Memory & Cognition (2022) 50:898–910
https://doi.org/10.3758/s13421-021-01256-0

Different types of semantic interference, same lapses of attention: Evidence from Stroop tasks

Michele Scaltritti[1] · **Remo Job**[1] · **Simone Sulpizio**[2,3]

Accepted: 1 November 2021 / Published online: 17 January 2022

Abstract

This research investigated the possibility that semantic control mechanisms are recruited only when the interfering semantic information does not overlap with task-relevant semantic dimensions. To reach this goal, we investigated two semantic types of Stroop interference—the semantic and the taboo Stroop effects—and used delta-plots to investigate the role of attentional and semantic control in these two interference phenomena. The semantic Stroop effect, where interference stems from the task-relevant color-related information, was absent in faster responses, whereas it steeply increased in the slowest ones. Contrary to our predictions, the same pattern was detected even for the taboo Stroop interference, with no trace of selective suppression of the interfering semantic connotation, despite its dissociation from any task-relevant semantic dimension. Further, there was a significant correlation between the increase of the two effects in the slowest responses, pointing towards a common underlying processing dynamic. We identified such common background with lapses of executive attention in maintaining task goals and schema, which in turn make the participants performance more prone to interference phenomena. Finally, the absence of any interference effects in the fastest responses suggests that an effective filtering of the distracting word stimuli can be implemented in the context of Stroop paradigms.

Keywords Visual word recognition · Stroop · Semantic control · Cognitive control · Taboo words

Introduction

Access to meaning is the ultimate goal of visual word recognition and reading. In this context, conceptual processing is usually envisaged as the retrieval of semantic features and/or representations from a long-term memory store. Yet the way in which we access this store and the information we retrieve may flexibly change as a function of the contextual goals. According to recent neurocognitive and computational studies on semantic processing, control mechanisms flexibly align conceptual processing with contextual goals and demands (e.g., Hoffman et al., 2018; Lambon Ralph et al., 2017). Here, we investigated the flexibility of semantic control processes by comparing two semantically driven interference effects triggered within the same task, namely the semantic and the taboo Stroop effects. While being similar in terms of their semantic origin and with respect to the overall task configuration, these two types of semantic interference differ with respect to their reliance on the task-relevant semantic dimensions (i.e., information related to the domain of colors), thus potentially offering a test-case for context-driven and goal-driven modulations of semantic control.

In the classic Stroop task, participants are presented with colored strings, and are asked to name the color and ignore the carrier string stimulus. The to-be-ignored stimulus is typically a color word (e.g., *red*). In the standard Stroop effect (Stroop, 1935), responses are slower (and less accurate) when the color–word is incongruent with the color response (e.g., the word *red* written in blue) compared with when the two are congruent (e.g., the word *red* written in red). According to a recent line of research, the detrimental influence exerted by the incongruent carrier words may vary as a function of the contextual factors, such as the proportion of neutral (e.g., Goldfarb & Henik, 2007; Kinoshita et al., 2018; Spinelli & Lupker, 2020a) or incongruent trials included in the experimental list

✉ Simone Sulpizio
simone.sulpizio@unimib.it

1 Department of Psychology and Cognitive Science, University of Trento, Trento, Italy

2 Department of Psychology, University of Milano-Bicocca, Milan, Italy

3 Milan Center for Neuroscience (NeuroMI), University of Milano-Bicocca, Milan, Italy

(e.g., Bugg, 2014; Bugg & Hutchison, 2013; Hutchison, 2011; see also Spinelli et al., 2019; Spinelli & Lupker, 2020b). This suggests that cognitive control and/or attentional mechanisms can intervene to regulate the activation of different sources of information in an adaptive fashion, thus modulating the Stroop interference (but see, e.g., Algom & Chajut, 2019; Schmidt, 2019, for different perspectives).

Moving from these considerations, we selectively focused on the semantic component of the Stroop interference, and investigated the potential unfolding of semantic control processes that may intervene in flexibly weighing conceptual activation as a function of the task context. To focus on the semantic component of the Stroop interference, we implemented Stroop tasks requiring manual responses. This (a) eliminated a first source of interference related to the competition between the color response and the carrier word at the level of phonological encoding (e.g., Kinoshita & Mills, 2020), and (b) minimized the task-conflict stemming from the prepotent association of the word distractors with the task-set of reading (Sharma & McKenna, 1998; see also Kinoshita, De Wit, & Norris, 2017). More crucially, we focused on two versions of the Stroop task exploiting interference effects that seem to more strongly hinge on semantic processing—namely, the semantic and the taboo Stroop.

In the semantic Stroop (Klein, 1964), carrier stimuli include color-associated words, such as *sky* (associated with blue) or *fire* (associated with red). When presented in an incongruent ink color (e.g., *sky* in red), color-associated words produce an interference effect, compared with words that are not associated with a specific color (e.g., *table* in red). The taboo Stroop effect, instead, refers to the interference effect detected when carrier stimuli are socially inappropriate words, compared with neutral control ones (Siegrist, 1995). Although both the semantic and the taboo interference effects require that the carrier word stimulus meaning is processed, the processing dynamics underlying the two semantically triggered effects can be different. The semantic Stroop effect has been considered a marker of semantic conflict, stemming from the activation of conflicting semantic codes in the color dimension. Moreover, it may avoid overlapping contamination from response conflict mechanisms, as carrier word stimuli are not part of the response set (e.g., Neely & Kahan, 2001; see also Augustinova et al., 2018). In contrast, the taboo interference effect has been linked to attentional capture phenomena triggered by task-irrelevant taboo words (e.g., Williams et al., 1996, see also Reynolds & Langerak, 2015), which would divert resources away from the main task. As a first step, we thus exploratively assessed correlations between the two effects across participants, to shed light on the overall overlap in terms of the underlying mechanisms. Given the putatively different processing dynamics underlying the two semantic interference phenomena (semantic conflict vs. attentional capture), potential correlations are not to be taken for granted, even in the context of highly similar tasks.

Then, we addressed the more specific issue concerning how cognitive control may react to these different semantic interferences. In both tasks, the access to the meaning of the task-irrelevant word stimuli (i.e., the color-associated and the taboo words in the semantic and in the taboo Stroop task, respectively) hinders the performance. Therefore, the intervention of control mechanisms may be invoked to control the activation of irrelevant interfering information. However, the semantic and the taboo Stroop task differ in terms of how the interfering information relates to the task-relevant color dimension. Evidence shows that the semantic features related to the task-relevant color dimension trigger more interference compared with features that are not related to colors. For example, as first reported by Klein (1964), color words that are not part of the response set yield stronger interference effects compared with color-associated words, because only for color-words the semantic features are entirely related to the task-relevant color dimension (Kinoshita et al., 2018). Further, words not associated with a specific color (*hat*) yield no more interference than pseudowords (*hix*) because the words semantic features are not diagnostic of colors (e.g., Kinoshita, De Wit, & Norris, 2017). Here we hypothesized that differences in the overlap between interfering and task-relevant semantic information across the taboo vs. the semantic Stroop effects might be coupled with the intervention of different mechanisms of semantic control. To gain a deeper insight into this issue, we resorted to delta-plot analyses (De Jong et al., 1994), in which the two Stroop effects are considered as a function of response speed.

