

We investigated whether, in visual word recognition, semantic interference triggers the activation of domain-general or domain-specific control mechanisms. To reach this goal, we adopted a correlational approach to investigate potential similarities between interference effects produced in the context of three different tasks: The lexical decision task with taboo words – which were proven to trigger semantic interference (e.g., Dhooge & Hartsuiker, 2011; Carretti et al., 2008; Sulpizio et al., 2019) –, the Simon task and the Semantic Stroop task – which were used as benchmarks of non-semantic and linguistic-semantic interference, respectively. Correlations were observed selectively between the lexical decision and the semantic Stroop tasks. In the remainder of this section, we begin by discussing the findings at the level of the overall means and their correlations. We then consider the findings from the delta-plot analyses as well as the correlations tested at this level. We conclude by highlighting general implications of our findings for visual word recognition.

In addition to replicating the classic Simon effect – i.e., slower responses for incompatible than compatible trials –, the analyses of the overall means showed a taboo-interference effect – i.e., slower responses for taboo than non-taboo words – and a Semantic Stroop effect – i.e., slower responses for color-associated than control words. The taboo-interference effect for words is in line with previous findings (e.g., Carretti et al., 2008; Madan et al., 2017; Sulpizio et al., 2019), confirming the detrimental effect that taboo connotation has

on word processing. No taboo interference surfaced with pseudowords, contrary to recent empirical evidence (Scaltritti et al., 2021; Sulpizio et al., 2020). This discrepancy between the current and previous studies may be due to a critical difference in the experimental setting. While the previous studies were conducted in a formal context – i.e., a university lab, at the presence of a senior experimenter –, the current experiments were run in an informal one – i.e., online at participants' home, with no contact with the experimenter. Since the effects of taboo connotation are highly sensitive to the context (e.g., Christianson et al., 2017; MacKay et al., 2004; 2015), this may have reduced the impact of taboo connotation, particularly for pseudowords, in which the access to the taboo connotation is mediated by the activation of their base-word. Additionally, the effect of taboo interference – but not the Simon and the Semantic Stroop effect – was subject to habituation, being maximal at the beginning of the experiment and disappearing later on. Habituation to taboo (and emotional) words has been repeatedly reported and has been ascribed to the repetition-induced dampening of the reaction to taboo stimuli (e.g., Bertels & Kolinsky, 2016; Harris & Pashler, 2004; MacKay et al., 2004; 2015).

With respect to the semantic Stroop effect, to the best of our knowledge, our experiment is the first to report evidence of this phenomenon in a transparent language such as Italian. Previous reports were all from opaque languages such as English (e.g., White et al., 2016; Kinoshita et al., 2018) or French (e.g., Augustinova et al., 2018). The replication of the semantic Stroop effect in Italian is not trivial, as it is traditionally assumed that, for transparent languages, orthographic processing may occur with a limited reliance on semantics (e.g., Katz & Frost, 1992; Schmalz et al., 2015). Instead, our finding of a semantic interference suggests that semantic processing may significantly affect word recognition even in languages with a high print-to-sound consistency (e.g., Job et al., 1998; Tabossi & Laghi, 1992; Wilson et al., 2013).

When considering these general effects, there was no correlation between the Simon effect and the semantic Stroop effect, confirming the pattern reported in the executive control literature (e.g., Fan et al., 2003; Stins et al., 2005; Shilling et al., 2002). More interestingly, the taboo-interference effect positively correlated with the semantic Stroop effect, but not with the Simon effect. We argue that the driving force behind this asymmetrical pattern of correlations is related to the presence/absence of semantic processing across the different tasks. Particularly, this finding seems to corroborate the notion that both the semantic Stroop and the taboo interference effect share a common origin at the level of semantic processing.

