

Cognitive control mechanisms in language processing: Are there both within- and across-task conflict adaptation effects?

Quarterly Journal of Experimental Psychology
1–23

© Experimental Psychology Society 2022



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/17470218221111789

qjep.sagepub.com



Nicoletta Simi¹ , Ian Grant Mackenzie¹, Hartmut Leuthold¹, Markus Janczyk² and Carolin Dudschig¹

Abstract

Cognitive conflict is regarded as a crucial factor in triggering subsequent adjustments in cognitive control. Recent studies have suggested that the implementation of control following conflict detection might be domain-general in that conflict experienced in the language domain recruits control processes that deal with conflict experienced in non-linguistic domains. During language comprehension, humans often have to recover from conflicting interpretations as quickly and accurately as possible. In this study, we investigate how people adapt to conflict experienced during processing semantically ambiguous sentences. Experiments 1 to 3 investigated whether such semantic conflict produces the congruency sequence effect (CSE) within a subsequent manual Stroop task and whether Stroop conflict leads to adjustments in semantic processing. Experiments 4 to 6 investigated whether semantic conflict results in conflict adaptation in subsequent sentence processing. Although processing conflict was consistently experienced during sentence reading and in the Stroop task, we did not observe any within-task or cross-task adaptation effects. Specifically, there were no cross-task CSEs from the linguistic task to the Stroop task and vice versa (Experiments 1–3)—speaking against the assumption of domain-general control mechanisms. Moreover, experiencing conflict within a semantically ambiguous sentence did not ease the processing of a subsequent ambiguous sentence (Experiments 4–6). Implications of these findings will be discussed.

Keywords

Cognitive control; language processing; semantic ambiguity; conflict adaptation

Received: 20 December 2021; revised: 23 March 2022; accepted: 23 March 2022

Cognitive control is a well-studied phenomenon of human cognition. It has been proposed that cognitive control processes allow us to adapt our behaviour in a flexible and goal-oriented manner to changing environmental demands (e.g., Gruber & Goschke, 2004). A prominent theory of cognitive control is the conflict monitoring theory (Botvinick et al., 2001). This theory assumes that the detection of conflict during information processing triggers a subsequent upregulation of cognitive control, for instance, by weakening the impact of task-irrelevant information and strengthening those of task-relevant information from trial to trial. These changes in control become then observable by affecting performance patterns in a trial as a function of whether the preceding trial entailed, for example, a conflict or an incorrect response. Such short-term conflict adaptation effects have been reported across a wide range of paradigms, including the

Stroop task (Forster & Cho, 2014; Funes et al., 2010; Kan et al., 2013), the Flanker task (Akçay & Hazeltine, 2011; Boy et al., 2010; Gratton et al., 1992; Janczyk & Leuthold, 2018) and the Simon task (Janczyk & Leuthold, 2018; Kunde et al., 2012; Kunde & Wühr, 2006; Stürmer et al., 2005). Errors have also been shown to trigger adaptation effects (cf. Dudschig & Jentzsch, 2009; Jentzsch & Dudschig, 2009). Moreover, convergent evidence from behavioural, neuroimaging, and neuropsychological

¹Department of Psychology, University of Tübingen, Tübingen, Germany

²Department of Psychology, University of Bremen, Bremen, Germany

Corresponding author:

Nicoletta Simi, Fachbereich Psychologie, Universität Tübingen, Schleichstr. 4, 72076 Tübingen, Germany.

Email: nicoletta.simi@uni-tuebingen.de

studies supports an interaction between cognitive control and language processing (Hsu & Novick, 2016; January et al., 2009; Thothathiri et al., 2012, 2018; Vuong & Martin, 2011). In fact, overlapping brain regions are shown to be activated by both standard cognitive control tasks (Stroop and Flanker) and language processing tasks where competitive alternatives are present (Hsu et al., 2009; Vuong & Martin, 2011). However, a current debate about conflict adaptation effects concerns the issue of whether conflict detection in one processing domain (e.g., sensorimotor) results in conflict adjustments in a different domain (e.g., linguistic). In this respect, recent studies of Novick and colleagues' research group (Hsu et al., 2017, 2021; Hsu & Novick, 2016; Kan et al., 2013; Novick et al., 2014) have attracted much attention. They showed cross-task adaptation effects from linguistic tasks involving syntactic conflict to both perceptual and sensorimotor tasks and vice versa. It remains an open issue whether (1) cross-task adaptation effects occur for different types of representational linguistic conflict (i.e., semantic conflict) and (2) whether within-task adaptation effects occur within the language domain. These issues are of great theoretical importance for two reasons. First, demonstrating cross-task adaptation effects for semantic instead of syntactic conflict would further strengthen the view of a domain-overarching cognitive control architecture. Second, demonstrating within-task adaptation effects would, at least to our knowledge, for the first time reveal that language comprehension processes at the sentence level are also subject to short-term cognitive control influences following semantic conflict. The goal of this study is to address these issues in a set of behavioural experiments either employing a self-paced reading task alone or in combination with a manual Stroop task.

Within standard tasks used to investigate conflict-triggered control implementations, stimuli typically consist of a relevant and an irrelevant task dimension. For example, in the classical Stroop task (for a review, see MacLeod, 1991), the relevant dimension is the print colour of the word, whereas the irrelevant dimension is the word meaning. In such a task, conflict originates if the word meaning and the print colour carry opposing information (incongruent trial; for example, the word *red* presented in blue colour). Responses are typically faster and less error-prone in congruent (C) than incongruent (I) trials. Another example is the Eriksen flanker task (Eriksen & Eriksen, 1974; Kopp et al., 1994; Verbruggen et al., 2006), where the central stimulus is the task-relevant dimension that is either congruent (<<<<<<) or incongruent (>>>>>>) with the surrounding stimuli. Crucially, short-term conflict adaptation effects—as proposed by conflict monitoring theory (Botvinick et al., 2001)—are typically revealed via congruency sequence analyses in such tasks; specifically, congruency on the current trial N (C vs. I) as a function of congruency on the preceding trial $N-1$ (c vs. i). One of the

first studies reporting such an analysis was Gratton et al. (1992). Using an Eriksen flanker task, the authors found that the congruency effect on trial N was reduced if it was preceded by an incongruent compared with a congruent trial $N-1$ (cI-cC > iI-iC). This phenomenon is referred to as the *congruency sequence effect* (CSE).

The actual mechanisms underlying the CSE are still debated. A full discussion of the proposed mechanisms is beyond the scope of the present paper, but some key discussions will be briefly summarised in the following. One issue regarding the CSE is the extent to which low-level stimulus-response bindings contribute to the effect (e.g., Hommel et al., 2004; Mayr & Awh, 2009; Mayr et al., 2003). Indeed, Mayr et al. (2003) argued that the adaptation effect pattern is specific to stimulus-response repetitions and may be the consequence of associative priming rather than top-down control initiated by conflict detection. Another debate has evolved around the question of whether control adjustments operate in a task-specific (Akçay & Hazeltine, 2011; Egner et al., 2010; Forster & Cho, 2014; Stürmer et al., 2005) versus a more domain-general manner (Freitas & Clark, 2014; Freitas et al., 2007; Kan et al., 2013). On one hand, if conflict adaptation is a manifestation of a domain-general system, then there should be clear evidence of cross-task CSEs. On the other hand, if adaptation reflects domain-specific mechanisms—for example, an incongruent flanker trial specifically resulting in more focusing on centrally presented information as suggested by the Spotlight model (Cohen et al., 1992; Eriksen & Eriksen, 1974; Eriksen & Hoffma, 1972; Miller, 1991; Yantis & Johnston, 1990)—there should be no evidence for such cross-task adjustments (cf. Braem et al., 2014). Whereas a large body of research indicates conflict adaptation effects to be task-specific, the evidence in favour of cross-task CSEs is rather limited (e.g., Freitas et al., 2007; Kleiman et al., 2014; Notebaert & Verguts, 2008; but see, for example, Akçay & Hazeltine, 2011; Funes et al., 2010). Such a result, however, would be of great theoretical interest since it can be taken as evidence against possible alternative explanations of the CSE in terms of low-level stimulus-response bindings (Hommel et al., 2004; Mayr et al., 2003).

A key study that was specifically designed to demonstrate conflict adaptation effects across distinctly different types of tasks, that is, either a language reading task involving syntactic conflict or a perceptual conflict task (Neckar cube) to a manual Stroop task, was conducted by Kan et al. (2013). In their study, cross-task congruency effects in the manual Stroop task were demonstrated when participants read on trial $N-1$ syntactically ambiguous sentences (Experiment 1) and also when a perceptually bistable Neckar cube was presented (Experiments 2 and 3). Most interesting for this study are the cross-task transfer effects between the reading task and the subsequent manual Stroop task. The employed sentences involved high

conflict trials with syntactically ambiguous (incongruent) sentences (e.g., “*The basketball player accepted the contract would have to be negotiated.*”) and low-conflict trials where no ambiguity (congruent) was present (e.g., “*The basketball player accepted that the contract would have to be negotiated.*”). The word-by-word presentation of the sentences using a self-paced reading paradigm was followed by the Stroop task that required manual choice responses as a function of font colour. A smaller Stroop effect in reaction time (RT) and error rate (ER) was found following incongruent as compared with congruent sentences. These findings accord with the idea that the detection of syntactic conflict initiates cognitive control processes that reduce conflict within subsequent Stroop trials, thus supporting the notion of a domain-general cognitive control mechanism (but see Aczel et al., 2021, for an unsuccessful replication attempt). This assumption was further substantiated by similar cross-task adaptation effects if the Stroop trials were preceded by perceptual conflict (Experiments 2 and 3). Whereas the Kan et al. (2013) study is illustrative, the authors argue that it would also be of interest to demonstrate cross-task adaptation effects from the Stroop to the language task, for which their analysis provided no evidence, as well as for interference tasks different from the Stroop task.

In an attempt to address this issue, Hsu and Novick (2016) combined the Stroop task with a language comprehension task in which syntactically ambiguous (e.g., “Put the frog on the napkin onto the box.”) and unambiguous sentences (e.g., “Put the frog that’s on the napkin onto the box.”) were auditorily presented while eye movements to objects in the visual world (e.g., an empty napkin, a frog on a napkin, a box, a horse) were recorded. The participants’ task was to use the mouse to move the target object (e.g., the frog on the napkin) to the correct goal location (e.g., box). Most interestingly, if incongruent compared with congruent Stroop trials preceded ambiguous sentences, eye movements occurred more frequently to the correct than the incorrect (e.g., empty napkin) goal location; that is, a conflict adaptation effect was present in the language task. Moreover, in a most recent study, Hsu et al. (2021) demonstrated that the recruitment of cognitive control via incongruent flanker trials also results in subsequent processing advantages for syntactically ambiguous sentences in the visual world language task, as in Hsu and Novick (2016). Specifically, the authors reported across several experiments that cognitive control recruited in the non-linguistic arrow flanker task speeds up revision processing during the reading of ambiguous sentences, resulting in fewer errors in the subsequent visual world comprehension task. These findings again strongly support the idea that conflict recruits cognitive control in a domain-overarching manner.

The most prominent line of research that focused on conflict adaptation within semantic tasks was carried out

by Nozari and colleagues (Freund et al., 2016; Nozari & Dell, 2011; Nozari & Novick, 2017). In their conflict-based account, the authors see conflict as a domain-general phenomenon as its signal is monitored in both linguistic and non-linguistic systems. However, in Nozari and Dell (2011), the authors provided evidence from computational modelling and individuals with brain damage, showing that the consequences of conflict detection (e.g., detecting errors) are specific to the source of conflict. In fact, the amount of conflict between lexical representations (e.g., cat and dog) only predicted the ability to detect semantic errors, while the amount of conflict between phonological representations (e.g., /k/ and /d/) only predicted the detection of phonological errors. Importantly, increased conflict at the lexical level did not lead to better detection of phonological errors and vice versa. According to the authors, this specificity arises because each layer of the production system generates conflict independently of other layers and presumably of other cognitive systems, and it is the internal dynamics of these conflict generators that determine the strength of the conflict signal. Thus, the model poses a domain-specific component to the monitoring process.

Surprisingly—despite a few studies addressing conflict adaptation between language comprehension and, for example, a Stroop task—it remains an open issue whether within-task CSEs occur in the language domain. Previous language studies have also investigated conflict adaption with the use of short, negated phrases (“not left,” “not right,” “now right,” and “now left”; Dudschig & Kaup, 2018, 2020). However, in these experiments, participants had to perform the according actions (left vs. right keypress according to phrase meaning), and therefore, the extent to which the observed CSE was purely linguistic or sensorimotor in nature remains debatable. In this study, we avoid potential confounds with sensorimotor processing by investigating CSEs effects within a sentence comprehension task proper (Experiment 4–6). This will allow us to gain a better understanding of how far conflict-related control implementations play a role within linguistic processing, and, consequently, whether conflict monitoring acts similarly in distinct cognitive domains (e.g., language domain and the sensorimotor domain). Surprisingly, and to the best of our knowledge, there are to date no studies investigating CSEs within sentence reading tasks (cf. Dudschig, 2022). From the perspective of the conflict monitoring model and the above reported results showing cross-task conflict adaptation between, for example, Stroop and language comprehension tasks, we suggest that also the monitoring of representational conflicts in language tasks should also lead to adjustments in the subsequent processing of linguistic input. Moreover, it is of interest whether the cross-task adaptation effects reported by Kan et al. (2013) and Hsu et al. (2021) are found in a similar manner when using a different type of representational language conflict than the

previously investigated syntactic conflict, that is, semantic conflict.