In principle, in the taboo Stroop task, the taboo connotation can be selectively suppressed to control its detrimental influence on the performance. A similar mechanism has been identified in the context of a lexical decision task (Scaltritti et al., 2021). Specifically, the taboo interference showed a reduction, and actually turned to a facilitation, in slower responses. This distributional pattern has been interpreted as a marker of a semantic suppression mechanism that needs time to fully accrue, thus becoming more evident only in the slowest responses. In the current experiment, we explored whether a similar suppression mechanism is invoked even in the context of a Stroop task featuring taboo words.

This sort of semantic suppression may not be invoked in the case of the semantic Stroop configuration. Here, the interfering semantic information activated by color-associated words cannot be suppressed as it pertains to the task-relevant dimension of color. Indeed, in the manual Stroop task, the semantic Stroop effect is enhanced in the slowest responses (Sulpizio et al., 2021). This pattern has been linked to the inability to consistently deploy inhibitory mechanisms due to a fluctuating efficiency of attentional control in maintaining the task goals and schemas, which would become particularly evident in the slowest responses (e.g., De Jong et al., 1999; see also San José et al., 2021; Scaltritti et al., 2015).

Interestingly, we recently observed (Sulpizio et al., 2021) that, when considering the slowest responses, the reduction of the taboo interference effect of the lexical decision task was inversely correlated with the enhancement of the semantic Stroop effect (but not with the reduction of compatibility effects of a Simon task). We speculated that participants for whom semantic information is more promptly available may reactively enhance semantic control via suppression in lexical decision, where the hindering taboo connotation is task irrelevant. Differently, an increased availability of semantic information in the context of the semantic Stroop task, where the task-relevant color dimension cannot be suppressed, would yield an enhanced semantic interference, particularly in the slowest trials where attentional control of the task schema (respond to the color and ignore the word) is less efficiently maintained. Albeit tentative, this speculative explanation suggested that semantic access may be regulated by control mechanisms that are flexibly implemented as a function of task configuration and goals.

In our current experiment, we further investigated the potential differences in the involvement of semantic control mechanisms based on selective suppression as a function of the overlap between the interfering semantic information and the task-relevant semantic dimension. Importantly, here we attempt to investigate this difference while relying on two instantiations of a manual Stroop task—namely, a semantic and a taboo Stroop experiment, rather than by comparing different experimental paradigms. The two types of semantically triggered interference effect may differ with respect to the involvement of semantic control mechanisms and, to investigate this possibility, we assessed the distributional profiles of the semantic and the taboo Stroop effect.

We focused on the slope of the last segment of the delta-plots, which has been consistently used to capture selective suppression (e.g., van den Wildenberg et al., 2010), a mechanism that needs time to fully accrue, and that becomes evident in a negative slope of this segment. This measure should also capture fluctuations in the deployment of attentional control, which are particularly reflected in the slowest latencies (e.g., De Jong et al., 1999). We thus tested for potential correlations across tasks with respect to the slope of the last segment of the delta plots. Albeit the slope may be expected to show a different direction across the two tasks (positive for semantic Stroop, signaling an enhancement of the effect in the slowest responses; negative in the taboo Stroop, signaling a suppression-driven reduction in the slowest latencies), a (inverse) correlation can be nonetheless hypothesized due to the common semantic origin of the interference effects. Participants may vary in their ability to access word meaning (e.g., Pexman & Yap, 2018), thus making the interfering semantic information more or less available during the Stroop task. In line with observations from our previous work (Sulpizio et al., 2021), we predict that the extent to which semantic information becomes available may have different consequences with respect to the mechanisms deployed to overcome semantic interference, as a function of the overlap with the task-relevant semantic dimension. In the semantic Stroop task, where color-related information cannot be blocked, an increased availability of the semantic information may determine an enhanced interference, particularly within trials in which task control is less effectively deployed (i.e., slowest trials). In contrast, in the taboo Stroop task, where the interfering taboo connotation can be suppressed, a higher availability of semantic information may prompt a stronger reliance on inhibitory mechanisms, with the aim to control the detrimental influence of performance. As a result, the same participants showing an enhanced interference in the semantic Stroop task, might also display a stronger reduction of the taboo interference in the taboo Stroop task. Such correlation would support the notion that semantic control mechanisms are flexibly implemented as a function of the task set.

Method

Participants

Ninety Italian native speakers took part in the experiments (33 females, mean age = 25.96 years, *SD* = 5.24, range: 18–39 range). Three participants were recruited among direct contacts of the authors, whereas 87 participants were recruited via the research platform Prolific Academic (Palan & Schitter, 2018), and rewarded with £3.75. All participants reported normal or corrected-to-normal vision and no history of learning disabilities. One participant was excluded from the sample due to the low accuracy in one of the two tasks (proportion of correct responses = .40), and another one because too few trials were retained in the final data file (16, eight for each experimental procedure) due to a failure in data transfer. The final sample thus consisted of 88 participants. The study was approved by the Ethical Committee of the University of Milano-Bicocca (protocol n.: RM-2020-279).

Stimuli

For both the semantic and the taboo Stroop tasks, there were four response colors: Green (RGB: 0,155,0), red (RGB: 255,0,0), blue (RGB: 0,170,255), and yellow (RGB: 255,255,0). For the semantic Stroop task, four color-associated words were selected, *prato* (lawn), *fragola* (strawberry), *cielo* (sky), and *limone* (lemon). These items were selected relying on previous experiments (e.g., Kinoshita et al., 2018), including our own (Sulpizio et al., 2021). Four words not associated with colors were selected to serve as control stimuli. These were *mazzo* (deck), *cratere* (crater), *bagno* (bath), and *salita* (hill). Control words were selected to be matched, as closely as possible, with color-associated words in terms of raw and log lexical frequency, number of letters, number of syllables, orthographic neighborhood size, and orthographic Levenshtein distance (see Table 1). Words and colors (*verde*, *rosso*, *blu*, and *giallo*, in Italian) did not share their initial phonemes. Color-associated words were presented only in combination with unrelated colors (e.g., *strawberry* was presented only in green, blue, and yellow). Likewise, each corresponding control word appeared only in three colors (e.g., *crater* was presented only in green, blue, and yellow).