Although seminal perspectives suggest that the semantic Stroop effect arises for a conflict at the level of response selection (e.g., Roelofs, 2003), i.e., the distractor produces interference because of its association with the response set colors, we endorse a strictly semantic locus of the phenomenon. This choice relies on evidence coming from the extant literature as well as on the results reported in the current study. An important line of evidence comes from recent studies that have explicitly tried to tease apart the different sources of conflict occurring during the Stroop tasks, that are task conflict (e.g., *DOG* in green – *XXX* in green), semantic conflict (e.g., *SKY* in green – *DOG* in green) and response conflict (e.g., *BLUE* in green – *SKY* in green). Using both verbal and manual Stroop tasks, Augustinova, et al. (2018) showed that only semantic conflict was identical in the two task versions – task conflict was present in verbal Stroop only, and response conflict was smaller in the manual Stroop (the insensitivity of semantic conflict to task modality was also reported by other researchers, e.g., Brown & Besner, 2001; Kinoshita et al., 2018). The same pattern (semantic conflict: manual Stroop = verbal Stroop; response conflict: manual Stroop < verbal Stroop) was reported by Augustinova et al. (2019), who also found similar semantic facilitation (*SKY* in blue < *DOG* in blue) in both task modalities, an

evidence inconsistent with the view that the semantic Stroop effect would measure response conflict. As the authors state: “if semantic effects were due to connections at the response level, one would expect to see simultaneous modification of the semantic- and response-level effects” (p. 11). Finally, also the selective correlation we report between the taboo interference effect and the semantic Stroop effect seems to point toward the same direction. In our lexical decision experiment any form of response conflict is absent given the task configuration, in which the taboo connotation is not predictive of stimulus lexicality. Thus, the correlation between the two tasks indicates a shared semantic origin, with the activation of prepotent semantic information hindering the performance across different task configurations.

With respect to this perspective, we should however acknowledge a potential limitation. In fact, the variance associated with the semantic Stroop and the Simon effect was markedly different, being larger for the former compared to the latter. We cannot thus exclude this difference may have contributed to the asymmetry in terms of the presence/absence of correlation between the taboo-interference effect on the one hand, and the semantic Stroop and Simon effect on the other. To shed light on the nature of these interference effects, it thus becomes even more cogent to specifically test for potential similarities and differences with respect to the control mechanisms that they would trigger to compensate for their detrimental influence.

Specifically, in order to identify the involvement of domain-general control mechanisms across the tasks, we assessed the correlations in the slopes of the final segment of the delta-plots, corresponding to the difference between the interference effects detected in the fifth vs the fourth quantile. A negative slope has been repeatedly reported for the Simon effect and is interpreted as a marker of selective suppression (e.g., Burle et al., 2002; Ridderinkhoff, 2002; for review, van

den Wildenberg et al., 2010), a mechanism that dampens prepotent yet task-irrelevant information. The same negative slope was evident for the taboo-interference effect, at least when considering the second half of the lexical decision experiment. This potentially suggests that the habituation effect, i.e. the vanishing of the taboo interference in the second half of the trials, may in part reflect an explicit suppression mechanism that, during the course of the experiment, is implemented to attenuate the detrimental effect of taboo connotation. We maintain, in fact, that although the Simon effect and the interference exerted by taboo words are clearly different phenomena, the presence of a similar negative slope in the two tasks – a pattern rarely reported in visual word recognition (but see Scaltritti et al., 2021, for a similar pattern with taboo stimuli) – suggests some similarity in the way cognitive control deals with the two types of interference.

Importantly, however, there seems to be no correlation between the slopes of the last segment of the delta plots for the two tasks. This finding challenges the notion of a domain-general selective suppression that operates across multiple types of representations, irrespective of the differences in terms of stimuli and tasks configurations. Instead, we speculate that specific suppression mechanisms are differentially implemented in the Simon and in the lexical decision task, in order to counteract prepotent and interfering information as a function of the context and the to-be-suppressed content.

When turning to the semantic Stroop task, one first thing to note is that the unfolding of the semantic Stroop effect as a function of response speed followed an unexpected direction. In fact, White et al. (2016) applied ex-Gaussian analyses to the semantic Stroop effect and showed that the semantic Stroop interference yields a distributional shift towards a slower range of latencies, without a specific involvement of the slower tail of the RTs distribution. Differently, our results show an increase of the semantic Stroop effect in the slowest quantiles, as indexed by

the positive slope between the fourth and the fifth quantile. This discrepancy might arise from a significant methodological difference between the two studies, that is the use of a vocal vs. manual version of the Stroop task (e.g., Sharma & McKenna, 1998). Consistently, Hasshim et al. (2019) recently reported a similar pattern in a manual version of the Stroop task, where semantic interference was selectively indexed by color-words (e.g., green) that are not part of the response set (e.g., white, blue, orange). It thus seems that in the manual version of the paradigm, the semantic component of the Stroop interference is enhanced in slower RTs.