To address these issues, we conducted six experiments using a self-paced sentence reading task and a manual Stroop task similar to Kan et al. (2013). However, as mentioned above, this study utilised semantic conflict within sentence materials (e.g., Musz & Thompson-Schill, 2016; Novick et al., 2005; Rayner, 1998; Rodd et al., 2010) to investigate the Generalisability of cross-task conflict adaptation between language comprehension and a Stroop task (in contrast to syntactic conflict materials used by the group of Novick (Hsu et al., 2017, 2021; Hsu & Novick, 2016; Kan et al., 2013; Novick et al., 2014). Semantic conflict is an interesting phenomenon to investigate as it is ubiquitous in language comprehension. Also, the retrieval of multiple meanings when reading an ambiguous homonym somewhat mirrors closely the processes involved when processing incongruent Stroop trials. Thus, when encountering a word like ball (which is semantically ambiguous), it is assumed that we immediately recall its dominant meaning (the round toy) and access its subordinate meaning (dance) only in the case where the surrounding context allows us to disambiguate it from its dominant meaning. During this information processing phase, one of the two meanings has to be inhibited in favour of the other (January et al., 2008; Moss et al., 2005; Novick et al., 2005; Thompson-Schill et al., 1997). Thus, we assume that cognitive control influences semantic disambiguation to avoid representational conflict at the semantic level.

In the present paper, in Experiments 1 to 3, the sentence reading task was intermixed with standard Stroop trials—as in the study by Kan et al. (2013)—to investigate cross-task transfer effects of conflict adaptation for a different type of representational, that is, semantic conflict. Then, we investigate the role of conflict monitoring within the language domain in Experiments 4 to 6. Specifically, these experiments investigate whether CSEs can be observed within a sentence reading task. In other words: does reading a semantically ambiguous sentence facilitate the processing of a subsequent semantically ambiguous sentence? We hypothesise that conflict-related effects in reading times should be observed on words following the semantically ambiguous one, that is, in the post-critical region. Longer reading times for incongruent than congruent conditions for this region would indicate the occurrence of a semantic conflict. Likewise, the finding of a CSE in this region would point towards the role of a conflict monitoring system within the language domain. If Stroop trials show adaptation as a function of semantic conflict, this would accord with the assumption of a domain-general process of control adaptation. The absence of cross-task adaptation, however, would rather support some level of domain-specificity of cognitive control. In addition, all experiments included a separate manual Stroop task block to demonstrate that the experimental setup allows measuring a standard CSE within this interference task.

Method

Several procedural details across the six subsequently reported experiments are identical; thus, for brevity, they will be reported here. Importantly, in each experiment, a manual Stroop task was implemented to ensure that standard conflict adaptation effects within this task can be assessed in the current experimental setup. Experiments 1 to 3 were identical, with the only difference being the presentation mode of the self-paced reading task. The same applied for Experiments 4 to 6. Presentation mode was manipulated (see Figure 1) to investigate whether any specific type of presentation mode taps into measuring specific sentence processing mechanisms in self-paced reading paradigms. In Experiments 1 and 4 (cumulative mode), words remained on screen, allowing participants the opportunity to look back to and reread critical regions (e.g., Felser et al., 2003); in Experiments 2 and 5 (non-cumulative mode), words were masked by dashes with the onset of the next word, thus removing the opportunity to reread (e.g., Beck & Weber, 2020; Ferreira & Henderson, 1990; Gibson & Warren, 2004; Koornneef & Van Berkum, 2006; Schneider et al., 2020); in Experiments 3 and 6 (centre non-cumulative mode), words were centrally displayed and replaced each other, thus removing advance information regarding exact sentence length (e.g., Ditman et al., 2007; Payne & Federmeier, 2017).

To the best of our knowledge, no study has directly investigated the role of presentation mode on sentence reading times. We implemented these three widely used self-paced reading setups to investigate whether there are any systematic influences of the paradigm on sentence processing effects and to replicate our findings across several slightly different setups. The cumulative presentation paradigm (Experiments 1 and 4) is more similar to the natural way of reading sentences and allows participants to go back and reread parts of the presented sentences again. The non-cumulative presentation (Experiments 2 and 5) corresponds to the one used by Kan et al. (2013) and gives a more accurate idea of how people process sentences online compared with the cumulative presentation. However, one disadvantage of the non-cumulative paradigm is that participants know how long the sentences are based on the dashes pre-sent on the screen. Knowledge of the length of a sentence and how close a word is to the end of the sentence can cause the development of expectations and predictions about the incoming words. This is impossible in the centre non-cumulative presentation (Experiments 3 and 6) because participants can see only one word/phrase at a time at the centre of the screen, and they do not have any clues about the length of the sentence. This presentation mode is also of interest because it is most widely used in event-related potential (ERP) studies of sentence comprehension (e.g., Kutas & Hillyard, 1980).

Based on the means and standard deviations provided by Kan et al. (2013; Table 1) for the interaction effect

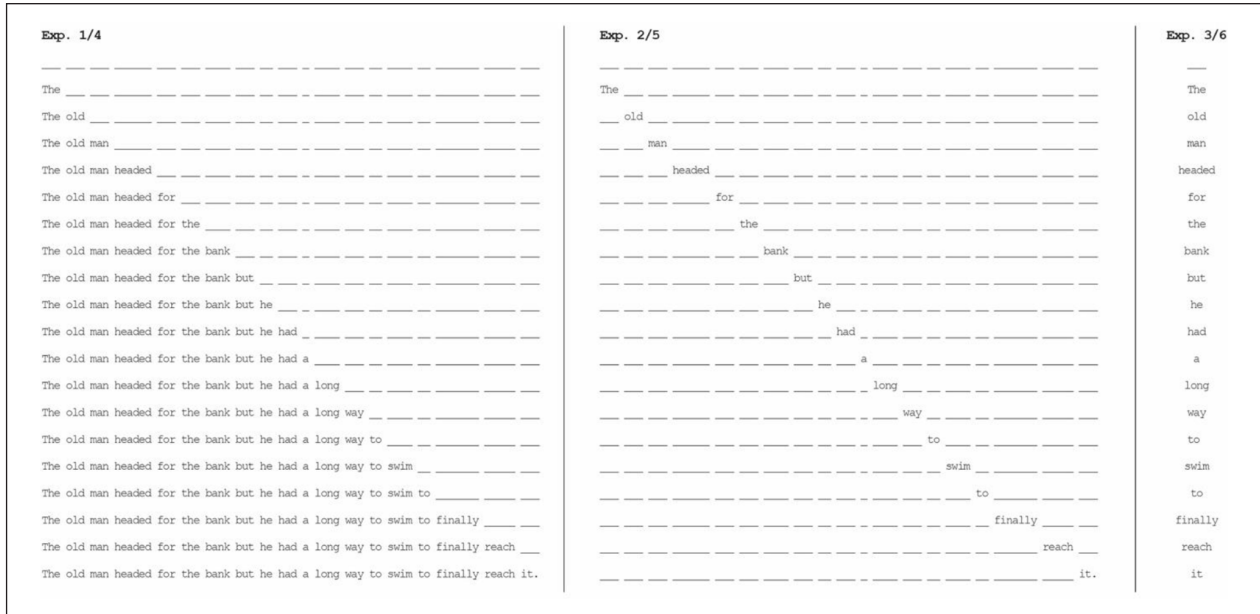


Figure 1. Presentation mode for the self-paced reading task across the six experiments.

Table 1. Example item in both the congruent and incongruent condition.

Condition	Region 1	Region 2		Region 3	
		Main noun		Disambiguating word	
Incongruent	The old man headed for the	bank	but he had a long way to	swim	to finally reach it.
Congruent	The old man headed for the	boat	but he had a long way to	swim	to finally reach it.

A single participant would only see a single item in either the congruent or incongruent condition.

(CSE) in the Stroop task, we determined the number of participants required using the R package powerbydesign (Papenmeier, 2018). This analysis showed for $\alpha=.05$ and between-condition correlations $r=.82$ (estimates were obtained from similar Stroop data from our lab with r s ranging from .82–.96; estimates from a reading time experiment showing r s $>.95$; see also Brysbaert, 2019) that a sample size of 63 participants was required to achieve a power of $\beta=.80$. Therefore, we aimed at including a minimum of 63 complete data sets in each experiment. To reveal potentially smaller experimental effects that are not noticed in the analysis of individual experiments, we also conducted a combined analysis of data sets from Experiments 1 to 3 and Experiments 4 to 6, which only differed with respect to the presentation mode.

Participants

All participants were recruited using the Amazon Turk platform, with the request that only native English speakers participate. All participants indicated informed consent before the experiment began. The experiment took approximately 20 minutes to complete, with participants being

paid \$3.50. This study was approved by the University of Tübingen Faculty of Science ethical committee (Ref.-Nr. 0831_132). Each experiment consisted of two critical parts. A small number of participants did not continue with the second part of the experiment following the completion of Part 1. Only participants who completed both parts 1 and 2 are included in the subsequent analysis. In addition, we used the Stroop accuracy performance in Part 2 to exclude participants who did not demonstrate adequate task performance; specifically, we removed participants with an overall error rate greater than 20% from both the analysis of the word-by-word self-paced reading task and the Stroop task.

Materials

The sentence material was adapted from Blott et al. (2020). The original materials varied in the number of words present within the region following the critical word. Specifically, there were three or four words in the original material. We decided to adapt this in a way so that the region following the disambiguating word would always have four words. The ambiguous sentences, which we will

refer to as incongruent (e.g., “The old man headed for the bank but he had a long way to swim to finally reach it.”), were created to guide the initial interpretation towards the dominant meaning of the homonym (e.g., bank as a financial institution). The homonym was then followed by a disambiguating region which forced the readers to reanalyze the homonym towards its unexpected subordinate meaning (e.g., bank as a side of a river). In incongruent conditions, readers typically slow down after encountering the word “swim,” which conflicts with the initial interpretation and drives the readers to reinterpret the word bank towards its less common subordinate meaning (see Blott et al., 2020). Congruent sentences (e.g., “The old man headed for the boat but he had a long way to swim to finally reach it.”) did not require any reinterpretation to comprehend the sentence correctly since the homonym was replaced with an unambiguous noun which was compatible in meaning with the disambiguating region. The wrap-up region of the sentence was identical in congruent and incongruent conditions (see Table 1 for an example item and the online Supplementary Material A for the complete item list). The items used within the Stroop task consisted of the words “blue” and “green” presented in either green or blue font colour.

Procedure

All experiments were written in JavaScript using the JsPsych library (De Leeuw, 2015) and run online in a standard web browser. Participants were informed that the browser would enter “Full-Screen” mode, and an initial screen-size check procedure would require a minimum screen resolution of $1,280 \times 720$ pixels to proceed. Participants provided age, gender, and handedness information and informed consent before beginning the experiment. The first task involved a calibration routine that required participants to adjust a small rectangular box until it matched the size of a standard bank card. It was not possible to participate using a smartphone or tablet device.

Experiments 1 to 3 consisted of three parts: (1) the word-by-word reading task, (2) the manual Stroop task, and (3) the recall phase.¹ In Part 1, the word-by-word reading task and the Stroop task were combined (see the respective Method sections for further details). Experiments 4 to 6 were identical except that Part 1 consisted of only the word-by-word reading task. Within the word-by-word reading task, the sentence stimuli were counterbalanced across participants so that participants encountered only one of the sentences within each pair (congruent vs. incongruent). In the end, we randomised 48 experimental trials (24 congruent sentences and 24 incongruent ones). Therefore, we had a slightly larger trial number than the original Kan et al. (2013) study (21 congruent and 21 incongruent sentences for each participant). In all six experiments, Part 2 involved a manual two-choice Stroop

task that required responding to the font colour (blue vs. green) of a centrally presented word “blue”/ “green” using the keys Q and P, with the left and right index fingers, respectively. The mapping of response-key to colour was randomly assigned across participants. The 48 Stroop trials were presented in a single block. Each trial started with a 500-ms fixation cross followed by the colour-word stimulus, which remained on the screen until response onset. Then the feedback screen followed, indicating if the response was correct or incorrect. This screen was displayed for 500 ms, followed by a 500-ms blank inter-trial-interval, after which the next trial started. In Experiments 1 to 3, the Stroop task block was preceded by a practice block of 10 trials to familiarise participants with the response-key mappings. A custom self-paced-reading plugin for JsPsych was developed for Part 1, with further details provided in the online Supplementary Material B.

Data analysis

For the analysis of the word-by-word reading data, each sentence was split into three regions (see Table 1): (1) a pre-ambiguity region that started with the first word of the sentence and included all words up to but excluding the first critical word (Region 1), (2) the ambiguous region up to but excluding the disambiguating word (Region 2) and, (3) the disambiguating region until the last word in the sentence (Region 3). The data preparation for the word-by-reading data involved the following steps. First, extreme values at the word level ($>2,000$ ms) were identified. Any sentence with such an extreme value at the word level was excluded from the subsequent analyses. Second, the average reading time per word was calculated for each participant and sentence region, with a low and high cut-off criterion being defined as ± 2.5 SDs of the calculated mean. Individual responses beyond these cut-off values were replaced by the respective cut-off values as in Kan et al. (2013).² Finally, paired *t*-tests were conducted for each of the three regions. For the analysis of Sentence-to-Sentence congruency effects in reading times, repeated-measures ANOVAs with the within-subject factors previous sentence congruency (congruent vs. incongruent) and current sentence congruency (congruent vs. incongruent) were conducted.³ For the analysis of the Stroop data, trials with RTs shorter than 150 ms or longer than 2,000 ms were classified as outliers (too fast and too slow, respectively) and were excluded from the subsequent analysis. In addition, trials with incorrect responses were excluded, and only trials that were preceded by a correct trial were considered in the sequential RT analysis. For RT and ER, repeated-measures ANOVAs with the within-subject factors previous Stroop congruency (congruent vs. incongruent) and current Stroop congruency (congruent vs. incongruent) were conducted. The inferential statistics for the CSE analysis of all 6 experiments are reported in Table 2.

Table 2. Summary of the statistical data for all experiments. The reading task measures refer only to those of Region 3.