For the taboo Stroop task, 72 taboo words (from ITA-BOO; Sulpizio et al., 2020) and 72 control words (from Italian adaptation of the Affective Norms for English Words; Montefinese et al., 2014) were selected. Compared with the semantic Stroop task, we selected a larger number of stimuli as taboo-interference is modulated by habituation (e.g., Sulpizio et al., 2021; see also the Appendix). Taboo stimuli consisted of socially inappropriate words referring to the domains of sexuality, insults, severe illness, and disgust. Control stimuli were neutral, socially appropriate words. For both taboo and control words, we tried to avoid stimuli that had an obvious association with specific colors. Taboo and control words differed in terms of arousal and valence, whereas they were comparable with respect to a series of psycholinguistic variables (see Table 1).

Table 1 Psycholinguistic properties of the words in the semantic and in the taboo Stroop experiments

Variables	Semantic Stroop		Taboo Stroop		
	Color-associated	Control	Taboo	Control	*t*-value
Valence	–	–	3.91	5.76	−9.15***
Arousal	–	–	5.11	4.86	2.12*
Concreteness	–	–	5.89	5.60	1.16
Familiarity	–	–	5.74	6.07	−1.58
Imageability	–	–	5.88	6.18	−1.54
Freq. (log)	7.21	6.99	5.74	5.87	−0.39
N. of Letters	5.75	5.75	7.75	7.67	0.25
N. of Syllables	2.50	2.50	3.28	3.25	0.18
Orth. N	5.25	6.50	3.18	3.43	−0.32
OLD	1.71	1.63	2.35	2.16	1.30

Note. *Freq. (log)* = log lexical frequency; *N. of Letters* = number of letters; *N. of Syllables* = number of syllables; *Orth. N* = number of orthographic neighbors; *OLD* = orthographic Levenshtein distance (Yarkoni et al., 2008). Frequency values (log-transformed) were taken from the SUBTLEX-IT database (Crepaldi et al., 2013). Number of orthographic neighbors and OLD were computed on the PhonItalia database (Goslin et al., 2014) using the vwr package (Keuleers, 2013) in R. Valence, arousal, concreteness, familiarity, and imageability scores were taken from ITA-BOO (Sulpizio et al., 2020) and the Italian adaptation (Montefinese et al., 2014) of ANEW (Bradley & Lang, 1999), for taboo and control words, respectively. The *t*-values result from independent-sample two-tailed *t* tests conducted to compare taboo and control words. Tests were not conducted for color-associated and control stimuli due to the low number of items (4) in each category. ***$p < .001$; **$p < .01$; *$p < .05$

Apparatus and procedure

The experimental procedures for both the semantic and the taboo Stroop experiments were programmed with the Open Sesame software (Version 3.2.8; Mathôt et al., 2012). The experiments were administered online, and online data collection was managed using JATOS (Version 3.5.3; Lange et al., 2015). At the beginning of the experiment, participants were asked to close all the other windows in the browser and all the other applications, as well as to set the browser to full-screen mode. Participants were first presented with an informed consent screen and asked whether they wanted to proceed. After acceptance, participants provided information regarding their age and were then directed to the first experimental procedure.

Each participant performed both the semantic and the taboo Stroop experiment. The order of the two tasks was counterbalanced across participants. The overall structure of the two tasks, as well as the trial events, were the same across both Stroop experiments. Participants were instructed to categorize word stimuli as a function of the color in which they were written by pressing one of four buttons (red: Z; yellow: X; green: N; blue: M), using their right and left index and middle fingers (one finger per response button). In each trial, a fixation cross was presented at the center of the screen for 450, 500, or 550 ms. After a 50-ms blank screen, the target stimulus (colored words) was displayed until response. When participants failed to respond before the allotted time (1,500 ms), a feedback screen ("too slow!") was displayed for 300 ms. The beginning of the next trial occurred after a blank screen lasting 800 ms.

In the semantic Stroop, each word (four color-associated and four control words) was presented in each of the three possible colors 12 times for a total of 288 trials. The whole set of 288 trials were divided in two equal blocks of 144 trials. Each color-word combination occurred equally often across the two blocks. Participants could take a self-terminated break in between the two blocks. Similarly, in the taboo Stroop experiment each participant was administered two blocks of trials (144 trials per block), with a self-terminated break in between. All the taboo (72) and control (72) words were presented once in each block, in a different color across the two blocks. The lists of stimuli were created so that each color appeared equally often across taboo and control stimuli within participants, whereas each word was displayed equally often in each possible color across participants.

Before beginning the first experimental session, participants performed two practice sessions. The first consisted of a response mapping training (following Kinoshita et al., 2018), in which the stimuli consisted of colored strings of six hashmarks (######). Participants were instructed to respond on the basis of their color. Each response color was presented six times, for a total of 24 trials. In the second practice session, four words were presented instead of the stings of hash marks.

Words were different from those selected for the experiments. Participants were asked to respond on the basis of the color in which words were written. Each word appeared three times in three of the four colors, for a total of 36 trials. In both practice sessions, the trial events were the same as in the experimental tasks, except for the fact that a feedback screen (300 ms) was delivered not just when participants failed to respond within the allotted time, but also in case of incorrect responses ("ERROR"). To facilitate the color–response association, during both practice sessions, four small colored squares were constantly displayed in the lower part of the screen, in correspondence to the associated response button (on the left side, red = Z, yellow = X; on the right side, green = N, blue = M).

Statistical analyses

Reaction times (RTs) were analyzed via linear mixed-effects models and response accuracy via generalized linear mixed effects models, using the lme4 library (Version, 4_1.1-21; Bates et al., 2015) in R (R Core Team, 2021). Models included random intercepts for participants, response colors, and words. Fixed effects were assessed using likelihood ratio tests to compare models in which the fixed effect under examination was present versus a version of the model in which it was absent. Fixed terms were retained only if their inclusion determined a significant increase in goodness of fit. In case any interaction resulted significant, all the involved lower-order terms were retained.

For RTs, analyses were conducted only on correct responses. Additionally, responses faster than 200 ms were considered as anticipations and were not included in the analyses. We first analyzed the overall effects of the two experimental manipulations (semantic association and taboo connotation), separately within the semantic and taboo Stroop tasks. We then assessed the correlations between the overall semantic and taboo Stroop effects detected across the two experiments, using Spearman's rho.

In a second step, we considered variations of both Stoop effects as a function of response speed, focusing on a delta-plot analysis. Specifically, within each participant and within each condition, RTs were partitioned into five quantiles. The first quantile thus included the fastest 20% of responses, the second quantile the next fastest 20%, and so on, until the fifth quantile, which included the slowest 20% of the responses. To assess changes in the Stroop effects as a function of response speed, the variable quantile was considered as a fixed effect within subsequent statistical models. Particularly, we assessed potential interaction between the variable quantile and the effect of the experimental manipulations as markers of changes in the unfolding of the Stroop effects as a function of response speed. Nonlinear relationships between quantiles and Stroop effects were assessed using orthogonal quadratic polynomials when fitting the quantile variable. Polynomial terms were retained only if they improved goodness of fit, as assessed by likelihood ratio tests, compared with a model including only linear relationships.