Within standard Stroop tasks, the involvement of the slower tail of the RTs distribution has been considered as a selective index of task-conflict (e.g., Steinhauser & Hübner, 2009). However, in our version of the task, task conflict is arguably controlled across conditions, given that the semantically interfering stimuli (color-related words) and the control ones (color-unrelated words) are both words and hence comparable with respect to task conflict (Kinoshita et al., 2018). Differently, one possibility is that the enhancement of the semantic Stroop interference in the slowest RTs reflects the inability to consistently deploy inhibitory mechanisms due to fluctuant attentional efficiency in controlling interfering information (e.g., de Jong et al., 1999). When attention is properly focused on task goals and related schemas, responses are fast and irrelevant distracting information is successfully inhibited. In contrast, when attention operates less efficiently, responses are slower and more prone to interference. In this context, the manual version of the semantic Stroop may offer more room for lapses of attentional control, possibly because of the highly arbitrary and artificial mapping between stimulus (color) and responses (button-press), compared to a situation in which vocal responses are required. In turn, this would alter the distributional profile of the semantic Stroop effect, with an enhancement in the slower tail of the RTs distribution signaling failures, on these trials, to



consistently maintain task goals and related schemas.

Importantly, the positive slope of the last segment in the delta-plot of the semantic Stroop effect was correlated with the negative slope of the last segment for the taboo-interference effect (observed in the second half of the experiment). Participants showing stronger semantic interference in the slower latencies of the Stroop task were also displaying an attenuated taboo interference in the slower latencies of the second half of the lexical decision task. A speculative (and post-hoc) way to explain this finding is to consider the role that the interfering semantic information has in the two tasks. In the semantic Stroop the interfering information (i.e., the color associated with the word distractor) pertains to the task-relevant semantic domain. Differently, in the lexical decision task the interfering information (i.e., the taboo connotation of the word meaning) is completely irrelevant and has no overlap with the response dimension (taboo connotation does not overlap with stimulus lexicality).

Assuming that participants vary in their ability to access word meaning (e.g., Pexman & Yap, 2018), the more easily semantic information is available, the more the need to block it when judged irrelevant for the task's goal, as in the case of taboo connotation in the lexical decision task, where the reduction of the taboo interference in the slowest quantiles suggests the deployment of semantic suppression. In the semantic Stroop task, a higher semantic availability would again make the interfering information – i.e., the color associated with the word – more easily activated. However, in this task configuration the interfering semantic dimension (color) is harder to block, because of its relevance in the task's configuration, with the consequence that semantic suppression is not implemented. As such, the availability of semantic information would determine an enhanced interference, particularly in those cases in which the task set is less efficiently maintained (i.e., slower trials). This interpretation is in line with a recent perspective

outlined by Kinoshita and colleagues (2017; 2018; see also Norris, 2006), suggesting that semantic processing in the Stroop task is goal-directed, and semantic features are activated (and thus play a role) as a function of task requests. Thus, for example, in a Stroop task, words not associated with a specific color (e.g., HAT) yield the same interference as nonwords (e.g., HIX), because their semantic features are not diagnostic of colors (Kinoshita et al., 2017). Similarly, as originally reported by Klein (1964), color words that are not part of the response set (e.g., *white*) triggers more interference compared to color-associated words (e.g., *lemon*), as the semantic features of the formers wholly pertain to the task-relevant semantic dimension of colors, whereas for the latter not all the semantic features pertain to the color domain (Kinoshita et al., 2018). This reasoning seems to resonate with our speculative proposal that semantic control mechanisms, such as semantic suppression, are goal-directed and task dependent. They do not operate in the context of the semantic Stroop task as the task-relevant color dimension cannot be blocked.

This proposal calls for additional research. Although the correlational approach is widely used in the literature to assess the role of executive control in language processing (e.g., Boned et al., 2021; Crowther & Martin, 2014; Korko et al., 2021), it comes with notable limitations, as correlations might surface due to unidentified factors shared across tasks. Albeit in our data this concern may be mitigated by the selectivity of the reported correlations, which consistently involved only semantically-driven effects, future studies may further attempt to directly manipulate the overlap between the semantic dimension triggering the interference and the one that is relevant for the task goals, in order to assess corresponding variations in the implementation of (semantic) control mechanisms.