Part 1															Part 2														
Reading Task		Sentence-Stroop CSE					Error Rate					Stroop Task CSE					Error Rate												
		RTs					Error Rate					RTs					Error Rate												
N	Diff (ms)	t	p	d	CSE (ms)	F	p	η_p^2	CSE (%)	F	p	η_p^2	CSE (%)	F	p	η_p^2	CSE (%)	F	p	η_p^2									
Experiment 1	13	3.53	**	0.43	-4	0.04	.84	.00	-0.9	0.45	.50	.01	18	4.26	*	.06	1.8	3.59	.06	.05									
Experiment 2	7	3.09	*	0.37	7	0.22	.64	.00	-0.4	0.09	.76	.00	25	3.75	.06	.05	3.1	5.92	*	.08									
Experiment 3	2	1.03	.31	0.12	-14	0.94	.34	.01	2.2	1.92	.17	.03	21	3.04	.09	.05	1.0	1.06	.31	.01									
Part 2																													
Reading Task		Sentence-Sentence CSE					Error Rate					Stroop Task CSE					Error Rate												
		RTs					Error Rate					RTs					Error Rate												
N	Diff (ms)	t	p	d	CSE (ms)	F	p	η_p^2	CSE (%)	F	p	η_p^2	CSE (%)	F	p	η_p^2	CSE (%)	F	p	η_p^2									
Experiment 4	18	3.44	**	0.44	-2	0.08	0.08	.00	.78	.00	.00	.00	17	3.72	.06	.06	2.8	9.18	**	.13									
Experiment 5	6	2.14	*	0.27	-1	0.05	0.05	.00	.83	.00	.00	.00	26	8.13	**	.11	3.2	7.96	**	.11									
Experiment 6	9	3.43	**	0.41	-1	0.06	0.06	.00	.81	.00	.00	.00	21	4.33	*	.06	1.7	2.13	.15	.03									

CSE: congruency sequence effect.
 * $p < .05$, ** $p < .01$, *** $p < .001$.

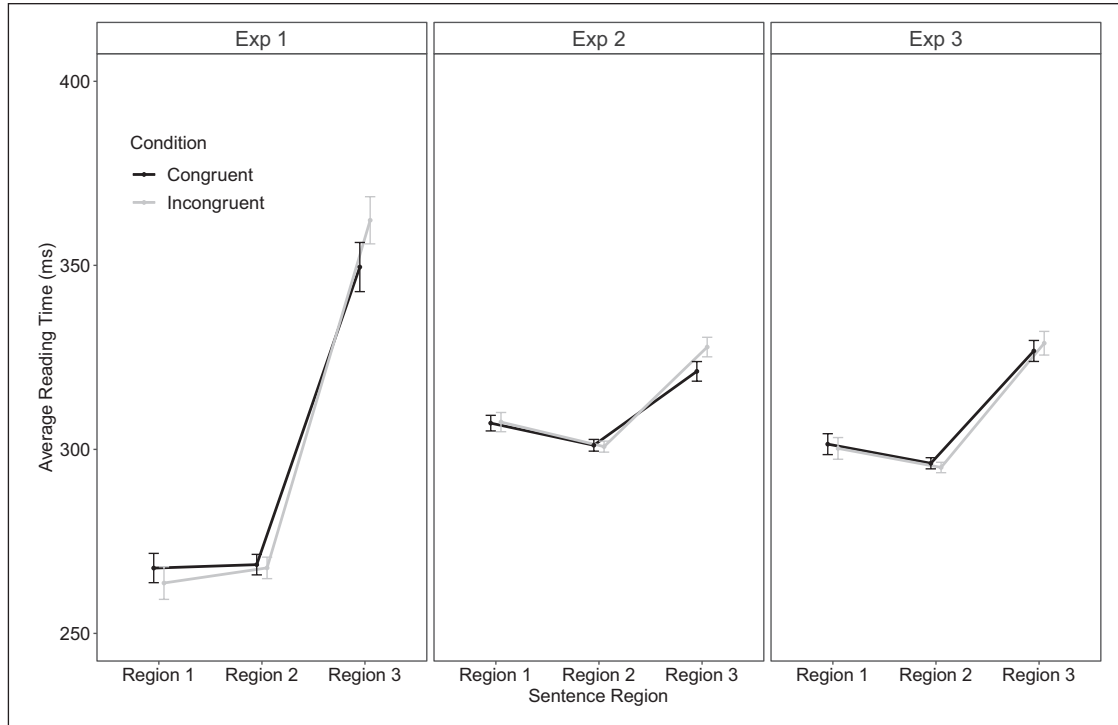


Figure 2. Average reading time (ms) as a function of sentence region and congruency condition within Part I of Experiments 1 (left), 2 (middle) and 3 (right) column.

The error bars represent ± 1 within-subject standard error (Morey, 2008).

Experiment 1

Participants

Seventy-two American English native speakers participated (34 females, $M_{age} = 45.22$, $SD_{age} = 11.36$, 65 right-handed), with 69 full datasets from both parts 1 and 2 remaining. One participant was removed from the subsequent analyses due to poor performance in the Stroop task during Part 1 (ER = ~31%). Thus, 68 participants remained in the subsequent analyses.

Procedure

In Part 1, participants read the sentences in a self-paced fashion by pressing the spacebar to reveal one word at a time. Each trial began with a full mask represented by a string of underscore dashes replacing each word. The sentences were presented in 20px monospaced font and did not require line breaks. The dashes were substituted by the first/next word in the sentence with each subsequent keypress. In this version of the word-by-word reading task, as each word appeared, the previous words remained on the screen until the end of the trial (see Figure 1, left column). Such a presentation mode potentially allows participants to reread the previous words in line with eye-tracking experiments (Rayner, 1998; Witzel et al., 2012), and thus might result in longer reading times on the last word and, as a consequence, in the last region of our analysis. Stroop

and sentence trials were randomly intermixed in the first part of the experiment. Here, we pseudo-randomised 144 experimental trials (48 congruent Stroop, 48 incongruent Stroop, 24 congruent sentences, 24 incongruent sentences) such that between 1 and 4 Stroop trials always separated a sentence trial (i.e., the pseudo-randomisation procedure ensured that two sentence trials were not presented sequentially). In the end, this allowed us to analyse cross-task CSEs in two separate ways. First, we analysed Stroop task performance depending on the congruency of the previous sentence congruency: $\text{Congruent}_{\text{Sentence}} - \text{Incongruent}_{\text{Stroop}}$, $\text{Congruent}_{\text{Sentence}} - \text{Congruent}_{\text{Stroop}}$, $\text{Incongruent}_{\text{Sentence}} - \text{Incongruent}_{\text{Stroop}}$, $\text{Incongruent}_{\text{Sentence}} - \text{Congruent}_{\text{Stroop}}$. Second, reading times were analysed depending on the congruency of the preceding Stroop trial: $\text{Congruent}_{\text{Stroop}} - \text{Incongruent}_{\text{Sentence}}$, $\text{Congruent}_{\text{Stroop}} - \text{Congruent}_{\text{Sentence}}$, $\text{Incongruent}_{\text{Stroop}} - \text{Incongruent}_{\text{Sentence}}$, $\text{Incongruent}_{\text{Stroop}} - \text{Congruent}_{\text{Sentence}}$. Before starting the experiment, participants could familiarise themselves with the task, with one sample sentence demonstrating the self-paced reading procedure.

Results

For the intermixed tasks, we first analysed the reading time and Stroop data (RT, ER) separately to allow a comparison with the results of Experiments 4 to 6, and then conducted the CSE analysis of the Stroop effect according

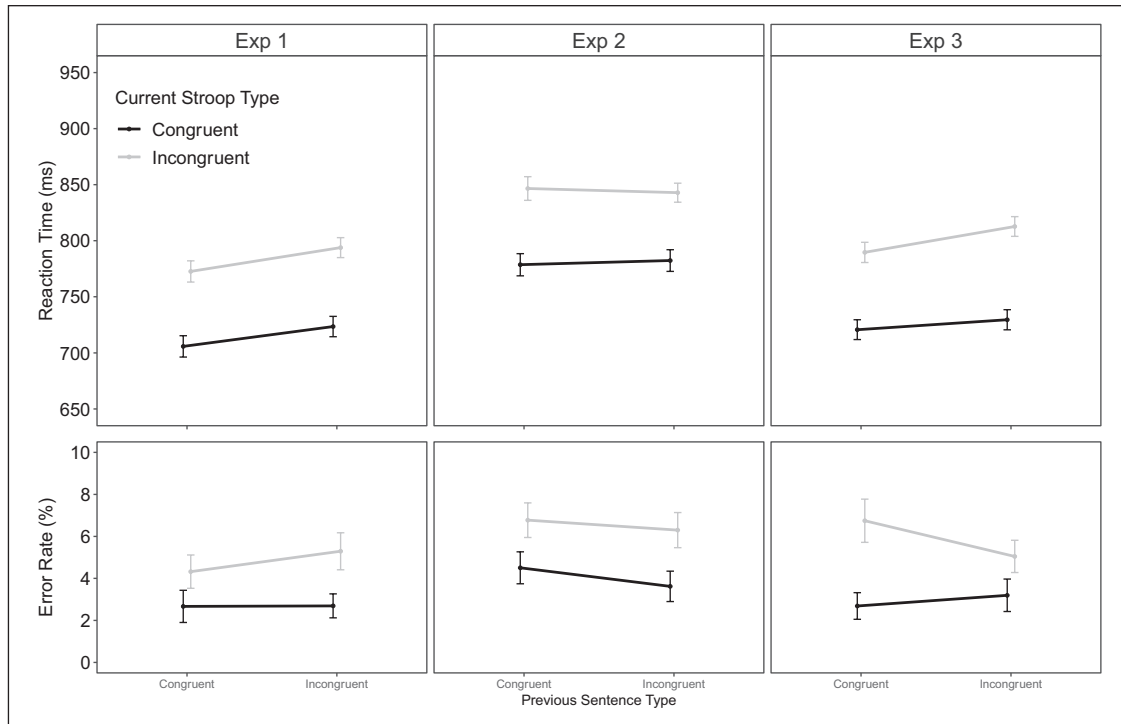


Figure 3. Stroop condition means for reaction time (RT; top row) and error rate (ER; bottom row) as a function of Previous Sentence Congruency and Current Stroop Congruency within Part I of Experiments 1 (left), 2 (middle) and 3 (right) column. The error bars represent ± 1 within-subject standard error (Morey, 2008).

to previous sentence congruency and vice versa. That is, Stroop trials were coded according to trial history (congruent vs. incongruent) with RT and ER being analysed with a repeated-measures ANOVA with the within-subject factors previous sentence congruency (congruent vs. incongruent) and current Stroop congruency (congruent vs. incongruent). In addition, reading times for the critical region 3 were analysed with a repeated-measures ANOVA with the within-subject factors previous Stroop congruency (congruent vs. incongruent) and current sentence congruency (congruent vs. incongruent).

Reading task (Part 1). Extreme data points accounted for less than 0.5% of the data points and were excluded. Subsequent outliers (~3.4% of data points) were replaced by the respective cut-off values. Condition means for the three regions are displayed in Figure 2 (left column). The congruency effect (13 ms) was significant in Region 3, $t(67)=3.53$, $p < .001$, $d_z=0.43$, 95% CI=[6, 20]ms.⁴ Region 1 demonstrated a significant congruency effect in the reverse direction, $t(67)=2.07$, $p=.04$, $d_z=0.25$, 95% CI=[-8, 0]ms, with 4 ms shorter reading times in the incongruent (264 ms) than the congruent condition (268 ms).⁵ Region 2 did not show a significant difference, $t(67)=-0.4$, $p=.67$, $d_z=0.05$, 95% CI=[-5, 3]ms.

Stroop task (Part 1). For RT, there was a significant Stroop effect with faster responses to congruent (656 ms) than

incongruent trials (723 ms), $t(67)=7.46$, $p < .001$, $d_z=0.90$, 95% CI=[49, 85]ms. For ER, there was also a significant Stroop effect with more errors to incongruent trials (3.6%) compared with congruent trials (2.1%), $t(67)=3.79$, $p < .001$, $d_z=0.46$, 95% CI=[0.7, 2.2]%

Sentence-to-Stroop CSE. Condition means are displayed in Figure 3 (left column). For RT, there was a significant main effect for Current Stroop Congruency, $F(1, 67)=40.51$, $p < .001$, $\eta_p^2=0.38$, with longer RTs to incongruent trials (783 ms) compared with congruent trials (715 ms). There was no Previous Sentence Congruency \times Current Stroop Congruency interaction, $F(1, 67)=0.04$, $p=.84$ ($BF_{10}=0.19$). For ER, there was a significant main effect for current Stroop congruency, $F(1, 67)=8.81$, $p=.004$, $\eta_p^2=0.12$, which was not modulated by previous sentence congruency, $F(1, 67)=0.45$, $p=.504$ ($BF_{10}=0.03$). Thus, both RT and ER findings indicated the absence of a cross-task CSE.

Stroop to sentence CSE. One participant did not show data in all four CSE conditions after removing outliers and was excluded from data analysis. A repeated-measures ANOVA with the within-subject factors current sentence congruency and previous Stroop congruency revealed no significant interaction in reading time in the final sentence region, $F(1, 66)=0.07$, $p=.796$. The congruency effect was of similar magnitude both with preceding congruent and incongruent Stroop trials (~12 to 14 ms).

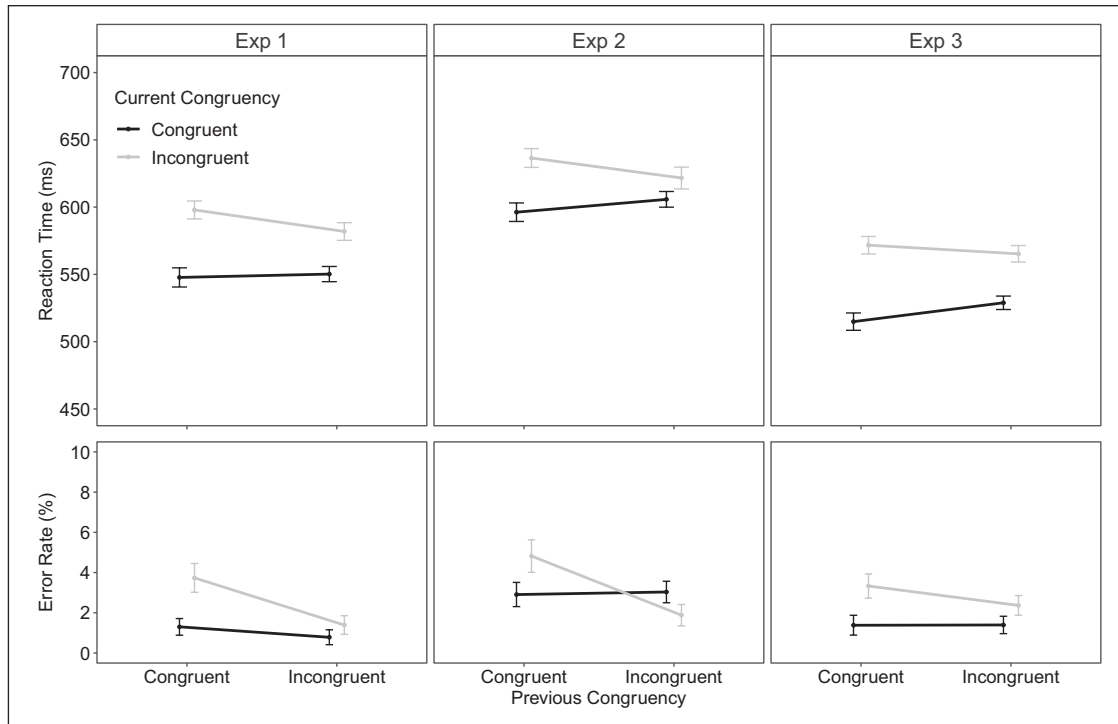


Figure 4. Stroop condition means for reaction time (RT; top row) and error rate (ER; bottom row) as a function of Previous Trial Congruency and Current Trial Congruency within Part 2 of Experiments 1 (left), 2 (middle) and 3 (right) column. The error bars represent ± 1 within-subject standard error (Morey, 2008).

