

**No pupillometric evidence for effortful proactive control in the proportion-congruent Stroop paradigm**

Journal:	<i>Quarterly Journal of Experimental Psychology</i>
Manuscript ID	QJE-SIP-23-149.R2
Manuscript Type:	Special Issue Paper
Date Submitted by the Author:	06-Dec-2023
Complete List of Authors:	Hasshim, Nabil; University of Salford Carruthers, Molly; De Montfort University Leicester FERRAND, Ludovic; CNRS, LAPSCO Augustinova, Maria ; Université de Rouen UFR Lettres et Sciences Humaines, Parris, Benjamin; Bournemouth University, Psychology
Keywords:	Cognitive control, Proportion congruency, Stroop task

SCHOLARONE™  
Manuscripts

**No pupillometric evidence for effortful proactive control in the proportion-congruent Stroop paradigm**

Nabil Hasshim<sup>1,2</sup>, Molly Carruthers<sup>2</sup>, Ludovic Ferrand<sup>3</sup>, Maria Augustinova<sup>4</sup> &

Benjamin A. Parris<sup>5</sup>

<sup>1</sup> University of Salford, UK

<sup>2</sup> De Montfort University, UK

<sup>3</sup> Université Clermont Auvergne, CNRS, LAPSCO, Clermont-Ferrand, France

<sup>4</sup> Normandie Université, UNIROUEN, CRFDP, Rouen, France

<sup>5</sup> Bournemouth University, UK

Corresponding Author:

Nabil Hasshim

School of Health and Society, University of Salford

Email: nabil.hasshim@gmail.com

Authors' note

This research was supported by the Experimental Psychology Society via the EPS Small Grants Scheme. The writing of the report; and the process related to submission of this article for publication was additionally supported by Agence nationale de la recherche (ANR Grant ANR-19-CE28-0013) et Réseau d'Intérêt Normandie (RIN Tremplin Grant 19E00851).

The preregistered design and analysis plans are available at the following OSF registrations page:

[https://osf.io/syzhx/?view\\_only=21613b8e058941758e43681337e7aa8c](https://osf.io/syzhx/?view_only=21613b8e058941758e43681337e7aa8c). The raw data (.edf and .csv formats), analysis files (JASP file format) are available at the

following OSF project page:

[https://osf.io/bweg8/?view\\_only=36ab2ae24ad949d7ab09b26992fd9c77](https://osf.io/bweg8/?view_only=36ab2ae24ad949d7ab09b26992fd9c77).

Peer Review Version

## Abstract

Cognitive control is the ability to allocate attention away from stimuli that are irrelevant to achieving a goal, towards stimuli that are. When conflict is anticipated, attention is biased in a global, top-down manner called proactive control and this effortful type of cognitive control is engaged before stimulus onset. The list-wise congruency proportion (LWPC) effect, where the Stroop congruency effect is reduced when there are more incongruent than congruent trials compared to vice versa, has been viewed as one of the prime signatures of this type of cognitive control. However, there has been recent debate about the extent to which this effect should be attributed to proactive control instead of alternative explanations such as simpler associative learning or reactive control. Thus, by using pupillometry (i.e., an indicator of cognitive effort), the present study investigated the extent to which LWPC effects result from effortful proactive control. Experiment 1 employed a classic proportion congruency manipulation while Experiment 2 replaced congruent trials with neutral trials to control for potential effects of associative learning. While in line with past findings, proportion congruency effects were obtained in response times of both experiments and pupillometry showed both proportion congruency and Stroop effects after stimulus onset, no differences in pupil sizes were found during the preparatory phase. Therefore, these results do not support the idea that the observed LWPC effects are due to participants engaging in effortful proactive control.

**Keywords:** *Cognitive control, Proportion congruency, Stroop task*

### No pupillometric evidence for effortful proactive control in the proportion-congruent paradigm

Cognitive control is the ability to orient our thoughts and actions towards an internal goal (Miller & Cohen, 2001). One of the most popular laboratory task that experimental psychologists use to study this ability is the Stroop task (Stroop, 1935; see MacLeod, 1991; Parris et al., 2021 for reviews) in which the goal is to identify the colour in which different colour-words are printed. Because for *incongruent* stimuli (e.g., the word 'blue' printed in green), attention needs to be shifted away from reading the word and applied to identifying the colour, poorer performance (e.g., slower responses, and more errors) is typically observed for these items as compared to control stimuli (e.g., *congruent* stimuli such as 'blue' printed in blue).

The efficiency of cognitive control is inferred from differences in magnitudes of this latter difference – referred to as the Stroop congruency effect such that smaller magnitudes are thought to reflect more efficient cognitive control (Braem et al., 2019). These differences can be observed both between participants (i.e., as a function of age for instance, see e.g., Bugg et al., 2007; Burca et al., 2022) and within participants. Indeed, magnitudes of the Stroop congruency effect are known to be substantially smaller when participants go through a block of trials which are mostly incongruent (e.g., 20% congruent and 80% incongruent) than when they go through a mostly congruent block (e.g., 80% congruent and 20% incongruent; see e.g., Lindsay & Jacoby, 1994; Logan & Zbrodoff, 1979).

This so-called list-wise congruency proportion effect (LWPC) along with other proportion congruency (PC) effects (see e.g., Bugg et al., 2008; Crump et al., 2006) suggests that cognitive control can be triggered in a top-down manner. Therefore, studies of these effects were crucial for subsequent theoretical development of cognitive control (e.g., Botvinick et al., 2001; Braver, et al., 2007; Norman & Shallice,

1  
2  
3 1986; Posner, 2012). Recently however, there is a debate about the extent to which  
4 these effects can actually be accounted for in a simpler, perhaps a more  
5 parsimonious way (see below). Therefore, by shedding some additional light on list-  
6 wise congruency proportion effect (LWPC), the present study was aimed to  
7 contribute to this ongoing debate.  
8  
9  
10  
11  
12  
13