An additional way to further refine our hypothesis would be to exploit complementary

sources of evidence provided by EEG. Research focusing on event-related potentials (ERPs), has highlighted the involvement of the N2 component, a fronto-central negativity peaking around 200-500 ms after stimulus presentation (e.g., Folstein & Van Petten, 2008), across a variety of tasks triggering reactive control in response to conflicting and/or interfering information (for a more general review, see Gratton et al., 2018). The literature on the Stroop task, instead, has consistently reported so-called N450 effects, stemming from an enhanced medial frontal negativity for incongruent compared to color neutral trials, peaking around 450 ms after stimulus presentation (e.g., Liotti et al., 2000). Importantly, this component has been linked with purely semantic conflict in the context of semantic Stroop effects (e.g., Augustinova et al., 2015). Thus, it remains to be investigated how different indexes, potentially pointing to different conflict-processing dynamics, are affected in the context of the lexical decision task, which strictly dissociates the semantic dimension responsible of the interference (e.g., taboo connotation) from the response-relevant information about stimulus lexicality, thus avoiding both semantic and response conflict.

In conclusion, the present findings offer initial evidence that, during visual word recognition, the control of interfering semantic information may require the involvement of a specific control mechanism, as suggested by the presence of selective correlations between the two tasks requiring semantic processing. This similarity is not trivial as it shows up despite the huge differences between the two tasks. Additionally, semantic control appears to differentially operate as a function of whether the semantic dimension triggering interference overlaps or not with task- and response-relevant dimension.

The notion of a semantic-specific control mechanism may be hard to reconcile with the more traditional view of the relationship between language and executive control processes, in

which domain-general control mechanisms are assumed to operate on semantic contents and, more in general, on linguistic information in a similar vein as they do for non-linguistic information (e.g., Declerck et al., 2017; Hussey & Novick, 2012; Thomas & Allport, 2000; Ye & Zhou, 2009). These models assume that executive control regulates language processing *via* a control system working at the more general level of goals planning (Dijkstra & van Heuven, 2002; Green, 1998; Green & Abutalebi, 2013). However, if control mechanisms recruited to resolve language-related conflicts are the same as those recruited to control non-linguistic conflicts, then we should expect some sort of correlation in the deployment of executive control across linguistic and non-linguistic tasks.

Our findings, instead, are more in line with recent proposals assuming that domain-general control mechanisms are flanked by distinct semantic-specific control mechanisms (Gao et al., 2020), which are implemented when flexible semantic processing is required to meet the task's goals and constraints (e.g., Davey et al., 2016; Lambon Ralph et al., 2017). According to Lambon Ralph et al. (2017), in fact, there is a graded organization of the semantic control network, with some regions contributing to more domain-general controls and others involved in more specific semantic computations. These latter mechanisms are assumed to operate within the semantic system by guiding activation of semantic knowledge, and to facilitate the interplay between information resulting from semantic activation and domain-general executive control processes (e.g., Davey et al., 2016; Hoffman et al., 2018; Lambon Ralph et al., 2017). Semantic-specific control mechanisms may thus intervene during visual word recognition whenever a reader has to face prepotent semantic information interfering with the accomplishment of the task's goal, such as the case of taboo connotation in lexical decision performance.

**Declaration of Conflicting Interests**

The authors declare that there is no conflict of interest.

## References

- 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., Parris, B. A., & Ferrand, L. (2019). The loci of stroop interference and facilitation effects with manual and vocal responses. *Frontiers in psychology, 10*, 1786.
- 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*, 543-549.
- Augustinova, M., Silvert, L., Spatola, N., & Ferrand, L. (2018). Further investigation of distinct components of Stroop interference and of their reduction by short response-stimulus intervals. *Acta Psychologica, 189*, 54-62.
- Balota, D. A., Paul, S. T., & Spieler, D. H. (1999). Attentional control of lexical processing pathways during word recognition and reading. *Language Processing, 15*-57.
- Balota, D., & Yap, M. (2006). Attentional control and flexible lexical processing: Explorations of the magic moment of word recognition. In S. Andrews (Ed.), *From inmarks to ideas: Current Issues in Lexical Processing* (pp. 229–258). Hove, UK: Psychology Press.
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., ... & Bolker, M. B. (2015). Package 'lme4'. *Convergence, 12*, 2.
- Belke, E., & Stielow, A. (2013). Cumulative and non-cumulative semantic interference in object naming: Evidence from blocked and continuous manipulations of semantic context. *Quarterly Journal of Experimental Psychology, 66*, 2135-2160.
- Bertels, J., & Kolinsky, R. (2016). Disentangling fast and slow attentional influences of negative and taboo spoken words in the emotional Stroop paradigm. *Cognition and Emotion, 30*,

1137-1148.