Accuracy analyses are instead reported only for the sake of completeness. Response accuracy was modeled as a binomial variable within generalized linear mixed-effects models. Correlations across experimental effects were not tested, as accuracy was very high, and possibly at ceiling, in both experiments (semantic Stroop: $M = .94$, $SD = .05$, range: .68–1; taboo Stroop: $M = .94$, $SD = .05$, range:.76–.99).

Results

Overall effects

Semantic Stroop

The semantic Stroop effect was significant, $\chi^2(1) = 6.28$, $p = .012$, with slower RTs for color-associated compared with neutral words, $b = 7.27$, $SE = 2.43$, $t = 2.99$ (Table 2).

In terms of response accuracy, there was no difference between color-associated and control words, $\chi^2(1) = 0.04$, $p = .84$ (see Table 2).

Taboo Stroop

There was a significant effect of taboo interference, $\chi^2(1) = 14.89$, $p < .001$. Responses were slower in the taboo condition, compared with the neutral ones, $b = 9.98$, $SE = 2.53$, $t = 3.95$ (Table 2).

There was no significant interference on response accuracy, $\chi^2(1) = 0.01$, $p = .94$ (Table 2).

Correlations

With respect to the RTs, there was no significant correlation between the semantic and the taboo Stroop effect, $r_s = -0.05$, $p = .64$ (Fig. 1).

Delta plots

Semantic Stroop

For models presented in this section, the random intercept for words had a variance close to zero, and was thus dropped to aid models' convergence. There was a significant interaction between quantile and experimental condition (control vs. color associated), $\chi^2(1) = 23.35$, $p < .001$. Fitting the quantile variable using a quadratic orthogonal polynomial increased goodness of fit, $\chi^2(2) = 3613$, $p < .001$. Parameters of the final model are listed in Table 3.

Table 2 Mean response latencies (RTs in ms) and proportion of accurate responses in the two Stroop tasks

Condition	RTs		Accuracy	
	M	*SE*	*M*	*SE*
Semantic Stroop				
Semantic	644	7.24	.94	.005
Control	637	7.69	.94	.006
Difference	7	2.52	0	.002
Taboo Stroop				
Taboo	654	8.30	.94	.005
Control	644	8.35	.94	.005
Difference	10	2.51	0	.003

Note. *M* mean; *SE* standard error of the mean

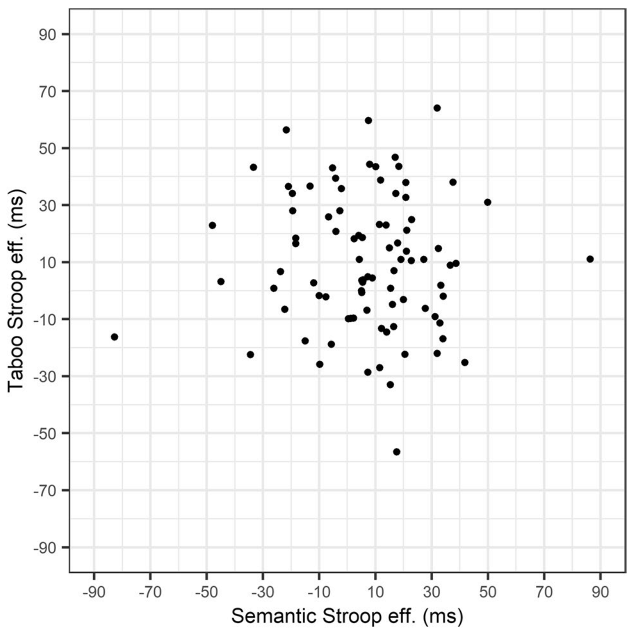

Fig. 1 Scatterplot of the semantic Stroop effect and the taboo Stroop effect. Each point represents an individual participant. There was no correlation between the two effects

As can be seen in Fig. 2a, the semantic Stroop interference is absent in the fastest RTs. It begins to be detected on modal RTs and appears to be strongly enhanced in the slowest RTs.

Taboo Stroop

The random intercept for words had 0 variance and was thus removed from all analyses. The quantile by condition (control vs. taboo) interaction was significant, $\chi^2(1) = 60.49$, $p < .001$. Fitting the quantile variable using a quadratic orthogonal polynomial increased goodness of fit, $\chi^2(2) = 3629$, $p < .001$. Parameters of the model are reported in Table 4.

Table 3 Parameters of the model for the quantile analysis in the semantic Stroop task

Random effects	Variance	*SD*	
Participant	4,721.92	68.72	
Color	3.95	1.99	
Residual	7785	88.23	
Fixed effects	*b*	*SE*	*t*
Intercept	637.16	7.44	85.68
Condition (color-associated)	7.19	1.15	6.28
Quantile (linear)	23,621.92	125.69	187.94
Quantile (quadratic)	5,450.63	124.91	43.63
Condition × Quantile (linear)	909.50	176.48	5.15
Condition × Quantile (quadratic)	168.91	176.47	0.96

Note. *SD* standard deviation; *SE* standard error

As Fig. 2b shows, the effect is absent in the first quantiles. It begins to appear only in modal quantiles and is sharply enhanced in the slowest RTs.

Correlations

For both the semantic and the taboo interference effect, for each participant, we computed the difference between the effect detected in the fifth versus the fourth quantile, thus capturing the slope of the last segment of the delta plot (e.g., van den Wildenberg et al., 2010). The results highlighted a positive correlation with respect to this index across the two tasks, $r_s = .26$, $p = .02$ (Fig. 2c).[1]

General discussion

The experiments highlighted reliable semantic and taboo Stroop effects,[2] with no sizable correlation amongst the two, at least at the level of mean differences between conditions. This may suggest that the two phenomena, despite being both elicited by semantic information, tap into different mechanisms. Consistently, the literature suggests that whereas the semantic Stroop effect stems from conflicting

[1] All the correlation analyses were replicated after transforming RTs into within-participants *z*-scores (Faust et al., 1999), in order to control for overall differences across participants in terms of response speed.

[2] Both the overall semantic and the taboo Stroop effects were rather small. For the semantic Stroop effect, we note that the size of the effects is similar to the one reported in other experiments using a manual variant of the Stroop paradigm (e.g., Kinoshita et al., 2018), including our previous one, featuring partially different control words (Sulpizio et al., 2021). Other researchers have highlighted that the semantic Stroop effect is usually small (e.g., Levin & Tzelgov, 2016; see also Parris et al., 2021), but our data (including Sulpizio et al., 2021) suggest it appears as a rather reliable phenomenon, even in a transparent language such as Italian.