Stroop task CSE. Less than 1% of trials were excluded as outliers ($\sim 0.80\%$). Condition means are displayed in Figure 4 (middle column). For RT, there was a significant main effect of current Stroop congruency with faster responses to congruent trials (549ms) compared with incongruent trials (590ms), $F(1, 67)=20.54$, $p < .001$, $\eta_p^2 = .23$. The interaction between current Stroop congruency and previous Stroop congruency was significant, $F(1, 67)=4.26$, $p = .04$, $\eta_p^2 = .06$, indicating a CSE of 18 ms.

For ER, there was a significant main effect of current Stroop congruency, $F(1, 67)=8.74$, $p = .004$, $\eta_p^2 = .12$, with more errors to incongruent (2.6%) than to congruent trials (1.0%). The Current Stroop Congruency \times Previous Stroop Congruency interaction was not significant, $F(1, 67)=3.59$, $p = .062$, $\eta_p^2 = .05$. However, numerically, the congruency effect was larger following congruent trials (2.4%) compared with incongruent trials (0.6%).

Experiment 2

Experiment 2 was identical to Experiment 1, with the following exception: the word-by-word self-paced reading task employed a routine whereby the mask also hid words that have already been read (see Figure 1, middle column). Essentially, such a procedure removes the opportunity for participants to reread the sentence and hampers its reinterpretation.

Participants

Seventy-seven American English native speakers (31 females, $M_{\text{age}} = 40.39$, $SD_{\text{age}} = 10.17$, 73 right-handed) participated, with seventy-four full datasets from both parts 1 and 2 remaining. In addition, six participants were removed from the subsequent analysis due to poor performance in the Stroop task during Part 2 (ERs $> 20\%$). Thus, 68 participants remained in the subsequent analyses.

Results

Reading task (Part 1). Extreme data points ($> 2,000$ ms) occurred in less than 0.5% of the data, with subsequent outliers being adjusted according to the ± 2.5 SD cut-off ($\sim 3.2\%$ of the data). The condition means for the three regions are displayed in Figure 2 (middle column). The congruency effect was significant in Region 3, $t(67)=3.09$, $p < .01$, $d_z = 0.37$, 95% CI=[2, 11]ms, but not in Region 1, $t(67)=0.14$, $p = .89$, $d_z = 0.02$, 95% CI=[-4, 4]ms, or in Region 2 and $t(67)=0.20$, $p = .84$, $d_z = 0.02$, 95% CI=[-4, 3]ms.

Stroop task (Part 1). Paired t -tests were conducted on the Stroop trials both for reaction times (RT) and error rates (ER). For RT, there was a significant Stroop Effect with longer RTs to congruent trials (714ms) compared with incongruent trials (766ms), $t(67)=5.62$, $p < .001$, $d_z = 0.68$,

95% CI=[33, 70]ms. For ER, there was also a significant Stroop Effect with more errors to incongruent trials (4.9%) compared with congruent trials (3.0%), $t(67)=3.29$, $p=.002$, $d_z=0.40$, 95% CI=[0.8, 3.1]%

Sentence-to-Stroop CSE. Again, Stroop trials were subsequently coded according to $n-1$ trial history (Congruent vs. Incongruent) with RT and ER being analysed with a repeated-measures ANOVA. Condition means are displayed in Figure 3 (middle column). For RT, there was a significant main effect for current Stroop congruency, $F(1, 67) = 33.09$, $p < .001$, $\eta_p^2 = .33$, with longer RTs to incongruent trials (845 ms) compared with congruent trials (780 ms). There was no significant interaction between previous sentence congruency and current Stroop congruency, $F(1, 67) = 0.22$, $p = .638$, $\eta_p^2 < .01$ ($BF_{10} = 0.02$).

For ER, again, there was a significant main effect for current Stroop congruency, $F(1, 67)=7.35$, $p=.009$, $\eta_p^2=0.10$, but no reliable interaction with previous sentence congruency, $F(1, 67) = 0.09$, $p = .760$ ($BF_{10} = 0.03$).

Stroop to sentence CSE. The repeated-measures ANOVA for reading time revealed no significant interaction of previous Stroop congruency and current sentence congruency in the final sentence region, $F(1, 67)=1.21$, $p=.275$, $\eta_p^2=.02$; if anything, the congruency effect was numerically larger with a preceding incongruent Stroop trial (8 ms) than a congruent one (4 ms).

Stroop task CSE (Part 2). RT outliers (~1.4%) were removed from the analysis. Condition means are displayed in Figure 4. For RT, there was a significant main effect of current Stroop congruency with faster responses to congruent (601 ms) than to incongruent trials (629 ms), $F(1, 67)=12.82$, $p < .001$, $\eta_p^2 = .16$. The Current Stroop Congruency \times Previous Stroop Congruency approached significance, $F(1, 67)=3.75$, $p=.057$, $\eta_p^2 = .05$; numerically, the congruency effect was larger following congruent trials (40 ms) than incongruent trials (16 ms). For ER, there was no main effect of Current Trial Type, $F < 1$, however, the interaction between current Stroop congruency and previous Stroop congruency was significant, $F(1, 67)=5.92$, $p=.018$, $\eta_p^2 = .08$, with the congruency effect being larger following congruent trials (1.8%) compared with incongruent trials (-1.0%).

Experiment 3

Experiment 3 was identical to Experiment 1 and 2, with the following exception: the word-by-word self-paced reading task employed a routine whereby each word within the sentence was presented centrally (see Figure 1, right column). This is a well-established methodology in EEG/ERP studies (e.g., Kutas & Hillyard, 1980), and, in contrast with the other two presentation versions, it does not provide advanced knowledge regarding sentence length.

Participants

Seventy-seven American English native speaker participants' data (34 females, $M_{age} = 43.79$, $SD_{age} = 11.84$, 68 right-handed) participated, with seventy-six complete datasets from both Parts 1 and 2 remaining. In addition, one participant was removed from the subsequent analysis due to poor performance in the Stroop task during Part 2 (ERs > 20%). Thus, seventy-five participants remained in the subsequent analyses.

Results

Reading task (Part 1). The same data analysis procedures were applied as in previous experiments. Extreme data points (>2,000 ms) accounted for less than 0.1% of the data, with subsequent outliers being adjusted according to the $\pm 2.5 SD$ cut-off (~2.5% of the data). The condition means for the three regions are displayed in Figure 2 (right column). The congruency effect was not significant in Region 3, $t(74)=1.03$, $p=.308$, $d_z=0.12$, 95% CI=[-6, 2]ms. Regions 1 and 2 also showed no significant effects, $t(74)=0.76$, $p=.448$, $d_z=0.09$, 95% CI=[-2, 4]ms and $t(74)=0.80$, $p=.427$, $d_z=0.09$, 95% CI=[-2, 4]ms, respectively.

Stroop task (Part 1). For RT, there was a significant Stroop effect with faster responses to congruent trials (669 ms) compared with incongruent trials (745 ms), $t(74)=8.43$, $p < .001$, $d_z=0.97$, 95% CI=[58, 94]ms. There was also a significant Stroop Effect for ER with more errors to incongruent trials (4.3%) compared with congruent trials (2.6%), $t(74)=3.71$, $p < .001$, $d_z=0.43$, 95% CI=[0.8, 2.6]%

Sentence-to-Stroop CSE. Condition means are displayed in Figure 3 (right column). For RT, there was a significant main effect of current Stroop congruency, $F(1, 74)=53.74$, $p < .001$, $\eta_p^2=0.42$, with longer RTs to incongruent (801 ms) than to congruent trials (725 ms). There was no significant interaction between previous sentence congruency and current Stroop congruency, $F(1, 74)=.94$, $p=.335$, $\eta_p^2=.01$ ($BF_{10}=0.13$).

For ER, the main effect of current Stroop congruency was significant, $F(1, 74)=15.09$, $p < .001$, $\eta_p^2=0.17$, with more errors to incongruent (5.9%) than to congruent trials (2.9%). The interaction with previous sentence congruency was not significant, $F(1, 74)=1.92$, $p=.170$, $\eta_p^2=0.03$ ($BF_{10}=0.07$).

Stroop to sentence CSE. Like in Experiments 1 and 2, the repeated-measures ANOVA of reading time revealed no significant interaction of previous Stroop congruency and current sentence congruency in the final sentence region, $F(1, 74)=1.21$, $p=.274$, $\eta_p^2=.02$.

Stroop task (Part 2). We removed 0.8% of the trials as outliers. Condition means are displayed in Figure 4. For RT,

there was a significant main effect of current Stroop congruency with faster responses to congruent Trials (522 ms) compared with incongruent trials (568 ms), $F(1, 74)=43.09$, $p < .001$, $\eta_p^2 = .37$. The interaction between current Stroop congruency and previous Stroop congruency was not significant, $F(1, 74)=3.04$, $p = .085$, $\eta_p^2 = .04$. However, numerically, the congruency effect was larger when the previous trial was congruent (57 ms) compared with incongruent (36 ms).

For ER, there was a significant main effect of current Stroop congruency, $F(1, 74)=7.04$, $p < .01$, $\eta_p^2 = .09$, with more errors to incongruent trials (2.8%) than to congruent trials (1.4%), which was not reliably modulated by previous Stroop congruency, $F(1, 74)=1.06$, $p = .307$, $\eta_p^2 = .01$.

Combined analysis (Experiments 1–3)

Experiments 1 to 3 employed an identical design, with the only changes being the type of presentation mode employed in the self-paced reading task of Part 1 of the experiments. Thus, an analysis of the combined data was conducted.

Reading task. Separate ANOVAs for the three regions showed a significant main effect of Experiment for region 1, $F(2, 208)=3.55$, $p = .03$, $\eta_p^2 = .02$, but not for region 2, $F(2, 208)=2.23$, $p = .110$, and for region 3, $F(1, 208)=1.53$, $p = .218$. The effect of presentation mode for region 1 is mainly driven by the shorter reading times in Experiment 1 where stimuli remained on screen. The congruency effect was significant in Region 3, $t(210)=4.48$, $d_z = 0.31$, 95% CI=[4, 10]ms, whereas no significant difference was observed in Regions 1, ($t(210)=1.54$, $p = .124$, $d_z = 0.11$, or in Region 2, $t(210)=0.80$, $p = .426$, $d_z = 0.06$).

Sentence-to-Stroop CSE. The repeated-measures ANOVA of Stroop performance (RT, ER) showed a main effect of current Stroop congruency with shorter RTs to congruent trials (740 ms) than to incongruent trials (809 ms), $F(1, 210)=127.16$, $p < .001$, $\eta_p^2 = .38$. However, the interaction between current Stroop congruency and previous sentence congruency was not significant, $F(1, 210)=0.17$, $p = .679$. For ER, there was a main effect of current Stroop congruency with fewer errors to congruent (3.2%) compared with incongruent (5.7%) trials. Like for RT, the interaction between current Stroop congruency and previous sentence congruency was not significant, $F(1, 210)=0.17$, $p = .679$.

Stroop to sentence. The repeated-measures ANOVA of reading time revealed no significant interaction of previous Stroop congruency and current sentence congruency in the final sentence region, $F(1, 209)=1.01$, $p = .315$.

All three experiments failed to demonstrate evidence for Sentence-to-Stroop CSEs. One might argue that the conflict experienced within the sentences was not strong enough ($d_z = 0.43$, 0.37, and 0.12 across Experiments 1–3, respectively) to trigger a CSE in the Stroop task. Therefore, we selected a subset of participants with longer reading times within Region 3 for incongruent compared with congruent sentences was selected, Exp 1: $N=48$ out of 68, $t(47)=9.30$, $p < .001$, $d_z = 1.34$, 95% CI=[26, 40]ms; Exp 2: $N=41$ out of 68, $t(39)=7.31$, $p < .001$, $d_z = 1.16$, 95% CI=[15, 26]ms; Exp 3: $N=38$ out of 75, $t(37)=7.38$, $p < .001$, $d_z = 1.20$, 95% CI=[13, 22]ms. However, when the CSE was analysed using this subset of participants, still no evidence for a Sentence-to-Stroop conflict-related adaptation effect was observed within each experiment, all $F_s < 0.28$, $p_s > 0.602$, and also not in the analysis of the combined data sets, $F(1, 126)=0.03$, $p = .870$.

Stroop task CSE (Part 2). Here, a repeated-measures ANOVA within the within-subject factors Previous Trial Type (Congruent vs. Incongruent) and Current Trial Type (Congruent vs. Incongruent) was performed for both RT and ER. For RT, there was a significant main effect of Current Trial Type with faster responses to congruent Trials (556 ms) compared with incongruent trials (595 ms), $F(1, 210)=70.67$, $p < .001$, $\eta_p^2 = .25$. The interaction between Current Trial Type and Previous Trial Type was significant, $F(1, 210)=10.63$, $p < .001$, $\eta_p^2 = .05$. Here, the congruency effect was larger when the preceding trial was congruent (49 ms) compared with incongruent (28 ms). For ER, there was a significant main effect of Current Trial Type with more errors to incongruent (2.9%) compared with congruent (1.8%) trials, $F(1, 210)=12.02$, $p < .001$, $\eta_p^2 = .05$. The interaction between Previous Trial Type and Current Trial Type was also significant, $F(1, 210)=9.86$, $p = .002$, $\eta_p^2 = .04$, with the congruency effect being larger when Previous Trial Type was congruent (2.1%) compared with incongruent (0.2%).