### 14 *Alternative accounts to proactive control*

15  
16 Prominent theories of cognitive control often view LWPC (and other PC) effects as  
17 prime signatures of *proactive control* (e.g., Botvinick et al., 2001; Braver, et al., 2007;  
18 Norman & Shallice, 1986; Posner, 2012). This type of attentional control – triggered  
19 by participants' expectancies and motivations in response to different task  
20 environments – biases attention system in a global, top-down manner. For example,  
21 in an environment where conflicting information (e.g., an incongruent Stroop trial) is  
22 frequently encountered (in mostly incongruent or MI block), the conflict monitoring  
23 system (Botvinick et al., 2001) will signal for a shift towards greater proactive control  
24 (and away from reactive control) where more attentional resources to be utilised so  
25 as to aid task performance. The cognitive system will then react by engaging greater  
26 global cognitive control to bias attention away from the word, and towards the colour  
27 instead (Cheesman & Merikle, 1986; Lindsay & Jacoby, 1994; Lowe & Mitterer,  
28 1982; West & Baylis, 1998). Since the influence of the word is consequently  
29 lessened, smaller magnitudes of the Stroop congruency effect are observed than in  
30 situations where conflict is not frequently encountered (i.e., in a mostly congruent or  
31 MC block) and the conflict monitoring system is not triggered.  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53

54 However, there are also control accounts that do not assume a proactive  
55 component but posits the application of control can be stimulus driven instead. Much  
56 of the work comes from the demonstration of context-specific (e.g., fonts or  
57  
58  
59  
60

1  
2  
3 locations, Bugg et al., 2008) and item-specific proportion congruency effects (see  
4  
5 Bugg & Crump, 2012 for a review of PC effects). Although cognitive control is still  
6  
7 assumed, these accounts state that it is applied at the trial level, and not sustained at  
8  
9 the global level.  
10

11  
12 Since the frequencies of colour-word pairings in typical PC manipulations are  
13  
14 not equal (e.g., in MC blocks, the colour-word appears in the corresponding colour  
15  
16 most often), proponents of alternatives to cognitive control argue that this may  
17  
18 explain LWPC and other PC effects instead. For example, Schmidt and Besner  
19  
20 (2008) proposed a contingency learning explanation, driven by the fact that the  
21  
22 frequencies of the colour-word pairs making up the Stroop stimuli are confounded in  
23  
24 several PC paradigms (see also Schmidt, 2013, 2019). Because each colour-word is  
25  
26 more frequently paired with its corresponding colour (e.g., 'blue' in blue) than the  
27  
28 other colours (e.g., 'blue' in red, yellow, etc.), MC blocks lead to stronger (explicitly or  
29  
30 implicitly) learned stimulus-response contingencies and aiding participants' ability to  
31  
32 predict responses to more frequent word/colour pairs.  
33  
34  
35  
36

37  
38 The idea of how the imbalance of the word-colour pairings makes the word  
39  
40 dimension informative is not a new one as it had previously been proposed by  
41  
42 Dishon-Berkovits and Algom (2000). Unlike Schmidt and Besner's account that  
43  
44 works via pure associative learning, Dishon-Berkovits and Algom's model works via  
45  
46 adaptation. Since the colour-word correlation results in the word dimension of the  
47  
48 Stroop stimuli being informative of which potential response is more likely, this  
49  
50 encourages word reading, which is the process that we assume participants are  
51  
52 trying to suppress. In line with this reasoning, pairing frequencies of the stimuli has  
53  
54 been shown to reduce (and in some cases eliminate) PC effects, and Stroop effects  
55  
56 in general (see Algom & Chajut, 2019 and Schmidt, 2013; 2019 for reviews of the  
57  
58  
59  
60

1  
2  
3 topic, Spinelli & Lupker, in press, and Hasshim & Parris, 2021) for a direct  
4  
5 comparison of the two accounts).  
6

7  
8 In short, associative learning can explain PC effects such as LWPC without  
9  
10 the need to evoke control accounts. Therefore, there has been growing  
11  
12 acknowledgement that control processes are not always engaged or even necessary  
13  
14 (see Algom & Chajut, 2019; Algom et al., 2022; Schmidt, 2013; 2019). In an attempt  
15  
16 to reconcile these different explanations, Bugg (2014) tested the idea that top-down  
17  
18 proactive control is a process that the cognitive system can engage, depending on  
19  
20 the environmental context. This “associations as antagonists to top-down control”  
21  
22 (also called “last-resort”) account posits that proactive control does occur, but only in  
23  
24 certain situations. Specifically, it was not observed when responses can be predicted  
25  
26 via learned associations, but the cognitive system reverts to top-down control in  
27  
28 situations when such reliable stimulus-response cues are unavailable.  
29  
30  
31

32  
33 Pertinent to the current research, Spinelli and Lupker (2021) modified the  
34  
35 proportion congruency manipulation in LWPC by replacing congruent trials with  
36  
37 neutral ones to isolate the effects of conflict within the task, as without congruent  
38  
39 trials the word dimension never predicts the response. By demonstrating that in such  
40  
41 an environment, global level control is still engaged, they provided further evidence  
42  
43 that not only does proactive control exist, but it can be invoked even in case of  
44  
45 unpredictable word-response relationships (but see Bugg, 2014 above). In their more  
46  
47 recent work, Spinelli and Lupker (2022) provided additional evidence for the  
48  
49 existence of cognitive control by demonstrating its engagement even in situations  
50  
51 where contingency learning was encouraged.  
52  
53  
54

55  
56 To sum up, the aforementioned lines of work are asking the question of  
57  
58 whether proactive cognitive control exists and if so, when exactly it is triggered and  
59  
60



engaged. This question is still subject to much ongoing debate (see Abrahamse et al., 2016; Egnér, 2008; and Henik, et al., 2018; for reviews on the general topic of cognitive control) with one of the reasons for this being that the arguments tend to be relatively circulatory (i.e., going back and forth from constant, conditional and no proactive control). One possible reason for this is the fact that the vast majority of studies mentioned simply infer cognitive control and its different features from the modulation of the Stroop congruency effects (although see e.g., Blais & Bunge, 2010 for the use of fMRI and West & Alain, 2000 for the use of EEG to address these issues). Since proactive control is thought to be a resource-demanding, engagement of effortful proactive control can be measured more directly and more sensitively via pupillometry. This was precisely the goal of the current study.

### *Current study*

Pupillometry has previously been used as a measure of cognitive effort (Kahneman & Beattie, 1966; Laeng et al., 2012) in selective attention tasks such as Stroop paradigms (e.g., Hasshim & Parris, 2015; Hershman & Henik, 2019; Laeng et al., 2011) where more effort is typically exerted in trials that have more conflict (e.g., incongruent trials compared to congruent) leading to larger pupil sizes.