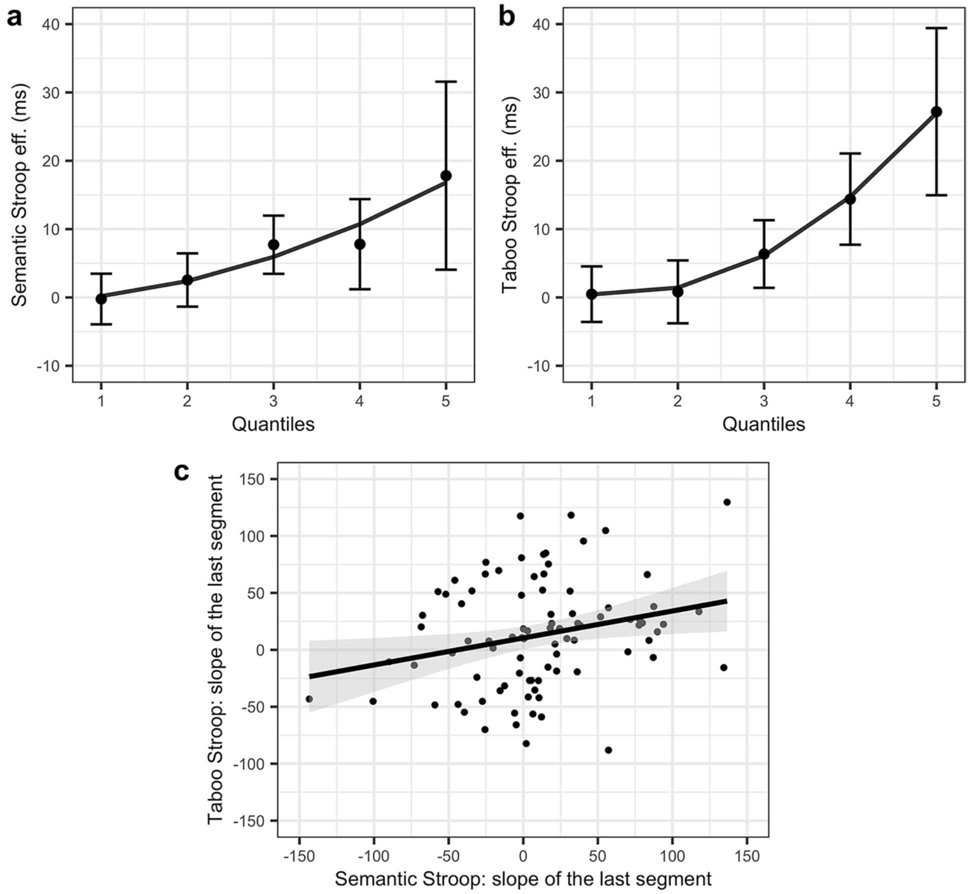

Fig. 2 Results of the delta-plot analyses. **a** Mean semantic Stroop effect (color-associated – control; *y*-axis) as a function of trial quantile (*x*-axis). Points represent empirical means, and error bars reflect corresponding 95% confidence intervals. The black line represents means predicted by the statistical model. **b** As in Panel **a**, for the taboo Stroop task. **c** Scatterplot of the slope of the last segment of the delta plot for the semantic Stroop effect (*x*-axis) versus the taboo Stroop effect (*y*-axis). Each point represents an individual participant

semantic codes activated by the color-associated word and the ink color (e.g., Augustinova et al., 2018; Seymour, 1977), the taboo Stroop effect is related to the attentional capture exerted by the inappropriate connotation of the carrier words, which divert resources away from the task-set (e.g., Williams et al., 1996). The lack of correlation between the two effects clearly warrants caution against any strong interpretation, particularly when considering limitations in the reliability of Stroop effects (and conflict effects more in general) when used to measure individual differences in correlational research (Hedge et al., 2018). The delta-plot analyses, however, revealed a different and more interesting pattern.

Table 4 Parameters of the model for the quantile analysis in the taboo Stroop task

Random effects	Variance	*SD*	
Participant	5,901.62	76.82	
Color	5.28	2.30	
Residual	7,363.73	85.81	
Fixed effects	***b***	***SE***	***t***
Intercept	643.75	8.31	77.49
Condition (color-associated)	9.94	1.11	8.92
Quantile (linear)	24,307.98	122.12	199.05
Quantile (quadratic)	5,145.03	121.49	42.35
Condition × Quantile (linear)	1449.63	171.64	8.45
Condition × Quantile (quadratic)	486.14	171.63	2.83

Note. *SD* standard deviation; *SE* standard error

The semantic Stroop interference increased as a function of response speed, being absent in fastest responses and displaying a steep enhancement in the slowest ones. The overall pattern replicates our previous observation with partially different stimuli (Sulpizio et al., 2021), thus solidifying the finding. As in Sulpizio et al. (2021), we interpret this pattern as an index of fluctuating efficiency in attentional control (e.g., De Jong et al., 1999), yielding lapses in the maintenance of task goals and schema. Detrimental fluctuation in control would become particularly evident in the slowest responses, whereas fastest ones would capture instances in which attentional control is tightly focused on task requests, thus minimizing distractors interference.

Delta-plot analyses of the taboo Stroop effect revealed a very similar pattern, with no interference in fastest responses, followed by an enhancement beginning within modal quantiles and becoming stronger in the slowest responses. Admittedly, this pattern contradicted our prediction concerning the possible engagement, for this particular type of Stroop interference, of a selective suppression mechanism that may intervene to dampen the detrimental influence of task-irrelevant taboo connotation, akin to what has been observed in lexical decision tasks (Scaltritti et al., 2021; Sulpizio et al., 2021). The rationale for our prediction was that the semantic feature of taboo connotation is completely unrelated with respect to the task-relevant semantic dimension pertaining to color-related information, in contrast to what occurs in the semantic Stroop effect, where the interfering semantic information carried by color-associated words is, by definition, related to the task-relevant color dimension. Whereas in the latter case selective suppression might not be an optimal way to control for interference, as it would also act on task-relevant information, for taboo stimuli, the interfering semantic connotation is different form the task-relevant one. We thus hypothesized that, in this second scenario, semantic control could be actively engaged to suppress the task-irrelevant taboo connotation of the carrier word stimuli. However, we did not find any trace of selective suppression. Like the semantic Stroop effect, the taboo interference revealed an enhancement in the slowest latencies. Additionally, the results revealed a significant correlation between the positive slopes of the last segment of the two effects. This potentially points towards a shared underlying phenomenon, that we identified with a fluctuating efficiency of attentional control in maintaining the task goal (i.e., categorize the color) thus limiting any interference from the irrelevant stimuli (i.e., words). Specifically, the correlation may reflect a general sensitivity towards different forms of (semantic) interference in the context of a complex task requiring the engagement of general control resources to maintain the task goal and the task-relevant schema. By requiring four different manual responses in the context of a purely arbitrary mapping, the manual version of the Stroop task may be particularly taxing in terms of task-set maintenance. Under these circumstances, participants might be reliably prone to suffer from different forms of interference, at least when the efficiency of attentional control decreases.