Discussion

Experiments 1 to 3 investigated whether—in line with Kan et al. (2013)—the processing of sentences containing representational (semantic) conflicts would result in the recruitment of control processes that subsequently enable the reader to deal with another conflict (e.g., incongruent Stroop trial) more efficiently. This particular result would provide strong support for the notion of domain-general control mechanisms. In the Kan et al. study, intermixing syntactically ambiguous sentences and Stroop trials, they showed conflict-related adaptation effects in both RT and ER. In contrast, none of our three experiments provided any evidence for such cross-task conflict adaptation (cf. Table 2). Indeed, the results showed that semantic conflict resulted in increased reading times in the disambiguating

region in two out of three experiments. However, no CSE was found from the sentence reading to the Stroop task and also not from the Stroop task to the sentence reading task. These results are inconsistent with the idea of a domain-general mechanism of cognitive control adjustments following conflict detection, at least as far as the present semantic conflict condition is concerned. Regarding the separate Stroop task run at the end of each experiment as a control condition, the results showed a CSE. Hence, at least the absence of the cross-task CSE in the current Stroop task cannot be attributed to this task being insensitive to reveal typical conflict adaptation effects in RT and ER since the present experiments by and large demonstrated within-Stroop task CSE patterns in either RT and/or ER. Given these results, in Experiments 4 to 6, we investigate whether conflict-related control adjustments are present within a sentence reading task involving semantically ambiguous (incongruent) and unambiguous (congruent) sentences. Specifically, we addressed the question of whether conflict experienced during processing ambiguous sentences results in processing adjustments that enable the reader to better deal with subsequent ambiguity. Previous studies showed that syntactically ambiguous sentences can trigger conflict adjustments in Stroop trials (Kan et al., 2013) and that a conflict in a Flanker task can adjust subsequent reading processing in ambiguous sentences (Hsu et al., 2021). However, to the best of our knowledge, no study addressed whether such adjustments can be observed within a purely linguistic reading task.

Experiment 4

The procedure in Experiment 4 differed from that of Experiment 1 only in that, in the first part of the experiment, only the reading task was presented. As in Experiment 1, each trial began with a full mask represented by a string of underscore dashes replacing each word. The sentences were presented in 20px monospaced font and did not require line breaks. The dashes were substituted by the first/next word in the sentence with each subsequent keypress. In this version of the word-by-word reading task, as each word appeared, the previous words remained on the screen until the end of the trial (see Figure 1, left column). Before starting the experiment, participants could familiarise themselves with the task, with one sample sentence demonstrating the self-paced reading procedure.

Participants

Seventy-two American English native speakers participated (36 females, $M_{age}=44.03$, $SD_{age}=14.56$, 69 right-handed) with sixty-five full datasets from both Parts 1 and 2 remaining. The performance exclusion criterion resulted in three participants being removed, with two participants essentially responding randomly in the Stroop task (46%

and 50% errors). Thus, 62 participants remained in the analysis reported below.

Results

Reading task. Extreme values at the word level occurred in less than 1% of data, while ~4% of data points were replaced by the respective cut-off values. Condition means for the three regions are displayed in Figure 5 (left column). The congruency effect was significant in Region 3, $t(61)=3.44$, $p=.001$, $d_z=0.44$, 95% CI=[7, 28]ms,⁶ whereas it was not significant within Region 1 and Region 2, $t(61)=0.08$, $p=.94$, $d_z=0.01$, 95% CI=[-5, 6]ms, and $t(61)=1.10$, $p=.27$, $d_z=0.14$, 95% CI=[-3, 9]ms, respectively.

Reading task CSE. Average reading times in Region 3 were analysed using a 2 Current Sentence Congruency \times 2 Previous Sentence Congruency repeated-measures ANOVA.⁷ Condition means are displayed in Figure 6 (left column). There was a significant main effect of current sentence congruency, $F(1, 58)=16.56$, $p<.001$, $\eta_p^2=0.22$, with longer reading times to incongruent trials (357ms) compared with congruent trials (338ms). Critically, this main effect was not modulated by previous sentence congruency, $F<1$ ($BF_{10}=0.03$), suggesting that the CSE was absent.

Stroop task CSE. Less than 1% of trials were excluded as outliers. Condition means are displayed in Figure 7 (left column). For RT, there was a significant main effect of current Stroop congruency with faster responses to congruent trials (510ms) compared with incongruent trials (545ms), $F(1, 61)=30.24$, $p<.001$, $\eta_p^2=.33$. The interaction between current and previous Stroop congruency, although not significant, demonstrated numerically the expected CSE in RT, $F(1, 61)=3.72$, $p=.06$, $\eta_p^2=.06$, with a larger congruency effect when the preceding Stroop trial was congruent (43ms) compared with incongruent (26ms).

For ER, there were numerically more errors when the Current Trial Type was incongruent (4.0%) compared with congruent (2.6%), $F(1, 61)=3.04$, $p=.086$, $\eta_p^2=.05$. The interaction between current Stroop congruency and previous Stroop congruency was significant, $F(1, 61)=9.18$, $p=.004$, $\eta_p^2=.13$, with the congruency effect being larger when the preceding Stroop trial was congruent (3.4%) compared with incongruent (-0.6%). This result indicated a standard CSE in ER within the current manual Stroop task.

Experiment 5

Experiment 5 was identical to Experiment 4 with the following exception: the word-by-word self-paced reading

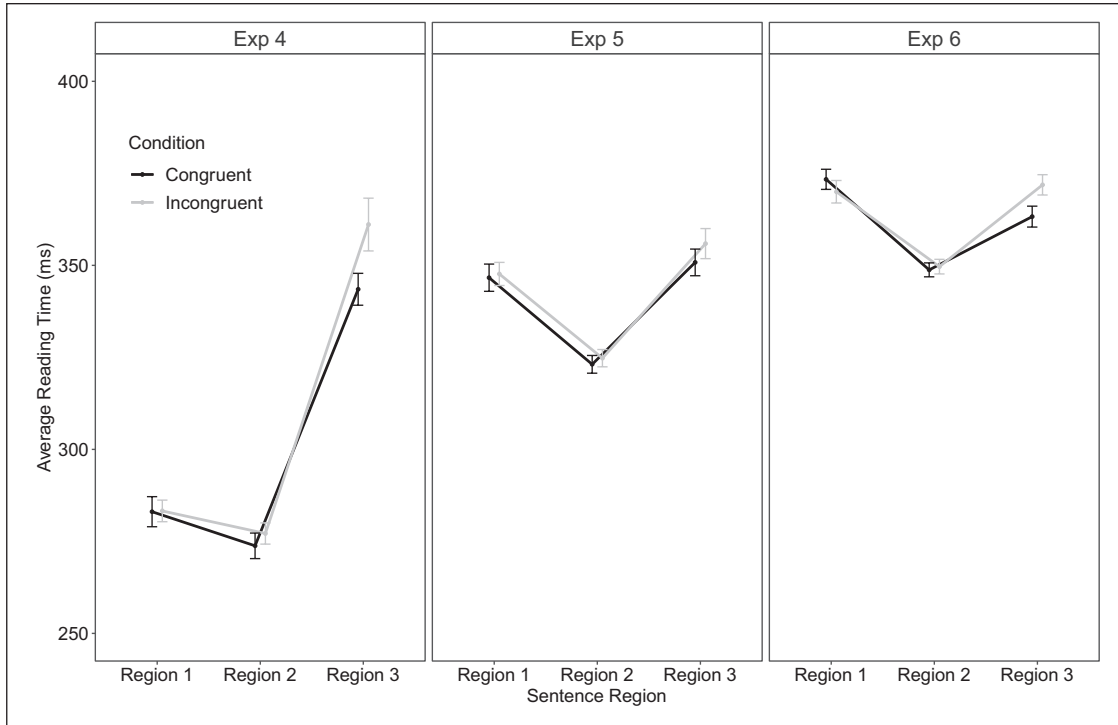


Figure 5. Average reading time (ms) as a function of sentence region and congruency condition within Part I of Experiments 4 (left), 5 (middle) and 6 (right) column. The error bars represent ± 1 within-subject standard error (Morey, 2008).

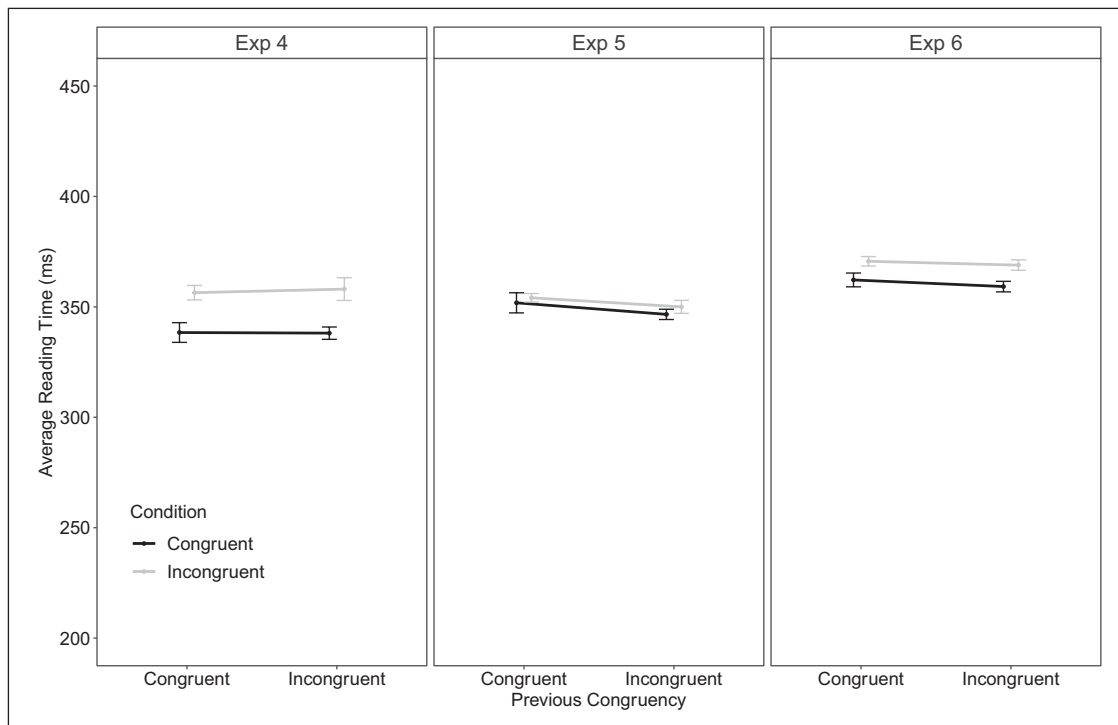


Figure 6. Average reading time within sentence region 3 as a function of previous trial congruency and current trial congruency within Part I of Experiments 4 (left), 5 (middle) and 6 (right) column. The error bars represent ± 1 within-subject standard error (Morey, 2008).

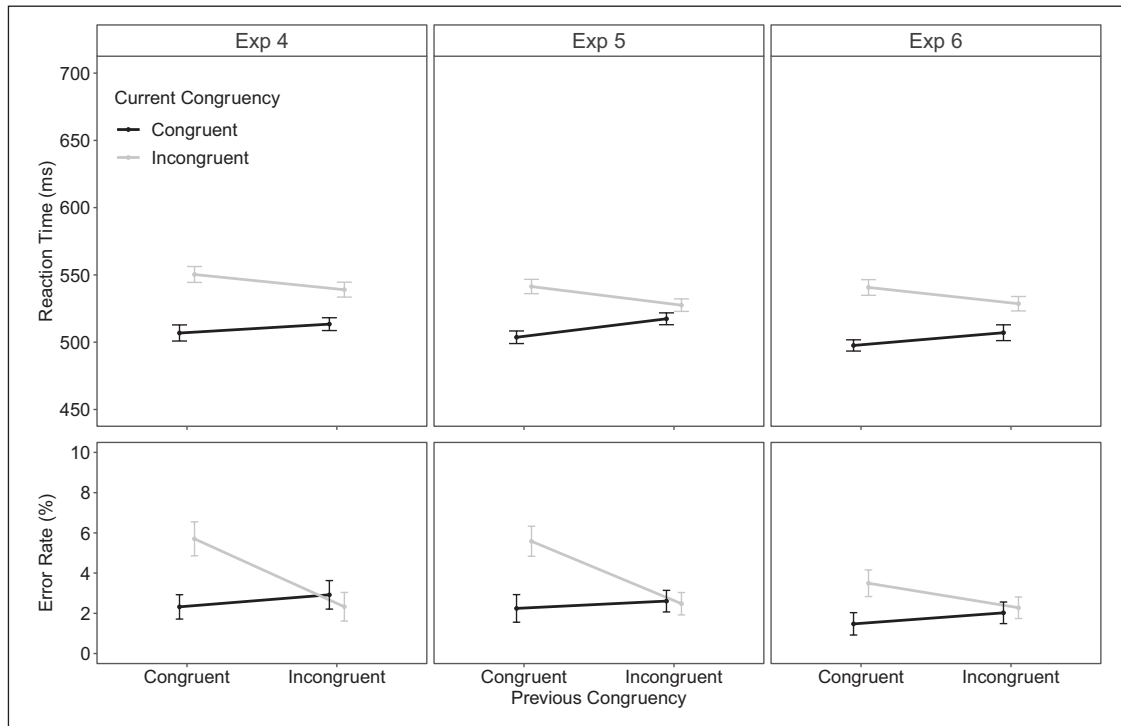


Figure 7. Stroop task condition means for reaction time (RT; top row) and error rate (ER; bottom row) as a function of Previous Trial Congruency and Current Trial Congruency within Part 2 of Experiments 4 (left), 5 (middle) and 6 (right) column. The error bars represent ± 1 within-subject standard error (Morey, 2008).

task employed a routine whereby the mask also hid words that have already been read (see Figure 1, middle column). Essentially, such a procedure removes the opportunity for participants to reread the sentence and hampers its reinterpretation.