The current study is specifically looking into the effortful and proactive aspect of cognitive control, as described in the previous section, and past research have used pupillometry to achieve this goal. For example, Parris et al. (2021) showed that the attenuation of Stroop interference when participants were under post-hypnotic suggestion corresponded with larger pupil sizes, which was interpreted as the improvement in performance being due to participants engaging in more effortful control, a form of demand characteristic. While Parris et al. measured effort throughout the task, Chiew and Braver (2013) had a similar interest to the current

1  
2  
3 study, in proactive control prior to target onset. They found pupil sizes within a pre-  
4  
5 trial time window were larger on trials in which participants were expecting a reward  
6  
7 for performing – demonstrating that pupillometry measures just before the onset of a  
8  
9 target reflect top-down activation of effort.  
10

11  
12 Another example of the use of pupillometry in PC Stroop designs is Diede and  
13  
14 Bugg (2017) who used pupillometry to specifically demonstrate context specific  
15  
16 proportion congruency manipulation (CSPC), where congruent and incongruent trials  
17  
18 each reliably appear in specific contexts. In this study using a Flanker task, targets  
19  
20 appeared in locations where trials were MI or MC and indeed, larger pupil sizes were  
21  
22 observed in the MI location<sup>1</sup> where more cognitive effort was expected. Somewhat  
23  
24 similarly, the current study used pupillometry to measure effortful control in LWPC  
25  
26 paradigms.  
27  
28  
29

30  
31 However, it is important to understand that the goal of the present study is not  
32  
33 to simply extend results of Diede and Bugg (2017) to another PC paradigm (applied  
34  
35 in the Stroop task). Indeed, CSPC manipulation used in their study provides  
36  
37 evidence for control being exerted at the item level instead of (or in addition to)  
38  
39 globally throughout the block as described in the proactive control accounts of LWPC  
40  
41 effect. Since the application of proactive control prior to target onset cannot account  
42  
43 for these latter effects, attentional processes are likely to be activated after the target  
44  
45 is presented (i.e., after stimulus onset). The proactive account of LWPC effect on the  
46  
47 other hand clearly anticipates effortful control to be engaged *before* the target is  
48  
49 present (i.e., before stimulus onset, e.g., Botvinick et al., 2001; Braver et al., 2007;  
50  
51 Kane & Engle, 2003), especially when conflict is anticipated (Braver et al., 2007;  
52  
53  
54  
55  
56  
57  
58

---

59  
60 <sup>1</sup> Trials in MI locations showed larger peak pupil sizes compared to MC. Average peak pupil sizes occurred ~800ms to ~1000ms after average response times.

Gonthier et al., 2016; Spinelli & Lupker, 2021). Therefore, the current study investigated whether and the extent to which the expected smaller Stroop congruency effects MI blocks (as compared to MC blocks) were indeed preceded by effortful proactive control engaged during the preparatory phase (i.e., before target onset). To this end, in Experiment 1, participants' pupil sizes before target onset were compared in MI as compared to MC block of the classic LWPC paradigm using congruent and incongruent trials (e.g., Lindsay & Jacoby, 1994). In Experiment 2, neutral trials replaced congruent trials as per Spinelli and Lupker (2021), so the influence of stimulus-response contingencies and the informativeness of the word – operating in Experiment 1 – were both removed.

## Experiment 1

### Method

#### *Ethical approval*

Ethical approval for this study was granted by the Faculty Research Ethics Committee, Faculty of Health & Life Sciences, De Montfort University; Reference:3427.

#### *Participants*

44 individuals recruited from the university community participated in the experiment (data from 7 participants were excluded as they did not meet the pre-defined minimum accuracy of 90%) and received either course credit for their undergraduate course or £10 shopping vouchers. The intention was to recruit a minimum of 44 participants which was the suggested sample size calculated through the jpower module of jamovi software (The jamovi project, 2020). The parameters of a minimally

1  
2  
3 interesting effect size ( $\delta$ ) of 0.5, power of 0.9, and type 1 error rate of 0.05 were  
4  
5 based on the sizes of the effects of interest in previous laboratory Stroop research  
6  
7 that utilised behavioural responses. This target sample size is larger than the sample  
8  
9 sizes used in Blais and Bunge (2010) with 16 participants, and West and Alain  
10  
11 (2000) with 18 participants, studies of similar designs that informed the current  
12  
13 research. As mentioned above, both studies investigated similar research questions  
14  
15 to the current experiments, but with the use of EEG and fMRI methodology.  
16  
17  
18

19 While the sample size was estimated from behavioural measures of the PC  
20  
21 effect, it is also larger than those reported in the earlier mentioned research that  
22  
23 measured cognitive effort using pupillometry. Parris et al. (2021) reported a  
24  
25 difference in pupil sizes averaged throughout trials with 16 participants, while Chiew  
26  
27 and Braver (2013) showed pre-trial pupillary differences with 33 participants. If we  
28  
29 were to assume that effortful proactive control was to be present in the current  
30  
31 design and would have a similarly sized effect as these two studies, a Bayesian  
32  
33 power analysis (estimating a minimum sample size required, if we expect to observe  
34  
35 an effect of the same size as in the literature) suggests that the sample size required  
36  
37 to show a  $BF > 3$  to be 6 participants and for a  $BF > 6$ , to be 15 participants.  
38  
39  
40  
41  
42

### 43 *Apparatus*

44  
45 Stimuli were presented to the participants using a standard PC (screen dimensions:  
46  
47 530mm x 300mm) running Experiment Builder software (SR Research Ltd.), which  
48  
49 was displayed on a colour monitor displaying at 1920px x 1080px with a refresh rate  
50  
51 of 60Hz. Pupillary data consisting of pupil and corneal reflection, as well as sampling  
52  
53 at 1000 Hz was recorded using the EyeLink 1000 Plus eye-tracker (SR Research  
54  
55 Ltd.) running in monocular mode. Data was processed offline using Data Viewer (SR  
56  
57 Research Ltd.) Blinks were automatically identified by the EyeLink Online Parser  
58  
59  
60

used by the EyeLink 1000 eye-tracker (SR Research). Data during blinks, along with those within the 10 frames (10ms) of the onset and offset of a blink event were ignored as recommended by Mathôt and Vilotijević (2022).