- 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, 162-168.
- Boned, J., Cardona, G., Jefferies, E., & Hernández, M. (2021). The influence of language dominance and domain-general executive control on semantic context effects. *Language, Cognition and Neuroscience*, 1-20.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108, 624.
- Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings* (Vol. 30, No. 1, pp. 25-36). Technical report C-1, the center for research in psychophysiology, University of Florida.
- Brown, M., & Besner, D. (2001). On a variant of Stroop's paradigm: Which cognitions press your buttons? *Memory & cognition*, 29, 903-904.
- Burle, B., Possamaï, C. A., Vidal, F., Bonnet, M., & Hasbroucq, T. (2002). Executive control in the Simon effect: an electromyographic and distributional analysis. *Psychological Research*, 66, 324-336.
- Carretié, L., Hinojosa, J. A., Albert, J., López-Martín, S., De La Gándara, B. S., Igoa, J. M., & Sotillo, M. (2008). Modulation of ongoing cognitive processes by emotionally intense words. *Psychophysiology*, 45, 188-196.
- Chiou, R., Humphreys, G. F., Jung, J., & Ralph, M. A. L. (2018). Controlled semantic cognition relies upon dynamic and flexible interactions between the executive 'semantic control' and hub-and-spoke 'semantic representation' systems. *Cortex*, 103, 100-116.

- Christianson, K., Zhou, P., Palmer, C., & Raizen, A. (2017). Effects of context and individual differences on the processing of taboo words. *Acta psychologica, 178*, 73-86.
- Crepaldi, D., Keuleers, E., Mandera, P., & Brysbaert, M. (2013). SUBTLEX-IT. Retrieved from <http://crr.ugent.be/subtlex-it/>
- Crowther, J. E., & Martin, R. C. (2014). Lexical selection in the semantically blocked cyclic naming task: the role of cognitive control and learning. *Frontiers in human neuroscience, 8*, 9.
- Davey, J., Thompson, H. E., Hallam, G., Karapanagiotidis, T., Murphy, C., De Caso, I., ... & Jefferies, E. (2016). Exploring the role of the posterior middle temporal gyrus in semantic cognition: Integration of anterior temporal lobe with executive processes. *Neuroimage, 137*, 165-177.
- De Jong, R., Berendsen, E., & Cools, R. (1999). Goal neglect and inhibitory limitations: Dissociable causes of interference effects in conflict situations. *Acta Psychologica, 101*, 379–394.
- De Jong, R., Liang, C.-C., & Lauber, E. (1994). Conditional and unconditional automaticity: A dual process model of effects of spatial stimulus-response correspondence. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 731-750.
- Declerck, M., Grainger, J., Koch, I., & Philipp, A. M. (2017). Is language control just a form of executive control? Evidence for overlapping processes in language switching and task switching. *Journal of Memory and Language, 95*, 138-145.
- Dhooge, E., & Hartsuiker, R. J. (2011). How do speakers resist distraction? Evidence from a taboo picture-word interference task. *Psychological Science, 22*, 855-859.



- Dijkstra, A. F. J., & van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language & Cognition*, *5*, 175-197.
- Fan, J., Flombaum, J. I., McCandliss, B. D., Thomas, K. M., & Posner, M. I. (2003). Cognitive and brain consequences of conflict. *Neuroimage*, *18*, 42-57.
- Folstein, J. R., and Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology*, *45*, 152–170.
- Gao, Z., Zheng, L., Chiou, R., Gouws, A., Krieger-Redwood, K., Wang, X., ... & Jefferies, E. (2020). Distinct and Common Neural Coding of Semantic and Non-semantic Control Demands. *bioRxiv*.
- Geer, J. H., & Bellard, H. S. (1996). Sexual content induces delays in unprimed lexical decisions: Gender and context effects. *Archives of Sexual Behavior*, *25*, 379–395.
- Gratton, G., Cooper, P., Fabiani, M., Carter, C. S., & Karayanidis, F. (2018). Dynamics of cognitive control: Theoretical bases, paradigms, and a view for the future. *Psychophysiology*, *55*, e13016.
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*, 67-81.
- Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*, 515-530.
- Harris, C. R., & Pashler, H. (2004). Attention and the processing of emotional words and names: Not so special after all. *Psychological Science*, *15*, 171-178.
- 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*, 231-238.