Participants

Seventy-one American English native speaker participated (37 females, $M_{age} = 40.07$, $SD_{age} = 10.25$, 67 right-handed), with sixty-five full datasets from both parts 1 and 2 remaining. The exclusion of participants with poor Stroop performance ($> 20\%$ overall error rate) resulted in the additional exclusion of one participant (ER=48%). Thus, sixty-four participants remained in the subsequent analysis.

Results

Reading task. Extreme data points accounted for less than 0.3% of the data, while less than 3% of data points were replaced by the respective cut-off values. Condition means for the three regions are displayed in Figure 5 (middle column). Paired t -tests showed a significant congruency effect in Region 3, $t(63)=2.14$, $p=.036$, $d_z=0.27$, 95% CI=[0.35, 10]ms.⁸ No significant congruency effect was found for Regions 1 or 2, $t(63)=0.29$, $p=.772$, $d_z=0.04$, 95% CI=[-6, 8]ms, and $t(63)=0.68$, $p=.496$, $d_z=0.09$, 95% CI=[-3, 7]ms, respectively.

Reading task CSE. Condition means of average reading times in Region 3 are displayed in Figure 6 (middle column). Neither the main effect of current sentence congruency, $F(1, 63)=0.94$, $p=.336$, $\eta_p^2=.01$,⁹ nor the interaction with previous sentence congruency was significant, $F(1, 63)=0.05$, $p=.826$, $\eta_p^2=.02$ ($BF_{10}=0.07$).

Stroop task CSE. Less than 1% of trials were removed as outliers. Condition means are displayed in Figure 7 (middle column). For RT, there was a significant main effect of current Stroop congruency with faster responses to congruent (511 ms) than incongruent trials (534 ms), $F(1, 63)=20.78$, $p<.001$, $\eta_p^2=.25$. The interaction between current Stroop congruency and previous Stroop congruency was significant, $F(1, 63)=8.13$, $p=.006$, $\eta_p^2=.11$, indicating a larger Stroop effect when the preceding Stroop trial was congruent (38 ms) compared with incongruent (10 ms).

For ER, there was a significant main effect of current Stroop congruency with more errors to incongruent (4.0%) compared with congruent (2.4%) trials, $F(1, 63)=5.50$, $p=.022$, $\eta_p^2=.08$. Again, the interaction between current Stroop congruency and previous Stroop congruency was significant, $F(1, 63)=7.96$, $p=.006$, $\eta_p^2=.11$. The Stroop effect was larger when the previous Stroop trial was congruent (3.3%) compared with incongruent (-0.1%). Overall, the present findings again suggest that within the Stroop task the standard CSE was observed.

Experiment 6

Experiment 6 was identical to Experiments 4 and 5 except that it used the non-cumulative presentation mode of Experiment 3 (see Figure 1, right column).

Participants

Seventy-eight American English native speaker participated (41 females, $M_{age}=41.94$, $SD_{age}=10.95$, 70 right-handed), with seventy-three full datasets from both parts 1 and 2 remaining. The exclusion of participants with poor Stroop performance ($>20\%$ overall error rate) resulted in the additional exclusion of three participants (ERs=40%, 31%, 21%). Thus, 70 participants remained in the subsequent analysis.

Results

Reading task. Extreme data points accounted for less than 0.2% of the data, while less than 3% of data points were replaced by the respective cut-off values. Condition means for the three regions are displayed in Figure 5 (right column). The congruency effect was significant in Region 3, $t(69)=3.43$, $p<.001$, $d_z=0.41$, 95% CI=[4, 14]ms.¹⁰ Regions 1 and 2 showed no significant congruency effects, $t(69)=-1.45$, $p=.151$, $d_z=0.17$, 95% CI=[-8, 1]ms, and $t(69)=0.43$, $p=.666$, $d_z=0.05$, 95% CI=[-3, 5]ms, respectively.

Reading task CSE. Condition means of average reading times in Region 3 are displayed in Figure 6 (right column). There was a significant main effect for current sentence congruency, $F(1, 69)=14.57$, $p<.001$, $\eta_p^2=0.17$, with longer reading times to incongruent (370 ms) than congruent sentences (361 ms). The interaction between current sentence congruency and previous sentence congruency was not significant, $F(1, 69)=.06$, $p=.809$ ($BF_{10}=0.04$), suggesting in line with results of Experiment 1 and 2 that no within-task CSE was present.

Stroop task CSE. Less than 1% of trials were excluded as outliers. Condition means are displayed in Figure 7 (right column). For RT, there was a significant main effect of current Stroop congruency with faster responses to congruent trials (502 ms) compared with incongruent trials (535 ms), $F(1, 69)=26.06$, $p<.001$, $\eta_p^2=.27$. The interaction between current Stroop congruency and previous Stroop congruency was significant, $F(1, 69)=4.33$, $p=.041$, $\eta_p^2=.06$, indicating the CSE as reflected by a larger congruency effect when the preceding Stroop trial was congruent (43 ms) than when it was incongruent (22 ms).

For ER, there was a significant main effect of current Stroop congruency with more errors to incongruent (2.9%)

compared with congruent (1.8%) trials, $F(1, 69)=4.31$, $p=.04$, $\eta_p^2=.06$. The interaction between current Stroop congruency and previous Stroop congruency was not significant, $F(1, 69)=2.13$, $p=.149$, $\eta_p^2=.03$. However, the congruency effect was numerically larger when the preceding Stroop trial was congruent (2.0%) compared with incongruent (0.3%).

Combined analysis (Experiments 4–6)

Like for Experiments 1 to 3, an analysis of the combined data sets was conducted for Experiments 4 to 6.

Reading task. Separate ANOVAs for the three regions showed a significant main effect of Experiment for Region 1, $F(2, 193)=14.33$, $p<.001$, $\eta_p^2=.13$, for Region 2, $F(2, 193)=12.24$, $p<.001$, $\eta_p^2=.11$, but not for Region 3, $F(1, 193)=0.47$, $p=.63$. The effect of presentation mode is mainly driven by the shorter reading times in Experiment 1 where stimuli remained on screen. Concerning the congruency effect, this was significant in Region 3, $t(195)=5.08$, $d_z=0.36$, 95% CI=[6, 14]ms, $p<.001$, but not in Regions 1 and 2, $t(195)=-0.49$, $p=.627$, $d_z=0.04$, and $t(195)=1.34$, $p=.183$, $d_z=0.10$, respectively.

Reading task CSE. The 2 Current Sentence Congruency \times 2 Previous Sentence Congruency repeated-measures ANOVA on mean reading time within Region 3 revealed a main effect of Current Congruency, $F(1, 192)=25.86$, $p<.001$, $\eta_p^2=.12$, due to faster reading times in congruent (350 ms) than in incongruent sentences (360 ms). There was no interaction between current sentence congruency and previous sentence congruency, $F(1, 192)=0.19$, $p=.67$.

All three experiments, as well as the combined analysis failed to demonstrate evidence for Sentence-to-Sentence CSEs. However, again, one might argue that the conflict experienced within the sentences experienced by some participants might not have been strong enough to trigger control adjustments. Indeed, while the conflict effect within Region 3 was significant in each of the experiments, the effects were relatively small ($d_z=0.44$, 0.27, and 0.41, respectively). As for Experiments 1 to 3, we selected within each experiment only those participants who showed longer reading times within Region 3 for incongruent compared with congruent sentences, Experiment 4, $N=44$ out of 62, $t(44)=6.23$, $p<.001$, $d_z=0.93$, 95% CI=[30, 52]ms; Experiment 5: $N=41$ out of 64, $t(40)=5.92$, $p<.001$, $d_z=0.93$, 95% CI=[9, 19]ms; Experiment 6: $N=50$ out of 70, $t(49)=8.14$, $p<.001$, $d_z=1.15$, 95% CI=[13, 22]ms. However, when the CSE was analysed using this subset of participants, still no evidence for Sentence-to-Sentence conflict-related adaptation effect was observed within each experiment, all $F_s < 1.47$, $p_s > 0.232$, and also not in the analysis of the combined data sets, $F(1, 133)=0.21$, $p=.65$.

Stroop task CSE (Part 2). The 2 Current Stroop Congruency \times 2 Previous Stroop Congruency repeated-measures ANOVA indicated faster responses to congruent (507 ms) than incongruent Stroop trials (538 ms), $F(1, 195)=76.54$, $p < .001$, $\eta_p^2=.28$. Crucially, the interaction between current Stroop congruency and previous Stroop congruency was significant, $F(1, 195)=15.62$, $p < .001$, $\eta_p^2=.07$, due to larger congruency effect when the preceding Stroop trial was congruent (41 ms) than when it was incongruent (19 ms).

For ER, there was a significant main effect of current Stroop congruency with more errors to incongruent (3.6%) compared with congruent (2.3%) trials, $F(1, 195)=12.47$, $p=.001$, $\eta_p^2=.06$. Also, the interaction between current Stroop congruency and previous Stroop congruency was significant, $F(1, 195)=17.57$, $p < .001$, $\eta_p^2=.08$, with the Stroop effect being larger when the preceding trial was congruent (2.9%) compared with incongruent (-0.1%).

Discussion

Experiments 4 to 6 investigated whether conflict-related control adjustments are present within a sentence reading task involving semantically ambiguous (incongruent) and unambiguous (congruent) sentences. Specifically, we addressed the question of whether conflict experienced during processing ambiguous sentences results in processing adjustments that enable the reader to better deal with subsequent ambiguity. Previous studies showed that syntactically ambiguous sentences can trigger conflict adjustments in Stroop trials (Kan et al., 2013), and that a conflict in a Flanker task can adjust subsequent reading processing in ambiguous sentences (Hsu et al., 2021). However, to the best of our knowledge, no study addressed whether such adjustments can be observed within a purely linguistic reading task. Importantly, the sentence manipulation utilising semantic ambiguity demonstrated the expected conflict pattern in our reading task (congruency effect). That is, we found significantly longer reading times in the incongruent than the congruent condition in the region following the disambiguating word in all three experiments. As expected, the previous regions did not show a difference, indicating that the disambiguating word triggered a conflict with regard to the initial interpretation within incongruent sentences (e.g., “The old man headed for the bank but he had a long way to swim to finally reach it.”).

Crucially, however, whereas semantic conflict prolonged reading times, the presumable detection of conflict did not trigger any adjustment in subsequent reading behaviour. That is, in all three experiments, the congruency effect in the current sentence was not modulated by the sentence congruency of the preceding trial, that is, whether or not it involved a semantic ambiguity. Thus, the three different self-paced presentation modes, that is, the cumulative mode with words remaining on the screen, and

the two non-cumulative modes with words disappearing and being masked with dashes after each space bar press or being presented at the screen’s centre and replaced by the subsequent word, did not influence these core findings. However, in the pure Stroop task block of the three experiments, we found a standard CSE either in RT and/or ER. Although this task is not of key theoretical relevance for this study, it clearly indicates that a CSE can be observed in this task and for the sample tested. Thus, the Stroop task used in the present experiments can in principle reveal a CSE and the observation of an absent within-task CSE cannot be attributed to participants being unable to adapt their processing following a conflict. We will elaborate on the current absence of a within-task CSE in the General Discussion.

General discussion

this study had two main goals. First, we wanted to investigate whether cross-task adaptation effects from sentence reading tasks to Stroop tasks could be observed (Experiments 1–3)—in line with the findings reported by Kan et al. (2013)—when using semantic conflict (see Blott et al., 2020) rather than syntactic conflict (as in Kan et al., 2013). All our experiments were accompanied by a CSE control task condition (the Stroop block at the end of each experiment)—ensuring that with the present experimental setup, it was possible to sensitively measure the conflict adaptation effects within the current manual Stroop task.

Second, we investigated whether standard CSEs can be observed within a language task with semantic ambiguity (Experiments 4–6). It is surprising that despite cross-task CSEs being observed from a sentence reading task to the Stroop task (i.e., Kan et al., 2013) and from a flanker task to a sentence comprehension task (Hsu et al., 2021), as well as the common finding of within-task CSEs (for a review, see Braem et al., 2014), there are to our knowledge no studies examining conflict adaptation effects in terms of a CSE within a sentence reading task.

An additional aim of this study was to analyse the role of presentation mode in self-paced reading setups for observing longer reading times in ambiguous sentences, a question that has not yet been systematically investigated. Regarding our first goal—investigating whether cross-task adaptation effects from sentence reading tasks to Stroop tasks can be observed, that is, whether conflict monitoring and control processing is shared between semantic processing and a task that does not involve likewise linguistic representations (i.e., Stroop)—the results showed a consistent pattern. Specifically, in three experiments (Experiments 1–3, and a combined analyses), both the sentences and the Stroop trials demonstrated the to-be-expected conflict effect (i.e., longer reading times in ambiguous compared with unambiguous sentences, and slower responses in incongruent compared with congruent

Stroop trials). However, no cross-task conflict adjustments were observed from sentences to Stroop trials and also not from Stroop trials to sentence reading trials.

Regarding our second goal—investigating whether within-task conflict adaptation patterns can be observed on subsequent sentence trials within a word-by-word self-paced reading task—again, the results showed a very straightforward pattern. Specifically, across three experiments (Experiments 4–6, and a combined analyses), we observed clear effects of sentence ambiguity, with longer reading times for ambiguous compared with unambiguous sentences in Region 3, which follows the region where the conflict arises. This finding is in line with Blott et al. (2020), supporting the idea that semantically ambiguous sentences produce processing conflicts during reading. However, no within-task conflict adaptation effect was observed, suggesting that any processing conflict experienced during the semantic processing of a sentence did not facilitate the processing of a subsequent similar representational conflict. There are several possible reasons for this null result. For example, timing parameters might play a crucial role as Egner et al. (2010) demonstrated decreasing CSEs in a face-word Stroop task with increasing time interval between successive trials. Specifically, the within-task conflict adaptation effect was absent with ISIs longer than 4,000 ms and RSIs longer than 2,500 ms. Thus, the long Sentence-to-Sentence interval (~4,000 ms) and the processing of intervening words might result in a decay of control-related processing adjustments before the subsequent semantic conflict is encountered (but see Bratzke & Janczyk, 2021, for a CSE in a dual-task experiment with an even longer interval). Moreover, it is possible that CSEs within the linguistic domain are present but of relatively small size. In this respect, it is noteworthy that the analysis of the combined data set across Experiments 4 to 6 did not show any sign of a CSE either.