### *Design*

The experiment was a 2 (block type: mostly congruent vs. mostly incongruent) x 2 (trial condition: congruent vs. incongruent) fully within-participant design. In the MC block, 80% of the trials were congruent trials, while the remaining 20% were incongruent trials, with these proportions reversed in the MI block. The stimuli consisted of two sets of colours (see Material section below) and participants encountered a different set in each block. The presentation order of the MC and MI blocks, and the colour set used in each block was randomly determined using a four-sided die. The presentation order of the trials within each block was randomised by the computer.

### *Material*

Two sets of colours and their corresponding words were used as targets for the experiment: yellow, green, and red in one set, and blue, pink, and white in the other. The visual angle of each of the words on the screen were as follows: yellow ( $2.08^\circ \times 0.78^\circ$ ), green ( $1.89^\circ \times 0.65^\circ$ ), red ( $1.04^\circ \times 0.59^\circ$ ); and blue ( $1.17^\circ \times 0.52^\circ$ ), pink ( $1.11^\circ \times 0.65^\circ$ ), white ( $1.95^\circ \times 0.65^\circ$ ). The RGB values of each colour were: yellow – (208, 255, 93), green (0, 255, 0), red (255, 0, 0), blue (0, 0, 255), pink (204, 0, 204), and white (255, 255, 255)<sup>2</sup>. The background of the screen was black (0, 0, 0) throughout the experiment.

---

<sup>2</sup> Luminosity of each stimulus was measured by converting each stimulus (word/letter-string and the background of the screen) into greyscale and aggregating the pixels as a proportion between 0 (completely black) and 1 (completely white). The mean luminosity of the congruent, incongruent, and neutral (used in experiment 2) stimuli were 0.000825, 0.000708, and 0.000677 respectively.

1  
2  
3 Participants encountering colours from different sets in the two blocks.  
4  
5 Congruent trials were made up of words spelling out the colour it appears in, while  
6  
7 incongruent trials were made up of a word appearing in one colour but spelling out a  
8  
9 different colour. Each word appears in its congruent colour on congruent trials and  
10  
11 can appear in either of the two other colours in the set. As mentioned above,  
12  
13 Experiment 1 was meant to follow the classic PC paradigm that does not control for  
14  
15 stimulus contingencies. Indeed, each word stimulus appear more (less) often in its  
16  
17 corresponding colour during the MC (MI) blocks. All visual stimuli were presented in  
18  
19 the centre of the screen.  
20  
21  
22  
23  
24

### 25 *Procedure*

26  
27 Participants performed a manual version of the Stroop task, where they were asked  
28  
29 to respond to the target by pressing keys on a full-sized QWERTY keyboard  
30  
31 corresponding to the colour. The keys used were the 1, 2, and 3, keys of the number  
32  
33 pad of a standard keyboard. During eye-tracking participants positioned their heads  
34  
35 on a headrest placed at an eye to screen distance of 880mm.  
36  
37  
38

39 For the experiment, participants underwent 330 trials. In the first half of the  
40  
41 experiment, participants underwent a block consisting of 15 practice trials (either MC  
42  
43 or MI) followed by the corresponding experimental block of 150 trials, with the stimuli  
44  
45 made up of one of the two aforementioned sets. The second half of the experiment  
46  
47 repeats this process with the second set of stimuli presented. A break was  
48  
49 administered in between the four blocks, and each block was preceded by a 5-point  
50  
51 calibration and validation.  
52  
53

54 Each trial began with a 500ms blank screen, followed by a grey fixation cross  
55  
56 presented for 800ms, before being replaced by the target. The target stayed in view  
57  
58 for 500ms followed by a blank screen until the participant responded up till a  
59  
60

1  
2  
3 maximum of 1000ms, when the trial would then be classified as a non-response and  
4 recycled as an upcoming trial. Another blank screen followed which was displayed  
5 for 1000ms. The time-windows for sampling pupil size are defined as follows:  
6  
7  
8 baseline pupil size was the average pupil size within 500ms before the onset of the  
9  
10 fixation, and proactive control was the average pupil size within 500ms before target  
11  
12 onset (see Figure 1). Since proactive control is a tonic response that is assumed to  
13  
14 be activated generally during the block, pupillometry readings were measured as the  
15  
16 average pupil size during the proactive time window.  
17  
18  
19  
20

21  
22 To reduce noise in the pupillary measure, recommendations from Mathôt et  
23  
24 al. (2018) that could be implemented, were reflected in the procedure. For example,  
25  
26 all stimuli were presented in the centre of the screen to minimise artefacts from pupil  
27  
28 foreshortening and to ensure that participants were looking at the centre of the  
29  
30 screen during the critical time windows. On trials where the total fixation time on the  
31  
32 fixation cross (as determined by an invisible circular interest area subtending  $2.1^\circ$ )  
33  
34 was  $<300$ ms, a drift correction was conducted and the trial restarted as this indicated  
35  
36 that either the participant was not focusing on the centre of the screen or accurate  
37  
38 calibration was not maintained. Recalibration and validation were conducted in more  
39  
40 severe cases. Furthermore, to reduce pupillary effects due to differences in exposure  
41  
42 time to the visual stimuli, the duration for which each stimulus was presented was  
43  
44 kept consistent for all trials.  
45  
46  
47  
48

49  
50 The details of the procedure were carefully considered to accommodate the  
51  
52 use of pupillometry in the PC Stroop paradigm. While a longer inter-trial-interval (ITI)  
53  
54 is recommended to allow for pupil dilation to go back to a baseline level after each  
55  
56 trial, long ITIs (e.g.,  $> 2000$ ms) have been shown to be detrimental to performance in  
57  
58 the Stroop task by reducing participants' ability to focus on the goal/task (De Jong et  
59  
60



1  
2  
3 al., 1999; see also e.g., Augustinova et al., 2018; Parris, 2014). This would arguably  
4  
5 affect maintaining proactive control since it would need to be effortfully sustained  
6  
7 over a longer period in between trials. Thus, an interval of 1500ms was chosen as a  
8  
9 compromise between these two concerns (also note that the pupillometry study of  
10  
11 Chiew and Braver, 2013 reported only minimal behavioural differences between ITIs  
12  
13 of 250 vs 4000ms). A baseline correction was also done to reduce carryover effects  
14  
15 from the previous trial, where pupillometry was taken as the percentage change  
16  
17 compared to the baseline period (Mathôt et al., 2018). To further elicit potential  
18  
19 proactive control, the fixation cross acts as an implicit cue to anticipate the  
20  
21 impending trial. The potential effects of these design choices are discussed with the  
22  
23 findings.<sup>3</sup>  
24  
25  
26  
27  
28  
29  
30