This interpretation finds further support in the pattern observed for the fastest responses. Notably, both the semantic and the taboo Stroop effect are absent in the first two quantiles (see Fig. 2a–b), which capture a relevant portion (40%) of participants responses. In other words, across both tasks, participants fully experienced interference only in a subset of trials (i.e., the slowest ones). When their attention was fully oriented on the task goal (i.e., identify the color), as in faster trials, it is likely that participants could filter out the carrier-word without fully processing it (for the nonautomaticity of word meaning in the Stroop task, see Kinoshita et al., 2018). In contrast, when participants incurred into lapses of attention, the likelihood to process the irrelevant carrier-stimulus was higher, and semantic information had more chance to interfere with the task.

The absence of interference in the fastest responses further differentiates the present results from those we reported in our previous lexical decision studies with taboo words (Scaltritti et al., 2021; Sulpizio et al., 2021), where the taboo interference was already visible in the first quantile and turned into a facilitation within slower responses. Our tentative explanation for the discrepancies in the distributional features of the taboo interference between lexical decision and (manual) Stroop task relies on the differences between the two experimental paradigms in terms of task goals. In lexical decision, the task goal requires to decide whether a letter string is a word or not, a task set that encourages lexical-semantic processing (e.g., Balota et al., 1991; Pexman, 2012). This should limit any chance to filter out the distracting taboo connotation, which would need an "active" inhibition via selective suppression mechanisms requiring time to fully accrue. As a consequence, taboo interference in lexical decision is fully displayed even in the fastest responses and turns into a facilitation (or a null effect) only in slowest ones (Scaltritti et al., 2021; Sulpizio et al., 2021). In the case of Stroop tasks instead, the lexical-semantic information from the distractor words is, in principle, useless to accomplish the task goals. Here, attentional control may effectively filter out the taboo connotation, at least when attention is tightly focused on task goals and schema (fastest trials). In contrast, lapses of attentional control (slowest responses), would let the distracting taboo information hinder the performance. As a result, the taboo interference is absent in fastest responses, and greatly enhanced in the slowest ones.

Further, in the present experiments, the absence of a systematic interference may discourage the system to use a more specific control mechanism, which would be useless on almost half of the trials. This reasoning resonates with recent neurocognitive models of conceptual processing assuming partially distinct networks subserving semantic and domain general control (e.g., Hoffman et al., 2018; Lambon Ralph et al., 2017). According to these models, specific mechanisms for semantic control are functionally dissociated from more general control mechanisms and "are only recruited when conceptual information itself must be controlled, and not whenever semantic tasks become hard" (Gao et al., 2021, p. 2). Therefore, the sporadic interferences we reported here, surfacing when attentional control operates less efficiently, may not prompt the recruitment of any form of semantic control, and may simply reflect fluctuations in a general control mechanism working to maintain task goals and task schema.

We should acknowledge that the literature on Stroop effects also offers a different interpretation with respect to Stroop interference showing an enhancement as a function of response speed that is a positive slope in the delta plots. This perspective relies on the framework of drift diffusion models (e.g., Ratcliff et al., 2016), in which decision is viewed as a process of (noisy) evidence accumulation, continuously unfolding over time until reaching a response-triggering boundary. Here, the positive slope in the delta plot of an effect is usually related to differences in the rate of evidence accumulation as a function of experimental conditions. As in the Stroop task the distracting information (word) and the task-relevant one (color) are merged within one stimulus, competing evidence coming from both sources would be sampled in parallel during the decisional process, thus reducing the rate at which evidence pertaining to the target dimension is collected. The rather ubiquitous positive delta slope found across variants of Stroop interference may thus be mapped onto a change in the rate of evidence accumulation (for more discussion and evidence, see Kinoshita, De Wit, Aji, & Norris, 2017; Kinoshita, De Wit, & Norris, 2017; Pratte et al., 2010). In this context, our findings suggest that the taboo connotation of carrier words should be included in the types of information from the distractor that hinder evidence accumulation from the target, despite the fact that the taboo information per se does not seem to involve evidence incongruent with the target (i.e., color).

Another important conundrum that we should consider in the context of the present results is related to the interpretation of the semantic Stroop effect. In line with some of the extant literature, we have endorsed the notion that the difference between trials with color-associated words and trials with color-unrelated words represents a measure of semantic conflict. Briefly, our choice relied on evidence of dissociations between effects of semantic and response conflict. For example, whereas semantic conflict (as indexed by the comparison between, e.g., *sky* in green vs. *dog* in green) remains relatively constant across manual and verbal variants of the semantic Stroop paradigm, response conflict (as indexed by the comparison between, e.g., *blue* in green vs. *sky* in green) is reduced in the manual version of the task (Augustinova et al., 2018; see also Augustinova et al., 2019; Brown & Besner, 2001). The fact that semantic conflict remains comparable when response conflict is reduced or eliminated seems to support the notion that the two types of conflict can be differentiated, and that semantic conflict cannot be reduced to a form of response conflict.

In a different perspective, color-associated words would still trigger response-conflict, due to their ability to activate the set of response colors (e.g., Cohen et al., 1990; Roelofs, 2003), and dissociations would merely reflect quantitative differences in response competition (for further discussion, see Parris et al., 2021). Importantly, evidence suggests that response conflict in manual Stroop tasks may be displayed in slower RTs (Hasshim et al., 2019), a distributional pattern that resembles the one we reported for the semantic Stroop effect. Without a measure of response conflict, our experiment is unable to tease apart the contribution of the two forms of conflict to the semantic Stroop effect. It is however interesting to note that, in our experiment, the taboo Stroop effect displayed the very same distributional profile and the enhancement of the two effects in the slowest RTs was significantly correlated. As the taboo interference effect reported here seems hard to reconcile with any form of response conflict, the correlation would seem to capture, at least, a partial overlap of the two phenomena, that we ascribe to semantically driven interference effects resulting from attentional lapses in maintain task goals and schema.

In conclusion, the semantic and the taboo Stroop effects both seem to stem from similar processing dynamics, in particular from fluctuations in attentional control of task goals and schema. When attention is deployed less efficiently (i.e., in the slowest trials), (semantic) interference phenomena are enhanced. Moreover, we did not find any trace of a selective semantic suppression of the taboo interference (i.e., the semantic connotation not overlapping with the task-relevant dimension of colors). We speculate that this sort of control mechanism, which has been reported in lexical decision, was not implemented in the current task due to the inconsistent effect of taboo interference in the Stroop paradigm. In fact, both the taboo interference and the semantic Stroop effect, were virtually absent in almost half of the responses. This suggests that rather than unavoidable consequences of the automaticity of lexical-semantic access, these phenomena may be linked with lapses in attentional control, which make the performance more prone to general interference effects from distractors (i.e., words). In turn, the ability of the system to efficiently filter out hindering semantic information, at least in a substantial proportion of trials, further challenges the notion of full automaticity in visual word recognition.