Alternatively, one might argue that conflict-related adjustments do not exist within the language system. Such an interpretation might, at first sight, be in contrast with findings using negated phrases (i.e., “not left” vs. “not right”), which show that processing one negated phrase facilitates the processing of the subsequent negated phrase (Dudschig & Kaup, 2018, 2020). Importantly, in these studies, a setup was used where participants also had to perform the according responses (e.g., “not left” demanding a right-hand response). As a result, the observed conflict adaptation effect might reflect a response-related phenomenon. That is, since participants performed two-alternate forced-choice responses to a limited number of four stimuli (i.e., now right/left, not right/left), the task implemented by Dudschig and Kaup (2018, 2020) is similar to standard interference tasks such as the Flanker task (Eriksen & Eriksen, 1974) and the Simon task (Simon & Rudell, 1967) and subject to the confounds induced by the S-R sequence (cf. Hommel et al., 2004). Thus, as already

pointed out in the introduction, it is conceivable that cognitive control influences as revealed by CSEs in such studies are mainly driven by low-level stimulus and response feature bindings and the associated memory retrieval processes (Hommel et al., 2004; Mayr et al., 2003). Since the present linguistic stimuli avoided confounds of the congruency sequence with the S-R sequence, however, the absence of a reliable conflict adaptation effect for present task conditions would be expected according to these views. Still, to our knowledge, no previous language studies investigated whether within-task conflict triggers adjustments in subsequent linguistic processing. Thus, we consider it important for future studies to further investigate conflict adaptation within purely linguistic tasks while at the same time employing shorter time intervals between subsequent conflict trials.

Together, these findings provide no support for the idea that processing conflict within a semantically ambiguous sentence will facilitate the processing of a subsequent Stroop trial and vice versa. It must be noted that our results contrast with those reported by Kan et al. (2013) and by Hsu and Novick (2016), who demonstrated a clear-cut cross-task adaptation effect from the reading task, using syntactically ambiguous sentences, to the Stroop task. Whereas their studies provide evidence for a domain-general cognitive control mechanism, ours did not despite very similar experimental setups and timing of stimulus presentations.

There are, in fact, several potential reasons why our results differ from those of Kan et al. (2013). First, it is possible that semantic conflicts are processed differently than syntactic conflicts (e.g., as evident from Nozari group’s results; Nozari & Dell, 2011). However, semantic conflict is—like syntactic conflict—strongly associated with the activation of competing meaning representations—a typical sign of conflict during reading. Indeed, fMRI studies (January et al., 2009; Novick et al., 2005) have shown similar neural activations in both linguistic (syntactic and semantic ambiguity) and non-linguistic tasks (e.g., Stroop task). Therefore, it is unclear from a theoretical perspective why syntactic but not semantic conflict should trigger domain-general conflict adjustments. In other words, it can be assumed that if domain-general conflict adjustments between the linguistic processing domain and the Stroop task exist in such experimental setups, in principle, similar adjustments should be observed in this study. Still, this issue requires further investigation as it is possible that syntactic conflicts are longer-lasting or have more enduring effects on processing than semantic conflicts have, as will be further discussed below. In addition, our study had a slightly larger number of linguistic trials (24 congruent and 24 incongruent trials compared with Kan et al.’s (2013) 21 congruent and 21 incongruent trials] and we tested a larger sample of participants ($N=68$) than Kan et al. did ($N=41$). Thus, this study

should have more power to detect a subtle CSE. Also, like in Kan et al.'s (2013) study, our control conditions showed clear conflict adaptation patterns within the Stroop task, suggesting that our experimental setup and our participants were indeed suited to reveal a CSE. Together, we find it difficult to identify a specific reason in terms of the experimental and methodological specificities that would explain the divergent results.

It should be mentioned, however, that mixed findings concerning cross-task CSEs have been reported in the literature. For instance, Thothathiri et al. (2018), using a thematic role assignment task (which is arguably influenced by both syntactic and semantic cues), showed that encountering conflict on a previous Stroop trial modulated sentence comprehension towards the correct interpretation if the sentence contained a conflict. However, Freund et al. (2016) refuted a fully domain-general control system. In fact, they found no evidence of cross-task adaptation from a linguistic task and a non-linguistic task. They eventually conclude that their results support some specificity in the process of control regulation. Finally, a recent study (Aczel et al., 2021), which aimed at replicating Kan et al. (2013), did not observe an adaptation pattern from Sentence-to-Stroop trials, despite an increased sample size. Thus, Aczel et al. (2021) question the idea of domain-general cognitive control—specifically with regard to the language system. This leads us to tentatively conclude that, given its theoretical relevance, the boundary conditions of cross-task conflict adaptation effects from and to the language processing require further investigation.

Notably, inconsistent conflict-adjustment effects on sentence reading times in the available cross-task conflict adaptation studies using verbal and non-verbal stimuli might depend on the timing of stimulus presentation. Specifically, in Kan et al.'s (2013) and our study, the time interval between critical stimuli in the language task and the subsequent Stroop task was similar (~2,800–2,900 ms). Whereas Kan et al. observed a cross-task CSE in the Stroop task following a syntactic conflict, we failed to observe such a cross-task CSE in Stroop performance when semantic conflicts preceded. However, conflict-related adjustments from the Stroop task to the reading task were generally absent in both the study of Kan et al. and in our study. Since the time intervals between critical Stroop trials and language trials were longer in both Kan et al.'s (4,250 vs.) and our study (~4,000–5,000 ms) than for the reverse sequential task order, it is conceivable that the timing of sequential stimulus presentations and the decay of conflict-related information play a critical role for the CSE to appear. Fitting this picture, Hsu and Novick (2016) and Hsu et al. (2021) observed cross-task CSEs from a Stroop task and a flanker task, respectively, to a sentence processing task using a visual word paradigm. However, in their study, instruction sentences were spoken simultaneously to the presentation of task-relevant visual

objects, the time interval between critical trials in the two tasks was shorter than in studies using word-by-word presentation of sentences. Moreover, assuming that the timing of subsequent trials is critical and that syntactic conflict hypothetically lasts longer than semantic conflict, one might even account for the absence of a conflict adaptation effect within the linguistic domain when semantic instead of syntactic conflicts are employed. Certainly, to shed light on these issues, future studies should systematically manipulate the type of representational conflict in the language domain (e.g., syntactic, semantic, pragmatic), the interference tasks used, as well as the time interval between subsequent conflict stimuli.

In addition to studying conflict processing within the verbal domain and between verbal and non-verbal domains, we also investigated the role of presentation mode within self-paced reading setups, a setup that is used widely in the language processing literature (e.g., Kan et al., 2013). Despite self-paced reading being a standard measurement in language processing literature, we are not aware of a systematic investigation of the influence of self-paced reading procedure on the reported reading times. The use of a self-paced-reading setup is motivated by evidence that readers and listeners incrementally process sentences, and words and phrases are immediately integrated when encountered. Thus, when reading, we do not wait until the end of a sentence or even the end of a single word before starting to interpret what appears in front of us (e.g., Altmann & Kamide, 1999; Marslen-Wilson, 1987; Tanenhaus et al., 1995). To analyse the role of presentation mode on reading time measurements, we implemented three widely used self-paced reading paradigms (e.g., Beck & Weber, 2020; Ditman et al., 2007; Felsler et al., 2003; Ferreira & Henderson, 1990; Gibson & Warren, 2004; Koornneef & Van Berkum, 2006; Payne & Federmeier, 2017.) to compare any systematic influences of presentation mode on sentence processing effects triggered by semantic ambiguity. Overall, all three presentation modes show that semantically ambiguous sentences result in longer reading times following the disambiguating word. Only in Experiment 3—using the non-cumulative central presentation mode—no such effects were observed, potentially suggesting that this mode is the least likely to indicate semantic ambiguity-induced slowing of reading. Overall, however, the present analysis of word-by-word reading times across the three experimental setups indicated that presentation mode is not a particularly critical variable concerning the measurement of processing time costs due to semantic conflicts.

In conclusion, we did not observe evidence of conflict adaptation mechanisms in cross-task experiments intermixing the sentences with Stroop trials (Experiments 1–3; Sentence-to-Stroop) or within the linguistic system (Experiment 4–6; Sentence-to-Sentence) despite clear within-task CSEs in the manual Stroop task. Our lack of

evidence for domain-general as well as domain-specific control influences on language processing should be interpreted with caution since the relative timing of when conflicting information is presented in conjunction with the type of conflict might be critical. It is therefore important that future studies manipulate the interval between subsequent incongruent trials as well as conflict type (e.g., syn-tactic, semantic, pragmatic) to reveal the time course of cognitive control influences, that is, how fast they dissipate under different experimental task conditions. To this end, CSEs in the language domain could be investigated using word-based rather than sentence-based linguistic conflicts or using the visual world paradigm in combination with eye-tracking measures as promoted in the studies of Hsu and colleagues (e.g., Hsu et al., 2021; Hsu & Novick, 2016). This line of research should have both theoretical and applied impact and would shed further light on how language is processed by our cognitive systems.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Nicoletta Simi  <https://orcid.org/0000-0002-0966-6971>

Supplementary material

The supplementary material is available at qjep.sagepub.com

Author contributions

N.S.: Conceptualisation, Writing of Introduction, Discussion and Methods Sections, Data Curation, Manuscript Preparation.

I.G.M.: Conceptualisation, Programming, Data Curation and Analysis, Writing of Methods Section, Manuscript Preparation.

M.J.: Conceptualisation, Manuscript Preparation, Supervision.

H.L.: Conceptualisation, Manuscript Preparation, Supervision.

C.D.: Conceptualisation, Writing of Introduction and Discussion Sections, Manuscript Preparation, Supervision.

Notes

1. The short recall phase (~2/3 min) was implemented for motivational reasons to give participants a reason to really process the presented sentences—as they were instructed a recall phase will follow the experiment—and to avoid strategies such as clicking through the sentences without reading them. No analysis regarding this recognition task was planned or conducted.
2. Analyses were also performed excluding these trials, but this did not change any of the critical results.
3. During the review process, a Bayesian alternative to the frequentist repeated-measures ANOVA was requested. This is reported for the critical Previous Congruency × Current Congruency interaction for the Sentence-to-Stroop sequences within Experiments 1 to 3 (Part 1), and the Sentence-to-Sentence sequences within Experiments 4 to 6 (Part 1). The reported Bayes factor results from a comparison of a

model with a main effect of Current Congruency to a model also containing the interaction between Previous Congruency and Current Congruency. This analysis was implemented using the R function `lmbf` from the `BayesFactor` package (Morey & Rouder, 2021). For an interpretation of the Bayes Factor value in terms of evidence for the null versus the alternative hypothesis, see Lee and Wagenmakers (2014).

4. An identical analysis performed on the unadjusted data points showed the same data pattern, $t(67)=3.81, p<.001, d_z=0.46, 95\% CI=[8, 25]ms$.
5. Note that the labels *congruent* and *incongruent* are somewhat misleading here as sentence congruency is only defined with the disambiguating word presented within region 3. Note also that an identical analysis on the unadjusted datapoints did not show a significant effect, $t(67)=1.91, p=.060$. We therefore consider this a random effect.
6. An identical analysis performed on the unadjusted data points showed the same data pattern, $t(61)=3.72, p<.001, d_z=0.47, 95\% CI=[10, 33]ms$.
7. An additional three participants were removed from the sequence analysis due to missing cell data following the exclusion of trials where the current trial or the previous trial contained an extreme value.
8. Like in Experiment 1, an identical analysis performed on the unadjusted data points showed the same data pattern, $t(63)=2.18, p=.033, d_z=0.27, 95\% CI=[0.48, 11]ms$.
9. Note that fewer trials enter calculation of condition means compared with the initial congruency analysis disregarding trial sequences. Moreover, additional exclusion procedures were applied for the CSE analysis, for example, trials following an erroneous (i.e., outlier) sentence reading time were excluded. Together, this might account for the deviation of significance regarding the main effect of congruency across the two analyses.
10. Like in Experiment 1, an identical analysis performed on the unadjusted data points showed the same data pattern, $t(63)=3.44, p<.001, d_z=0.41, 95\% CI=[4, 15]ms$.