31 --- Insert Figure 1 about here ---  
32  
33  
34

35 The preregistered design and analysis plans are available at the following OSF  
36  
37 registrations page:

38 [https://osf.io/syzhx/?view\\_only=21613b8e058941758e43681337e7aa8c](https://osf.io/syzhx/?view_only=21613b8e058941758e43681337e7aa8c). The raw

39  
40 data (.edf and .csv formats), analysis files (JASP file format) are available at the  
41  
42 following OSF project page:  
43  
44

45 [https://osf.io/bweg8/?view\\_only=36ab2ae24ad949d7ab09b26992fd9c77](https://osf.io/bweg8/?view_only=36ab2ae24ad949d7ab09b26992fd9c77).  
46  
47  
48

## 49 Results

50  
51 Trials consisting of incorrect responses, or correct responses where responses were  
52  
53 not within 200-1500 ms of target onset, or with more than 30% missing pupil data  
54  
55

---

56  
57 <sup>3</sup> Although the 300ms period between the Baseline and Proactive periods was in place to minimise  
58 the effects of pupillary light reflex to the fixation cross, it is still a possibility. See the online  
59 Supplementary Material which shows the pattern for the pupillometry measures throughout the  
60 experiment that suggests that it is not due to pupillary light reflex.



1  
2  
3 were excluded from the analysis (only gaze data within the invisible central interest  
4 area defined above were considered 'valid'). Also, data from 7 participants who did  
5 not attain at least 90% accuracy was excluded from the analysis. Since the main  
6 analyses of interest are the comparison of pupil sizes between the MI and MC blocks  
7 with two directly competing predictions (i.e., proactive control accounts predict larger  
8 pupil sizes in MI, while other accounts do not), a Bayesian t-test was performed to  
9 complement the frequentist analysis. Bayes factors were calculated in JASP (JASP  
10 Team, 2022), using a half-Cauchy prior distribution (i.e., pupil size in MI>MC) scaled  
11 to 0.707 and robustness regions were reported to determine whether any  
12 interpretation from the Bayes factors were sensitive to the choice of prior distribution.  
13 The default prior specifications in JASP were also used for the Bayesian ANOVA  
14 analyses (i.e., r-scale fixed, random, and covariates values of 0.5, 1, and 0.354  
15 respectively). For replicability of the analyses reported in this manuscript, the random  
16 seeds were specified to be '999'. Table 1 summarises the RT and error data for both  
17 experiments.

18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38 --- Insert Table 1 about here ---  
39  
40  
41  
42

### 43 *Response times*

44 Response times were analysed to determine whether the pattern of results was in  
45 line with the expectations from research that use similar techniques, namely the  
46 classic LWPC effect.  
47  
48  
49  
50

51  
52 The 2 (block: MC vs MI) × 2 (trial condition: congruent vs incongruent)  
53 analysis of variance (ANOVA) of RT showed the block × trial condition interaction,  
54  $F(1, 36) = 34.94, p < .001, \eta_p^2 = 0.49$ , and main effect of trial condition,  $F(1, 36) =$   
55  $58.38, p < .001, \eta_p^2 = 0.62$ , to be statistically significant. This meant that while the  
56  
57  
58  
59  
60

1  
2  
3 Stroop congruency effect was statistically significant in both MI,  $t(36) = 3.85$ ,  $p <$   
4  
5  $.001$ ,  $d = 0.63$ , and MC,  $t(36) = 7.97$ ,  $p < .001$ ,  $d = 1.31$ , blocks, it was significantly  
6  
7 reduced in the former; a result that is consistent with the classic LWPC effect (see  
8  
9 figure 2 for a visualisation of the RTs). The main effect of block was not statistically  
10  
11 significant,  $F(1, 36) = 0.551$ ,  $p = .463$ ,  $\eta_p^2 = 0.015$ .

12  
13  
14 The corresponding Bayesian ANOVA analysis showed that the model which  
15  
16 included block, trial condition, and the interaction between the two, to be the best  
17  
18 model with  $BF_{10} = 8.0e+9$  when compared to the null model. The analysis of effects  
19  
20 comparing this model to one without the interaction showed decisive evidence in  
21  
22 favour of including it with a  $BF_{inclusion}$  of 407.08.

23  
24  
25  
26 --- Insert Figure 2 about here ---  
27  
28  
29  
30  
31  
32

### 33 *Error rates*

34  
35 The  $2 \times 2$  ANOVA for error rates revealed a statistically significant block  $\times$  trial  
36  
37 condition interaction  $F(1, 36) = 8.74$ ,  $p = .005$ ,  $\eta_p^2 = 0.20$ , and main effect of trial  
38  
39 condition,  $F(1, 36) = 22.85$ ,  $p < .001$ ,  $\eta_p^2 = 0.39$ , while the main effect of block was  
40  
41 non-significant,  $F(1, 36) = 1.31$ ,  $p = .260$ ,  $\eta_p^2 = 0.035$ . This meant that although the  
42  
43 Stroop congruency effect was significant in the error rates for both MC,  $t(36) = 4.36$ ,  
44  
45  $p < .001$ ,  $d = 0.717$ , and MI,  $t(36) = 2.41$ ,  $p = .021$ ,  $d = 0.397$ , blocks, the effect in the  
46  
47 former was larger. The results from the error rates mirrored that of the RT analysis.  
48  
49  
50  
51

### 52 *Pre-target Pupillometry*

53  
54 The comparison of preparatory pupil diameter between the two blocks is the primary  
55  
56 analysis of interest of this research. The results showed that the difference in dilation  
57  
58 between MI (0.44%) and MC (0.66%) blocks were statistically non-significant,  $t(36) =$   
59  
60

-0.690,  $p = .495$ ,  $d = -0.113$ ,  $BF_{01} = 8.91$ . The Bayes factor showed evidence supporting the null hypothesis and robustness check showed that for the  $BF$  to fall below 3, an interpretation of only anecdotal support for the null, the prior distribution would have to be scaled to be narrower than 0.193<sup>4</sup>.