Hoffman, P., McClelland, J. L., & Lambon Ralph, M. A. (2018). Concepts, control, and context: A connectionist account of normal and disordered semantic cognition. *Psychological Review*, *125*, 293-328.

Hussey, E. K., & Novick, J. M. (2012). The benefits of executive control training and the implications for language processing. *Frontiers in Psychology*, *3*, 158.

Job, R., Peressotti, F., & Cusinato, A. (1998). Lexical effects in naming pseudowords in shallow orthographies: Further empirical data. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 622-630.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In *Advances in Psychology* (Vol. 94, pp. 67-84). North-Holland.

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*, 369.

Kinoshita, S., & Lupker, S. J. (2002). Effects of filler type in naming: Change in time criterion or attentional control of pathways? *Memory & Cognition*, *30*, 1277-1287.

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*, 1730-1581.

- Klein, G. S. (1964). Semantic power measured through the interference of words with color-naming. *The American Journal of Psychology*, *77*, 576-588.
- Korko, M., Coulson, M., Jones, A., & de Mornay Davies, P. (2021). Types of interference and their resolution in monolingual word production. *Acta Psychologica*, *214*, 103251.
- Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, *18*, 42-55.
- Lange, K., Kühn, S., & Filevich, E. (2015). “Just Another Tool for Online Studies” (JATOS): An Easy Solution for Setup and Management of Web Servers Supporting Online Studies. *PloS one*, *10*, e0130834.
- Liotti, M., Woldorff, M. G., Perez III, R., & Mayberg, H. S. (2000). An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia*, *38*, 701-711.
- Madan, C. R., Shafer, A. T., Chan, M., & Singhal, A. (2017). Shock and awe: Distinct effects of taboo words on lexical decision and free recall. *Quarterly Journal of Experimental Psychology*, *70*, 793-810.
- MacKay, D. G., Johnson, L. W., Graham, E. R., & Burke, D. M. (2015). Aging, emotion, attention, and binding in the taboo Stroop task: data and theories. *International Journal of Environmental Research and Public Health*, *12*, 12803-12833.
- MacKay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., & Dyer, J. R. (2004). Relations between emotion, memory, and attention: Evidence from taboo Stroop, lexical decision, and immediate memory tasks. *Memory & Cognition*, *32*, 474-488.

- Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2014). The adaptation of the affective norms for English words (ANEW) for Italian. *Behavior Research Methods*, *46*, 887-903.
- Montefinese, M., Hallam, G., Thompson, H. E., & Jefferies, E. (2020). The interplay between control processes and feature relevance: Evidence from dual-task methodology. *Quarterly Journal of Experimental Psychology*, *73*, 384-395.
- Neely, J. H., & Kahan, T. A. (2001). *Is semantic activation automatic? A critical re-evaluation*. In H. L. Roediger III, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *Science conference series. The nature of remembering: Essays in honor of Robert G. Crowder* (p. 69-93). American Psychological Association.
- Norris, D. (2006). The Bayesian reader: explaining word recognition as an optimal Bayesian decision process. *Psychological review*, *113*, 327-357.
- O'Malley, S., & Besner, D. (2008). Reading aloud: Qualitative differences in the relation between stimulus quality and word frequency as a function of context. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1400-1411.
- Pexman, P. M., & Yap, M. J. (2018). Individual differences in semantic processing: Insights from the Calgary semantic decision project. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*, 1091-1112.
- Piai, V., Roelofs, A., Acheson, D. J., & Takashima, A. (2013). Attention for speaking: domain-general control from the anterior cingulate cortex in spoken word production. *Frontiers in Human Neuroscience*, *7*, 832.
- Pratte, M. S., Rouder, J. N., Morey, R. D., & Feng, C. (2010). Exploring the differences in

distributional properties between Stroop and Simon effects using delta plots. *Attention, Perception, & Psychophysics*, 72, 2013-2025.