References

- Aczel, B., Kovacs, M., Bognar, M., Palfi, B., Hartanto, A., Onie, S., Tiong, L., & Evans, T. (2021). Is there evidence for cross-domain congruency sequence effect? A replication of Kan et al. (2013). *Royal Society Open Science*, 8, Article 191353. <https://doi.org/10.1098/rsos.191353>
- Akçay, Ç., & Hazeltine, E. (2011). Domain-specific conflict adaptation without feature repetitions. *Psychonomic Bulletin & Review*, 18, 505–511. <https://doi.org/10.3758/s13423-011-0084-y>
- Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73, 247–264. [https://doi.org/10.1016/S0010-0277\(99\)00059-1](https://doi.org/10.1016/S0010-0277(99)00059-1)
- Beck, S. D., & Weber, A. (2020). Context and literality in idiom processing: Evidence from self-paced reading. *Journal of Psycholinguistic Research*, 49, 837–863. <https://doi.org/10.1007/s10936-020-09719-2>
- Blott, L. M., Rodd, J. M., Ferreira, F., & Warren, J. E. (2020). Recovery from misinterpretations during online sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 47(6), 968–997. <https://doi.org/10.1037/xlm0000936>
- Botvinick, M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive

- control. *Psychological Review*, *108*, 624–652. <https://doi.org/10.1037/0033-295x.108.3.624>
- Boy, F., Husain, M., & Sumner, P. (2010). Unconscious inhibition separates two forms of cognitive control. *Proceedings of the National Academy of Sciences*, *107*, 11134–11139. <https://doi.org/10.1073/pnas.1001925107>
- Braem, S., Abrahamse, E. L., Duthoo, W., & Notebaert, W. (2014). What determines the specificity of conflict adaptation? A review, critical analysis, and proposed synthesis. *Frontiers in Psychology*, *5*, Article 1134. <https://doi.org/10.3389/fpsyg.2014.01134>
- Bratzke, D., & Janczyk, M. (2021). Introspection about backward crosstalk in dual-task performance. *Psychological Research*, *85*, 605–617. <https://doi.org/10.1007/s00426-019-01282-3>
- Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition*, *2*, Article 16. <https://doi.org/10.5334/joc.72>
- Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, *280*, 747–749. <https://www.science.org/doi/10.1126/science.280.5364.747>
- Cohen, J. D., Servan-Schreiber, D., & McClelland, J. L. (1992). A parallel distributed processing approach to automaticity. *The American Journal of Psychology*, *105*(2), 239–269. <https://doi.org/10.2307/1423029>
- De Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioural experiments in a Web browser. *Behavior Research Methods*, *47*, 1–12. <https://doi.org/10.1371/journal.pone.0051382>
- Ditman, T., Holcomb, P. J., & Kuperberg, G. R. (2007). An investigation of concurrent ERP and self-paced reading methodologies. *Psychophysiology*, *44*(6), 927–935. <https://doi.org/10.1111/j.1469-8986.2007.00593.x>
- Dudschig, C. (2022). Language and non-linguistic cognition: Shared mechanisms and principles reflected in the N400. *Biological Psychology*, *169*, Article 108282. <https://doi.org/10.1016/j.biopsycho.2022.108282>
- Dudschig, C., & Jentsch, I. (2009). Speeding before and slowing after errors: Is it all just strategy? *Brain Research*, *1296*, 56–62. <https://doi.org/10.1016/j.brainres.2009.08.009>
- Dudschig, C., & Kaup, B. (2018). How does “not left” become “right”? Electrophysiological evidence for a dynamic conflict-bound negation processing account. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(5), 716–728. <https://doi.org/10.1037/xhp0000481>
- Dudschig, C., & Kaup, B. (2020). Negation as conflict: Conflict adaptation following negating vertical spatial words. *Brain and Language*, *210*, Article 104842. <https://doi.org/10.1016/j.bandl.2020.104842>
- Egner, T. (2008). Multiple conflict-driven control mechanisms in the human brain. *Trends in Cognitive Sciences*, *12*, 374–380. <https://doi.org/10.1016/j.tics.2008.07.001>
- Egner, T., Ely, S., & Grinband, J. (2010). Going, going, gone: Characterizing the time-course of congruency sequence effects. *Frontiers in Psychology*, *1*, Article 154. <https://doi.org/10.3389/fpsyg.2010.00154>
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon identification of a target letter in a non-search task. *Perception and Psychophysics*, *16*, 143–149. <https://doi.org/10.3758/bf03203267>
- Eriksen, C. W., & Hoffman, J. E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. *Perception & Psychophysics*, *12*(2), 201–204. <https://doi.org/10.3758/BF03212870>
- Fanelli, D., & Ioannidis, J. P. (2013). US studies may overestimate effect sizes in softer research. *Proceedings of the National Academy of Sciences*, *110*(37), 15031–15036. <https://doi.org/10.1073/pnas.1302997110>
- Ferreira, F., & Henderson, J. (1990). Use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(4), 555–568.
- Forster, S. E., & Cho, R. Y. (2014). Context specificity of post-error and post-conflict cognitive control adjustments. *PLoS ONE*, *9*, Article e90281. <https://doi.org/10.1371/journal.pone.0090281>
- Freitas, A. L., Bahar, M., Yang, S., & Banai, R. (2007). Contextual adjustments in cognitive control across tasks. *Psychological Science*, *18*, 1040–1043. <https://doi.org/10.1111/j.1467-9280.2007.02022.x>
- Freund, M., Gordon, B., & Nozari, N. (2016). Conflict-based regulation of control in language production. In A. Papafragou, D. Grodner, D. Mirman & J. C. Trueswell (Eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 1625–1630). Cognitive Science Society.
- Funes, M. J., Lupianez, J., & Humphrey, G. K. (2010). Analyzing the generality of conflict adaptation effects. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 147–161. <https://doi.org/10.1037/a0017598>
- Gibson, E., & Warren, T. (2004). Reading-time evidence for intermediate linguistic structure in long-distance dependencies. *Syntax*, *7*, 55–78. <https://doi.org/10.1111/j.1368-0005.2004.00065.x>
- Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General*, *121*, 480–506. <https://doi.org/10.1037/0096-3445.121.4.480>
- Gruber, O., & Goschke, T. (2004). Executive control emerging from dynamic interactions between brain systems mediating language, working memory and attentional processes. *Acta Psychologica*, *115*, 105–121.
- Hazeltine, E., Lightman, E., Schwarb, H., & Schumacher, E. H. (2011). The boundaries of sequential modulations: Evidence for set-level control. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1898–1914.
- Hommel, B., Proctor, R. W., & Vu, K. P. L. (2004). A feature-integration account of sequential effects in the Simon task. *Psychological Research*, *68*, 1–17.
- Hsu, N. S., Jaeggi, S. M., & Novick, J. M. (2017). A common neural hub resolves syntactic and non-syntactic conflict through cooperation with task-specific networks. *Brain and Language*, *166*, 63–77. <https://doi.org/10.1016/j.bandl.2016.12.006>

- Hsu, N. S., Kuchinsky, S. E., & Novick, J. M. (2021). Direct impact of cognitive control on sentence processing and comprehension. *Language, Cognition and Neuroscience*, 36(2), 211–239. <https://doi.org/10.1080/23273798.2020.1836379>
- Hsu, N. S., & Novick, J. M. (2016). Dynamic engagement of cognitive control modulates recovery from misinterpretation during real-time language processing. *Psychological Science*, 27(4), 572–582.
- Janczyk, M., & Leuthold, H. (2018). Effector system-specific sequential modulations of congruency effects. *Psychonomic Bulletin & Review*, 25, 1066–1072.
- January, D., Trueswell, J. C., & Thompson-Schill, S. L. (2009). Colocalization of Stroop and syntactic ambiguity resolution in Broca's area: Implications for the neural basis of sentence processing. *Journal of Cognitive Neuroscience*, 21, 2434–2444. <https://doi.org/10.1162/jocn.2008.21179>
- Jentsch, I., & Dudschig, C. (2009). Why do we slow down after an error? Mechanisms underlying the effects of posterror slowing. *Quarterly Journal of Experimental Psychology*, 62, 209–218. <https://doi.org/10.1080/17470210802240655>
- Kan, I. P., Teubner-Rhodes, S., Drummey, A. B., Nutile, L., Krupa, L., & Novick, J. M. (2013). To adapt or not to adapt: The question of domain-general cognitive control. *Cognition*, 129, 637–651. <https://doi.org/10.1016/j.cognition.2013.09.001>
- Kleiman, T., Hassin, R. R., & Trope, Y. (2014). The control-freak mind: Stereotypical biases are eliminated following conflict-activated cognitive control. *Journal of Experimental Psychology: General*, 143, 498–503. <https://doi.org/10.1037/a0033047>
- Koornneef, A. W., & Van Berkum, J. J. A. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *Journal of Memory and Language*, 54(4), 445–465. <https://doi.org/10.1016/j.jml.2005.12.003>
- Kopp, B., Mattler, U., & Rist, F. (1994). Selective attention and response competition in schizophrenic patients. *Psychiatry Research*, 53(2), 129–139. [https://doi.org/10.1016/0165-1781\(94\)90104-x](https://doi.org/10.1016/0165-1781(94)90104-x)
- Kunde, W., Augst, S., & Kleinsorge, T. (2012). Adaptation to (non) valent task disturbance. *Cognitive, Affective & Behavioural Neuroscience*, 12, 644–660. <https://doi.org/10.3758/s13415-012-0116-8>
- Kunde, W., & Wühr, P. (2006). Sequential modulations of correspondence effects across spatial dimensions and tasks. *Memory & Cognition*, 34, 356–367. <https://doi.org/10.3758/BF03193413>
- Lee, M. D., & Wagenmakers, E. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109, 163–203. <https://doi.org/10.1037/0033-2909.109.2.163>
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25, 71–102. [https://doi.org/10.1016/0010-0277\(87\)90005-9](https://doi.org/10.1016/0010-0277(87)90005-9)
- Mayr, U., & Awh, E. (2009). The elusive link between conflict and conflict adaptation. *Psychological Research*, 73(6), 794–802. <https://doi.org/10.1007/s00426-008-0191-1>
- Mayr, U., Awh, E., & Laurey, P. (2003). Conflict adaptation effects in the absence of executive control. *Nature Neuroscience*, 6(5), 450–452. <https://doi.org/10.1038/n1051>
- Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., & Kramer, A. F. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Cognitive Brain Research*, 12, 467–473. [https://doi.org/10.1016/S0926-6410\(01\)00076-3](https://doi.org/10.1016/S0926-6410(01)00076-3)
- Miller, J. (1991). The flanker compatibility effect as a function of visual angle, attentional focus, visual transients, and perceptual load: A search for boundary conditions. *Perception & Psychophysics*, 49(3), 270–288.
- Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Reason*, 4, 61–64.
- Morey, R. D., & Rouder, J. N. (2021). *Bayesfactor: Computation of Bayes factors for common designs* (R package version 0.9.12-4.3). <https://CRAN.R-project.org/package=BayesFactor>
- Musz, E., & Thompson-Schill, S. L. (2016). Tracking competition and cognitive control during language comprehension with multi-voxel pattern analysis. *Brain and Language*, 165, 21–32. <https://doi.org/10.1016/j.bandl.2016.11.002>
- Notebaert, W., & Verguts, T. (2008). Cognitive control acts locally. *Cognition*, 106, 1071–1080. <https://doi.org/10.1016/j.cognition.2007.04.011>
- Novick, J. M., Hussey, E. K., Teubner-Rhodes, S., Harbison, J. I., & Bunting, M. F. (2014). Clearing the garden-path: Improving sentence processing through cognitive control training. *Language and Cognitive Processes*, 29(2), 186–217. <https://doi.org/10.1080/01690965.2012.758297>
- Novick, J. M., Kan, I. P., Trueswell, J. C., & Thompson-Schill, S. L. (2009). A case for conflict across multiple domains: Memory and language impairments following damage to ventrolateral prefrontal cortex. *Cognitive Neuropsychology*, 26, 527–567. <https://doi.org/10.1080/02643290903519367>
- Nozari, N., & Dell, G. S. (2011). Selective attention and speech errors: Feature migration in time. In L. Carlson, C. Hölscher & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1370–1375). Cognitive Science Society.
- Nozari, N., & Novick, J. (2017). Monitoring and control in language production. *Current Directions in Psychological Science*, 26(5), 403–410. <https://doi.org/10.1177/0963721417702419>
- Papenmeier, F. (2018). *powerbydesign: Power estimates for ANOVA designs* (R package version 1.0.4). <https://CRAN.R-project.org/package=powerbydesign>
- Payne, B. R., & Federmeier, K. D. (2017). Pace yourself: Intraindividual variability in context use revealed by self-paced event-related brain potentials. *Journal of Cognitive Neuroscience*, 29(5), 837–854. https://doi.org/10.1162/jocn_a_01090
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>
- Rodd, J. M., Johnsruide, I. S., & Davis, M. H. (2010). The role of domain-general frontal systems in language comprehension:

- Evidence from dual-task interference and semantic ambiguity. *Brain and Language*, *115*, 182–188. <https://doi.org/10.1016/j.bandl.2010.07.005>
- Schneider, C., Bade, N., & Janczyk, M. (2020). Is immediate processing of presupposition triggers automatic or capacity-limited? A combination of the PRP approach with a self-paced reading task. *Journal of Psycholinguistic Research*, *49*, 247–273.
- Simon, J. R., & Rudell, A. P. (1967). Auditory S–R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, *51*, 300–304.
- Stürmer, B., Leuthold, H., Soetens, E., Schröter, H., & Sommer, W. (2002). Control over location-based response activation in the Simon task: Behavioural and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 1345–1363.
- Stürmer, B., Seiss, E., & Leuthold, H. (2005). Executive control in the Simon task: A dual-task examination of response priming and its suppression. *European Journal of Cognitive Psychology*, *17*, 590–618.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634. <https://doi.org/10.1126/science.7777863>
- Thompson-Schill, S. L., D’Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences*, *23*, 14792–14797. <https://doi.org/10.1073/pnas.94.26.14792>
- Thompson-Schill, S. L., D’Esposito, M., & Kan, I. P. (1999). Effects of repetition and competition on activity in left prefrontal cortex during word generation. *Neuron*, *23*, 513–522. [https://doi.org/10.1016/s0896-6273\(00\)80804-1](https://doi.org/10.1016/s0896-6273(00)80804-1)
- Thothathiri, M., Asaro, C. T., Hsu, N. S., & Novick, J. M. (2018). Who did what? A causal role for cognitive control in thematic role assignment during sentence comprehension. *Cognition*, *178*, 162–177. <https://doi.org/10.1016/j.cognition.2018.05.014>
- Thothathiri, M., Kimberg, D. Y., & Schwartz, M. F. (2012). The neural basis of reversible sentence comprehension: Evidence from voxel-based lesion symptom mapping in aphasia. *Journal of Cognitive Neuroscience*, *24*(1), 212–222. https://doi.org/10.1162/jocn_a_00118
- Verbruggen, F., Notebaert, W., Liefvooghe, B., & Vandierendonck, A. (2006). Stimulus- and response-conflict-induced cognitive control in the flanker task. *Psychonomic Bulletin & Review*, *13*, 328–333. <https://doi.org/10.3758/BF03193852>
- Vuong, L. C., & Martin, R. C. (2011). LIFG-based attentional control and the resolution of lexical ambiguities in sentence context. *Brain and Language*, *116*(1), 22–32. <https://doi.org/10.1016/j.bandl.2010.09.012>
- Witzel, J., Witzel, N., & Nicol, J. (2012). Deeper than shallow: Evidence for structure-based parsing biases in second-language sentence processing. *Applied Psycholinguistics*, *33*, 419–456. <https://doi.org/10.1017/S0142716411000427>
- Yantis, S., & Johnston, J. C. (1990). On the locus of visual selection: Evidence from focused attention tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1), 135–149.