Appendix 1

Analyses of the semantic and taboo Stroop effects as a function of trial progression

In this analysis, we assessed potential changes in the semantic and taboo Stroop effects as a function of the trial progression. Taboo interference effects are often subject to habituation phenomena (e.g., Bertels & Kokinsky, 2016; MacKay et al., 2004; 2015; Sulpizio et al., 2021).

For the semantic Stroop experiment, there was no evidence for a modulation of the effect during the course of the experiment, as the interaction between the type of carrier word (color-associated vs control) and trial number failed to reach significance, χ^2 (1) = 2.27, p = .13. The main effect of trial number, however, was significant, χ^2 (1) = 16.59, p < .001, indicating that RTs were getting progressively faster over the course of the experiment, b = -0.06, SE = 0.01, t = -4.07. Using orthogonal quadratic polynomial in fitting the variable of trial number significantly increased goodness of fit, χ^2 (1) = 4.09, p = .04. The unfolding of RTs across trials for the two conditions is represented in Appendix Fig. 3, panel a.

For the taboo Stroop experiment, instead, there was a significant interaction between the experimental condition (taboo vs control) and the trial number, χ^2 (1) = 5.26, p = .02, as the taboo interference declined over the course of the experiment, b = -0.07, SE = 0.03, t = -2.29. The use of a quadratic orthogonal polynomial in fitting the trial number variable significantly increased goodness of fit, χ^2 (2) = 25.06, p < .001. In fitting the models for the taboo Stroop experiment, we dropped the random intercept for words (i.e., the one associated with the smallest amount of variance) in order to aid model convergence and keep the random effects structure comparable across all models

The unfolding of RTs across trials for the two conditions is represented in Appendix Fig. 3, panel b.

Note that the correlation between the last segments of the delta-plots from the semantic and the taboo Stroop tasks remained significant (r_s = .26, p = .02) even when, for the taboo Stroop, we included in this analysis only the trials from the first half of the experiment (i.e., from trial 1 to trial 144), where the taboo interference was actually displayed.

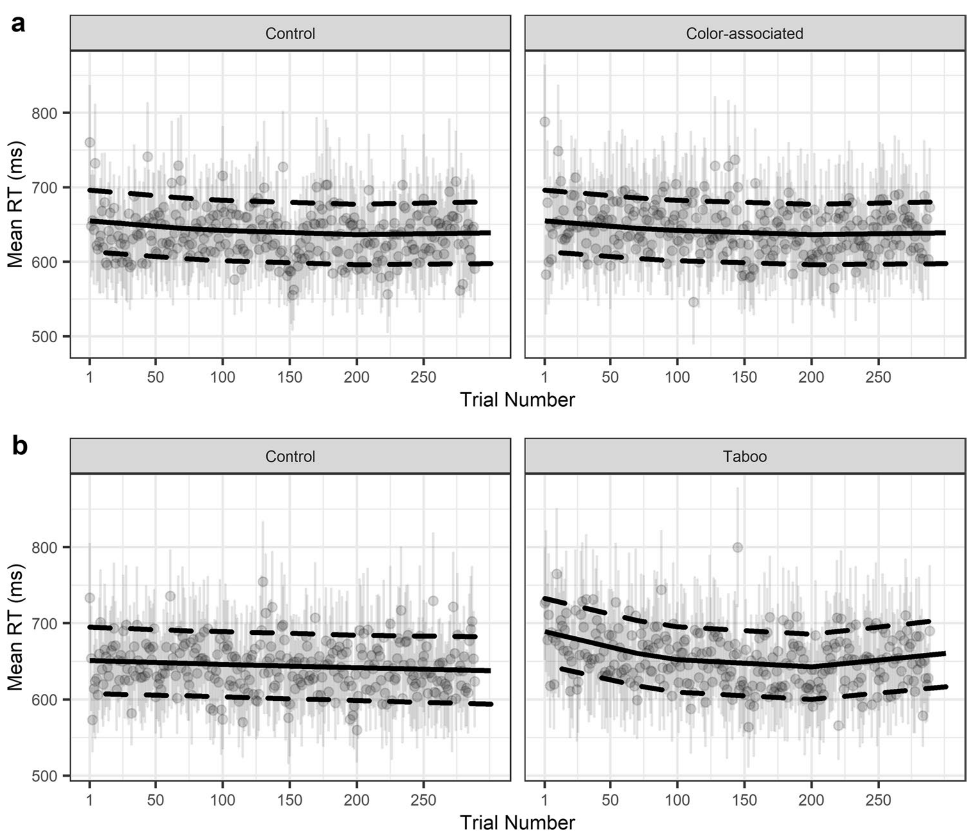

Fig. 3 Modulation of the RTs across different conditions as a function of trial progression. Panel a: mean RTs (y axis) as a function of trial number (x axis) and experimental condition (labels on top) for the semantic Stroop task. Points represent mean RTs, and error bars reflect corresponding 95% confidence intervals. Solid black line represents the effect of trial number predicted by the statistical model, whereas dashed lines represent the corresponding 95% confidence intervals. Panel b: as in panel a, for the taboo Stroop task

Appendix 2

Exploratory analyses of the frequency of occurrence of slow response across the trial sequence

We considered slowest RTs as indicative of trials in which attentional control was operating less efficiently (e.g., De Jong et al., 1999). In this exploratory analysis, we tried to assess how slow responses are distributed across the experiment. Separately for the semantic and the taboo Stroop tasks, we considered the frequency of occurrence of the slowest responses (i.e., those responses falling within the fifth quantile) as a function of trial progression (i.e., from trial 1 to trial 288). Specifically, we fitted linear regression models using trial serial number as the predictor variable, and the frequency of slow responses falling in the fifth quantile as the dependent variable. Both in the semantic (b = - 0.013, SE = 0.003, t = -4.21, p < .001) and in the taboo Stroop task (b = - 0.017, SE = 0.003, t = -5.97, p < .001), the frequency of very slow responses was reduced over the course of the experiment (Appendix Fig. 4). This pattern may be tentatively linked with practice effects, possibly reinforcing the (arbitrary) response mapping and thus attenuating the impact of attentional lapses.