### *Post-response Pupillometry (exploratory analysis)*

Although congruency effects in pupillometry are well established, predictions for reactive control accounts on LWPC effects have been less researched. Thus, post-response pupil sizes<sup>5</sup> were examined as exploratory analyses. The average pupil size during the 1000ms blank screen at the end of the trial (see “Post-response” in Figure 1) were sampled for each of the four conditions and compared, with the same baseline as the one used in the pre-response measures. This post-response time-window should be sensitive to congruency effects, following the time-series findings of Hershman and Henik (2019) who indicated pupil effects begin to emerge around 500ms (between neutral and incongruent trials) and 1000ms (between congruent and incongruent trials) post stimulus onset.

---

<sup>4</sup> It is possible that proactive control could be affecting pupil sizes during the baseline period as well if it is engaged throughout a block and thus occluding any effect. Although correction is recommended to reduce noise and improve statistical power, this is a valid concern. To check this, we compared the raw pupil sizes (pixels recorded by the eye-tracker) during the baseline period which showed no statistical difference between the blocks – Experiment 1:  $t(36) = 0.814$ ,  $p = .421$   $d = 0.134$   $BF_{01} = 4.16$ ; Experiment 2:  $t(35) = 0.342$ ,  $p = .734$   $d = 0.057$ ,  $BF_{01} = 5.29$ . This suggests that the pupil sizes during baseline were not different across blocks.

<sup>5</sup> The original pre-registration of this analysis indicated that the time-window for the exploratory post-target analysis starts immediately at target onset (i.e., while the target is still visible). However, while conducting the study, we concluded that this may not be a valid comparison since the pupillometry readings could be affected by the different luminosity of the stimuli on the screen. The time-window of data analysed and reported in the main manuscript excludes the period where any stimuli was visible. For transparency, analyses of pupil size in the time window between target onset and response showed the main effects and interactions to be non-significant for both experiments (all  $ps > .271$ ). Bayesian t-tests comparing each of the different conditions showed support for the null on all comparisons ( $BF_{01}s > 3$ ) apart from the difference between the neutral and incongruent trials in the MN block of Experiment 2, where  $BF_{01} = 2.97$ . This suggests that the different stimuli (colour and words) used in the different conditions did not differently affect pupil sizes. Observing pupillary Stroop effects only after response is also consistent with findings in the literature (e.g., Laeng et al., 2012; Hershman et al., 2023) showing the effects to only emerge after response.

1  
2  
3 As this was not the main analysis of interest and the design of the experiment  
4 was geared towards the pre-target analysis, it should be noted that some conditions  
5 (e.g., congruent trials in a Mostly Incongruent block) have a maximum of 30 trials  
6 presented to participants which meant that some conditions had very few valid  
7 observations. Thus, in addition to the trial and participant exclusion criteria that were  
8 previously specified, participants who did not have at least 10 valid trials for each  
9 condition (a further 7 participants) were also excluded from the analyses<sup>6</sup>.

10  
11  
12 The 2 (block: MC vs MI) × 2 (trial condition: congruent vs incongruent)  
13 ANOVAs of the mean pupil size showed the main effect of block to be statistically  
14 significant,  $F(1, 29) = 4.66$ ,  $p = .039$ ,  $\eta_p^2 = 0.14$ , as was the main effect of trial  
15 condition,  $F(1, 29) = 9.61$ ,  $p = .004$ ,  $\eta_p^2 = 0.25$ . This suggests larger post-response  
16 pupil dilation occurred during the MC block, and on incongruent trials. The block ×  
17 trial condition interaction was non-significant,  $F(1, 29) = 0.834$ ,  $p = .369$ ,  $\eta_p^2 = 0.028$ .  
18 The Bayesian analysis indicated the model containing the two factors being the best  
19 model, with a  $BF_{10}$  of 12.05 when compared to the null model. Compared to the null  
20 model, models including only the block; only trial condition; and both factors and their  
21 interaction returned  $BFs_{10}$  of 4.69, 2.13, and 4.12 respectively.

22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42 --- Insert Figure 3 about here ---  
43  
44  
45  
46

### 47 *Time-course Pupillary Changes*

48  
49 More recent research using pupillometry has moved to favour temporal analysis of  
50 pupil size changes (e.g., Hershman & Henik, 2019; Hershman et al., 2023) as it has  
51 several advantages, such as revealing when effects differ in their onset instead of  
52  
53  
54  
55  
56  
57

58  
59  
60  

---

<sup>6</sup> Analyses that included these participants showed similar patterns, although only the congruency effect in Experiment 2 was statistically significant.

1  
2  
3 magnitude, or those that only appear in a very small time-period compared to the  
4  
5 overall time window that is analysed. However, these are not the characteristics of  
6  
7 the proactive control effect that the current research endeavours to detect, but  
8  
9 instead the effortful proactive control is hypothesised to be sustained throughout the  
10  
11 preparatory period. Thus taking the average pupil sizes across the specified time-  
12  
13 windows were preferred instead.  
14  
15

16  
17 Nonetheless, it is possible that the temporal dynamics might reveal additional  
18  
19 insights beyond what was hypothesised. Thus, Figure 4 shows visualisations of pupil  
20  
21 size during the critical pre-target period, from the two experiments. The initial pupil  
22  
23 size is larger than baseline, which is likely attributed to a phasic response to the  
24  
25 fixation cross. The overall pattern does not suggest any hidden small effects within  
26  
27 the time-period. Although there is a visual difference between the two lines, note that  
28  
29 the analyses showed no significant differences between the blocks, and that the  
30  
31 pattern on neither graph is consistent with the hypothesis of generally larger pupil  
32  
33 sizes in the MI condition.  
34  
35

36  
37  
38 --- Insert Figure 4 about here ---  
39  
40  
41

## 42 Discussion

43  
44 The RT data from the experiment replicated the classic PC effect, where Stroop  
45  
46 interference was smaller in the MI block compared to the MC block. The post-  
47  
48 response pupillometry analysis showed a congruency effect of larger pupil dilation  
49  
50 after accurately responding to incongruent trials compared to congruent ones,  
51  
52 replicating typical pupillary Stroop effects. The main effect of block was also  
53  
54 significant with generally larger post-response pupil dilation during the MC block  
55  
56 which is the opposite of what would be expected if more effort was exerted in the MI  
57  
58  
59  
60