Reynolds, M., & Besner, D. (2005). Contextual control over lexical and sublexical routines when reading English aloud. *Psychonomic Bulletin & Review*, 12, 113-118.

Ridderinkhof, R. K. (2002). Micro-and macro-adjustments of task set: activation and suppression in conflict tasks. *Psychological Research*, 66, 312-323.

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

Scaltritti, M., Balota, D. A., & Peressotti, F. (2013). Exploring the additive effects of stimulus quality and word frequency: The influence of local and list-wide prime relatedness. *Quarterly Journal of Experimental Psychology*, 66, 91-107.

Scaltritti, M., Job, R., & Sulpizio, S. (2021). Selective suppression of taboo information in visual word recognition: Evidence for cognitive control on semantics. *Journal of Experimental Psychology: Human Perception & Performance*. Advance online publication.  
<https://doi.org/10.1037/xhp0000917>

Schmalz, X., Marinus, E., Coltheart, M., & Castles, A. (2015). Getting to the bottom of orthographic depth. *Psychonomic Bulletin & Review*, 22, 1614-1629.

Sharma, D., & McKenna, F. P. (1998). Differential components of the manual and vocal Stroop tasks. *Memory & Cognition*, 26, 1033-1040.

Shilling, V. M., Chetwynd, A., & Rabbitt, P. M. A. (2002). Individual inconsistency across measures of inhibition: An investigation of the construct validity of inhibition in older

adults. *Neuropsychologia*, *40*, 605-619.

Simon, J. R., & Small Jr, A. M. (1969). Processing auditory information: interference from an irrelevant cue. *Journal of Applied Psychology*, *53*, 433-435.

Spagna, A., Mackie, M. A., & Fan, J. (2015). Supramodal executive control of attention. *Frontiers in psychology*, *6*, 65.

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

Stins, J. F., Polderman, J. C., Boomsma, D. I., & de Geus, E. J. (2005). Response interference and working memory in 12-year-old children. *Child Neuropsychology*, *11*, 191-201.

Sulpizio, S., & Navarrete, E. (2020). Outgroup faces hamper word recognition. *Psychological Research*, *84*, 2300-2310.

Sulpizio, S., Pennucci, E., & Job, R. (2020). The impact of emotional content on pseudoword recognition. *Psychological Research*. Advanced online publication.  
<https://doi.org/10.1007/s00426-020-01454-6>

Sulpizio, S., Toti, M., Del Maschio, N., Costa, A., Fedeli, D., Job, R., & Abutalebi, J. (2019). Are you really cursing? Neural processing of taboo words in native and foreign language. *Brain and language*, *194*, 84-92.

Sulpizio, S., Vassallo, E., Job, R., & Abutalebi, J. (2020). ITABÙ: Dati preliminari di un database delle parole tabù per l'italiano [ITABOO: Preliminary data for an Italian database for taboo words]. *Giornale Italiano di Psicologia*, *2*, 599-614.

- Tabossi, P., & Laghi, L. (1992). Semantic priming in the pronunciation of words in two writing systems: Italian and English. *Memory & Cognition*, *20*, 303-313.
- Thomas, M. S., & Allport, A. (2000). Language switching costs in bilingual visual word recognition. *Journal of Memory and Language*, *43*, 44-66.
- van den Wildenberg, W. P., Wylie, S. A., Forstmann, B. U., Burle, B., Hasbroucq, T., & Ridderinkhof, K. R. (2010). To head or to heed? Beyond the surface of selective action inhibition: a review. *Frontiers in Human Neuroscience*, *4*, 222.
- Ye, Z., & Zhou, X. (2009). Executive control in language processing. *Neuroscience & Biobehavioral Reviews*, *33*, 1168-1177.
- White, D., Risko, E. F., & Besner, D. (2016). The semantic Stroop effect: An ex-Gaussian analysis. *Psychonomic Bulletin & Review*, *23*, 1576-1581.
- Wilson, M. A., Ellis, A. W., & Burani, C. (2012). Age-of-acquisition affects word naming in Italian only when stress is irregular. *Acta Psychologica*, *139*, 417-424.
- Zevin, J.D. & Balota, D.A. (2000). Priming and attentional control of lexical and sublexical pathways during naming. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *26*, 121-135.