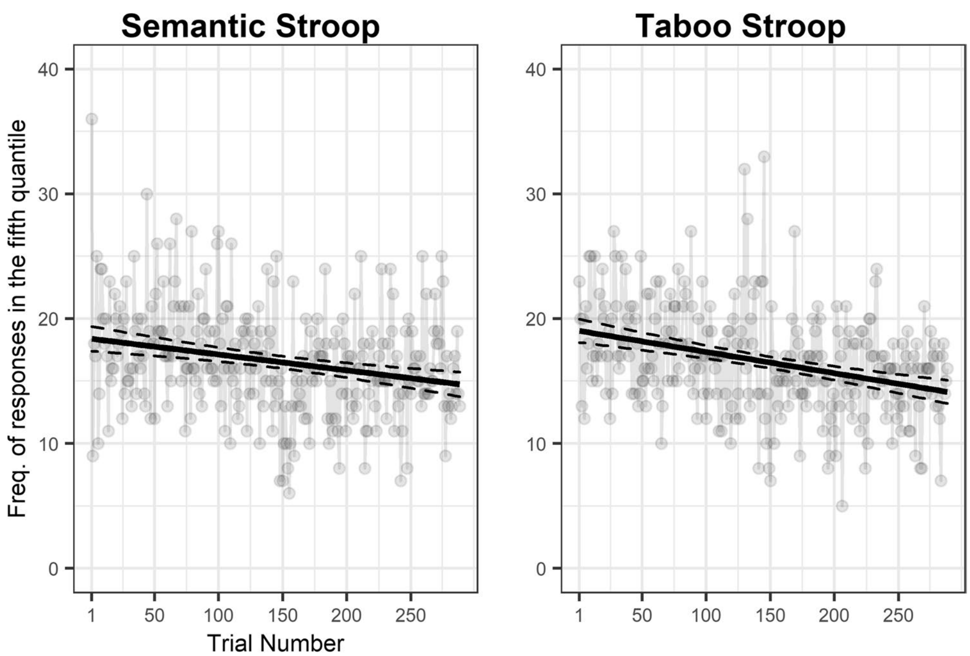

Fig. 4 Exploratory analysis of the distribution of slowest responses over the course of the experiments. Frequency of occurrence of the slow responses falling in the fifth quantile (y-axis) plotted as a function of trial progression (x-axis) in the semantic and taboo Stroop tasks. Grey points and lines represent the number of responses at each trial serial position (from 1 to 288). Solid black line represents the effect of trial number (i.e., trial ordinal position within the sequence) predicted by the statistical model, whereas dashed lines represent the corresponding 95% confidence intervals

Author note The authors are grateful to Paolo Leoni for useful suggestions with respect to the implementation of the online experiment.

Declarations

Conflicts of interest None of the authors has any conflict of interest to disclose.

References

Algom, D., & Chajut, E. (2019). Reclaiming the Stroop effect back from control to input-driven attention and perception. *Frontiers in Psychology, 10*, 1683.

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.

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.

Balota, D. A., Ferraro, F. R., & Connor, L. T. (1991). On the early influence of meaning in word recognition: A review of the literature. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 187–222). Erlbaum.

Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting Linear mixed-effects models using lme4. *Journal of Statistical Software,* 67(1)1, 1–48.

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings* [Technical Report C-1] (Vol. 30, No. 1, pp. 25–36). 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.

Bugg, J. M. (2014). Conflict-triggered top-down control: Default mode, last resort, or no such thing? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 567–587.

Bugg, J. M., & Hutchison, K. A. (2013). Converging evidence for control of color-word Stroop interference at the item level. *Journal of Experimental Psychology: Human Perception and Performance, 39*, 433–449.

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.

Crepaldi, D., Keuleers, E., Mandera, P., & Brysbaert, M. (2013). SUBTLEX-IT. http://crr.ugent.be/subtlex-it/

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.

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.

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), 777–799.

Gao, Z., Zheng, L., Chiou, R., Gouws, A., Krieger-Redwood, K., Wang, X., Varga, D., Lambon Ralph, M. A., Smallwood, J., & Jefferies, E. (2021). Distinct and common neural coding of semantic and non-semantic control demands. *NeuroImage, 236*, 118230

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.

Goslin, J., Galluzzi, C., & Romani, C. (2014). PhonItalia: A phonological lexicon for Italian. *Behavior Research Methods*, *46*, 872–886.

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.

Hedge, C., Powell, G., & Summer, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods, 50*(3), 1166–1186.

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.

Hutchison, K. A. (2011). The interactive effects of listwide control, item-based control, and working memory capacity on Stroop performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 851–860.

Keuleers, E. (2013). vwr: Useful functions for visual word recognition research (R Package Version 0.3.0). https://CRAN.R-project.org/package=vwr

Kinoshita, S., & Mills, L. (2020). Phonological encoding in the oral but not manual Stroop task: Evidence for the role of a speech production process. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*, 1494–1504.

Kinoshita, S., De Wit, B., Aji, M., & Norris, D. (2017). Evidence accumulation in the integrated and primed Stroop tasks. *Memory & Cognition*, *45*, 824–836.

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*, 369–384.

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.

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*, Article e0130834.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*, 314–324.

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.

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

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance, 17*, 22–27.

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.* Advance online publication https://doi.org/10.1007/s00426-021-01554-x

Pexman, P. M. (2012). Meaning-based influences on visual word recognition. In J. S. Adelman (Ed.), *Visual word recognition* (Vol. 2, pp. 24–43). Psychology Press.

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.

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.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences, 20*, 260–281.

Reynolds, M. G., & Langerak, R. M. (2015). Emotional Stroop dilution: The boundary conditions of attentional capture by threat words. *Acta Psychologica, 159*, 108–115.

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

San José, A., Roelofs, A., & Meyer, A. S. (2021). Modeling the distributional dynamics of attention and semantic interference in word production. *Cognition, 211*, 104636.

Scaltritti, M., Navarrete, E., & Peressotti, F. (2015). Distributional analyses in the picture–word interference paradigm: Exploring the semantic interference and the distractor frequency effects. *Quarterly Journal of Experimental Psychology, 68*, 1348–1369.

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, 47*, 934–945.

Schmidt, J. R. (2019). Evidence against conflict monitoring and adaptation: An updated review. *Psychonomic Bulletin & Review, 26*, 753–771.

Seymour, P. H. (1977). Conceptual encoding and locus of the Stroop effect. *Quarterly Journal of Experimental Psychology, 29*, 245–265.

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

Siegrist, M. (1995). Effects of taboo words on color-naming performance on a Stroop test. *Perceptual and Motor Skills*, *81*, 1119–1122.

Spinelli, G., & Lupker, S. J. (2020a). Proactive control in the Stroop task: A conflict-frequency manipulation free of item-specific, contingency-learning, and color-word correlation confounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. https://doi.org/10.1037/xlm0000820

Spinelli, G., & Lupker, S. J. (2020b). Item-specific control of attention in the Stroop task: Contingency learning is not the whole story in the item-specific proportion-congruent effect. *Memory & Cognition*, *48*, 426–435.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643-662.

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.

Sulpizio, S., Job, R., Leoni, P., & Scaltritti, M. (2021). Prepotent task-irrelevant semantic information is dampened by domain-specific control mechanisms during visual word recognition. *Quarterly Journal of Experimental Psychology*. Advance online publication. https://doi.org/10.1177/17470218211030863.

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.

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin, 120*, 3–24.

Open practice statement The data and materials for all experiments are available (https://osf.io/egsbh/).