1  
2  
3 blocks. It is possible that post-response pupil sizes reflect the amount of conflict  
4  
5 encountered in the trial. That is, pupil Stroop effects are commonly observed  
6  
7 ~500ms-600ms after participants make a response (see e.g., Hershman & Henik,  
8  
9 2019; 2020; Laeng et al., 2011), indicating that pupils reflect a delayed reaction to  
10  
11 the conflict encountered when attempting to respond (Simpson, 1969), and thus  
12  
13 pupil sizes would be overall larger in the MC block in which the RT congruency effect  
14  
15 was larger. However, if pupil sizes were a delayed reflection of pre-response conflict  
16  
17 encountered, we might expect to see a main effect of block in the RT data, but this  
18  
19 was not observed (see also the lack of main effects in the pupil data in Experiment  
20  
21 2). Thus, unfortunately, the results from this exploratory analysis were not consistent  
22  
23 with either account.  
24  
25  
26  
27

28  
29 For the main analysis of interest, the pre-target pupillometry data showed  
30  
31 evidence for no difference in pupil dilation during the preparatory phase between MI  
32  
33 and MC blocks. This means that while the LWPC effect was observed in the  
34  
35 behavioural data, it was not accompanied by a pupillary marker of effortful attention  
36  
37 control in anticipation of conflict.  
38  
39

40  
41 While a pupillary effect would have suggested that proactive control is still  
42  
43 engaged in the classic LWPC paradigm despite the associative learning confounds  
44  
45 highlighted earlier, the results from Experiment 1 do not necessarily provide  
46  
47 evidence against the proactive control account. As proponents of proactive control  
48  
49 accounts have argued, it might be a process that only emerges when other more  
50  
51 effective strategies are unavailable (e.g., Bugg, 2014; Spinelli & Lupker, 2021).  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Experiment 2

Experiment 2 aimed at tackling one of the major associative learning confounds found in the classic LWPC paradigm, the effect of word-response contingency (Schmidt, 2013, 2019). This was done by modifying the design of the experiment in two main ways. Firstly, the congruent trials were used in Experiment 1 were replaced by neutral trials (e.g., string of 'ssss' in red). As suggested by Spinelli and Lupker (2021), the advantage of utilising neutral trials over congruent ones in this context is that the former reduces the informativeness of the word/letter-string and negates any advantage of gleaning information from the text. Additionally, the effect of word-response contingency was further controlled for by mapping each (non)word equally often to two of the response options (and never to the third). This meant that for each letter-string and word making up the neutral and incongruent trials, there was an equal probability of it appearing in one of two colours (and never the third).

## Method

### *Participants*

41 individuals recruited from the university community participated in the experiment (data from 5 participants were excluded as they did not meet the pre-defined minimum accuracy of 90%) and received either course credit for their undergraduate course or £10 shopping vouchers.

### *Design, Procedure, & Material*

The design and procedure were identical to that of Experiment 1. The material was similar except that that congruent trials were replaced by repeated-letter neutral trials. Thus, instead of an MC block, participants went through a mostly neutral (MN) block. The repeated-letter sequences were zzzzzz ( $2.28^\circ \times 0.46^\circ$  of visual angle),



xxx (1.17° x 0.39°), jjjjj (0.91° x 0.78°) for one half of the trials, and the other half involved the targets ssss (1.30° x 0.39°), qqqq (1.43° x 0.52°), and hhhhh (1.69° x 0.52°).

To control for contingency effects, each repeated-letter string only appeared in two of the three colours, and equally often in each (e.g., 'xxx' appeared in green and yellow equally often, but never in red; 'jjjjj' appeared equally often in yellow and red, but never green; etc.). This mirrors the properties of incongruent trials.

## Results

Data were processed in the same way as in Experiment 1, with data from 5 participants excluded for not meeting the 90% minimum accuracy threshold.

### *Response times*

The 2 (block: MN vs MI) × 2 (trial condition: neutral vs incongruent) ANOVA of RT showed the block × trial condition interaction,  $F(1, 35) = 25.09, p < .001, \eta_p^2 = 0.42$ , and main effect of trial condition,  $F(1, 35) = 44.15, p < .001, \eta_p^2 = 0.56$ , to be statistically significant. The main effect of block was not statistically significant,  $F(1, 35) = 0.314, p = .579, \eta_p^2 = 0.009$ . This meant that while the Stroop interference effect was statistically significant in both MI,  $t(35) = 2.49, p = .018, d = 0.42$ , and MN,  $t(35) = 6.76, p < .001, d = 1.13$ , blocks, it was significantly reduced in the former.

Similar to Experiment 1 this is consistent with the classic LWPC effect (see Figure 5 for descriptive statistics).

The Bayesian ANOVA analysis revealed the model which included block, trial condition, and the interaction between the two, to be the best model with  $BF_{10} = 99658.60$  when compared to the null model. The analysis of effects comparing this model to one without the interaction showed strong evidence in favour of including it with a  $BF_{inclusion}$  of 44.03.



--- Insert Figure 5 about here ---

### Error rates

Unlike the RT data, none of the effects in the error rate analysis were statistically significant, interaction:  $F(1, 35) = 3.33, p = .077, \eta_p^2 = 0.087$ , main effect of block:  $F(1, 35) = 0.070, p = .794, \eta_p^2 = 0.002$ , and trial condition:  $F(1, 35) = 0.732, p = .401, \eta_p^2 = 0.020$ . This could be due to lower error rates in the incongruent trials, compared the incongruent trials in Experiment 1.

### Pre-target Pupillometry

The comparison of preparatory pupil sizes showed that difference in dilation between MI (-0.01%) and MN (0.15%) were statistically non-significant,  $t(35) = -0.468, p = .639, d = -0.078, BF_{01} = 7.72$ . Similar to the results from Experiment 1, the Bayes factor showed evidence supporting the null hypothesis with a robustness check showing that for the  $BF$  to fall below 3, the prior distribution would have to be scaled to be narrower than 0.228.

### Post-response Pupillometry (exploratory analysis)

Following the same exclusion criteria as in Experiment 1, a further four participants were excluded from this analysis. The 2 (block: MC/MN vs MI)  $\times$  2 (trial condition: congruent vs incongruent) ANOVAs of the mean pupil size showed only the main effect of trial-condition was statistically significant,  $F(1, 31) = 8.12, p = .008, \eta_p^2 = 0.21$ , with incongruent trials displaying larger pupil dilations compared to neutral trials. Neither the main effect of block,  $F(1, 31) = 1.44, p = .240, \eta_p^2 = 0.044$ , nor the block  $\times$  trial-condition interaction,  $F(1, 31) = 1.24, p = .274, \eta_p^2 = 0.038$ , were statistically significant. Although the pattern of results is similar to that of Experiment

1  
2  
3 1 (see Figure 6), the main effect of block was non-significant (unlike in Experiment 1)  
4  
5 while the Bayesian analysis indicated the null model to be the best model with a  
6  
7 model that included only trial condition having a  $BF_{10} = 0.935$ . This suggests that this  
8  
9 exploratory analysis was underpowered and conclusions should not be drawn from  
10  
11 this specific analysis.  
12  
13

14  
15 --- Insert Figure 6 about here ---  
16  
17

## 18 19 Discussion

20  
21 The primary results of interest from Experiment 2 closely followed those of  
22  
23 Experiment 1, showing that although the LWPC effect was reflected in participants'  
24  
25 RT, where the Stroop effect was larger in the MI compared to the MN block  
26  
27 (replicating the results of Spinelli & Lupker, 2021), this was not accompanied by  
28  
29 larger pupil dilation in the former. The analysis of the pupillometry data showed  
30  
31 evidence for the null hypothesis. However, it should be noted that in the exploratory  
32  
33 post-response pupillometry analysis, although trial congruency effects were shown in  
34  
35 the frequentist analysis, the data was not sensitive enough to show similar evidence  
36  
37 in the Bayesian analysis.  
38  
39  
40  
41  
42  
43  
44

## 45 General discussion

46  
47 The current research investigated whether there is evidence of effortful control of  
48  
49 attention being proactively applied when conflict is expected in a selective attention  
50  
51 task. Following previous work utilising pupillometry to evidence proactive control  
52  
53 (Chiew & Braver, 2013; Parris et al., 2021), the intention was to demonstrate a  
54  
55 putative marker of proactive control in the time-frame just before the presentation of  
56  
57 the target. If the attention control mechanism is engaged effortfully in anticipation of  
58  
59  
60

1  
2  
3 conflict, larger pupil sizes were expected before the Stroop stimulus was presented  
4  
5 during the MI block, where most of the trials were conflicting – providing convincing  
6  
7 evidence for proactive control. We did not observe this effect in our data.  
8  
9

10 In both experiments, while RTs displayed the classic LWPC effect of a smaller  
11  
12 Stroop congruency effect in the MI block, no pupillary effects were observed prior to  
13  
14 target onset. Experiment 1 used the typical LWPC paradigm of comparing an MI  
15  
16 block to an MC block (Lindsay & Jacoby, 1994), while Experiment 2 aimed at  
17  
18 addressing the potential confound of contingency effects by comparing an MI block  
19  
20 to an MN block. The results in the RT data are consistent with Botvinick et al.  
21  
22 (2001)'s model of proactive control since replacing congruent trials with neutral trials  
23  
24 will not affect the PC effect because the conflict monitoring module works via  
25  
26 sensitivity to the frequency of conflict trials (i.e., incongruent). The pattern of results  
27  
28 is also consistent with Tzelgov et al. (1992) who found larger interference effects  
29  
30 with increased proportion of neutral trials. However, the pupillometry results failed to  
31  
32 detect an accompanying marker for increased effort when participants were  
33  
34 anticipating target onset in the MI blocks. Indeed, descriptive statistics showed that  
35  
36 the amount of pupil dilation in general to be very small and even in the opposite  
37  
38 direction (albeit non-significant) to that which proactive control accounts would  
39  
40 predict. Additionally, the Bayes factors for both experiments favoured evidence  
41  
42 against the hypothesis of larger pupil sizes in the MI (as compared to MC/MN) block.  
43  
44  
45  
46  
47  
48

49 Interestingly, it is notable that the exploratory analysis showed that the pattern  
50  
51 of results from the (post-response) pupil sizes data do not mirror that of participants  
52  
53 RTs (non-significant block  $\times$  trial condition interaction in pupil sizes) or conform to  
54  
55 the expectation of more effort required in the MI block. This suggests that the post-  
56  
57 response pupil size might be sensitive to block level effects in addition to trial-level  
58  
59  
60

ones, something that future research using an appropriately powered design should explore.

### *Limitations of the current research*

In isolation, interpretation of the findings from the two experiments seems straightforward: while the LWPC effect was observed in the RTs, the absence of pupillary differences in the preparatory phase suggests that the effect was not due to more proactive control processes being engaged when conflict was more likely. However, although demonstrating an effect would have provided evidence for proactive control, the converse does not necessarily mean evidence against it as even though pupillary effects have been used in similar investigations (e.g., Chiew & Braver, 2013; Parris et al, 2021) they might not be sensitive enough to index effortful proactive control in the current study. Relatedly, the lack of effects could also be due to a lack of power for the analysis. However, the sample sizes were larger than studies that demonstrated pupillometry effects comparable to the ones of interest, and the BFs reported do not suggest an issue with sensitivity. Even if the results were to be taken at face-value i.e., evidence against greater effort, the only conclusion that can be made is that the mechanism (control or otherwise) behind the LWPC effect does not involve sustained effort.

Another potential limitation of the present study is the use of only manual (button-press) responses. Larger Stroop effects are typically found in vocal, compared to manual, versions of the task (e.g., Augustinova et al., 2019; White, 1969) and can differently affect processes within Stroop task performance (e.g., Sharma & McKenna, 1998). However, behavioural LWPC effects were observed in both experiments as were the typical pupillary Stroop effects, suggesting that the paradigm did manage to capture the effects of interest. Although there is no strong

1  
2  
3 indication that different anticipatory control mechanisms should be involved when the  
4  
5 subsequent response is vocal vs. manual, response modality should be something  
6  
7 that future research should consider as well.  
8  
9

10         Also, and importantly, control accounts (such as Braver et al., 2007 and  
11  
12 Crump et al., 2006) do highlight control that is applied reactively during the task.  
13  
14 Somewhat in line with reactive control, the secondary, exploratory analyses of post-  
15  
16 response pupil sizes hint at the influence of proportion congruency on the effort  
17  
18 applied after the target is presented. However, the data from the current study, which  
19  
20 was primarily aimed at detecting the larger block-level effects, were not sensitive  
21  
22 enough to examine the interactive effects of trial condition and proportion  
23  
24 congruency on post-response pupil dilation. Given this important limitation, further  
25  
26 research is also required to tease apart these processes that can independently  
27  
28 affect pupil sizes. In doing so, these studies also need to integrate pupillometry to  
29  
30 other variations that target reactive control (e.g., context-specific PC, Crump et al.,  
31  
32 2006).  
33  
34  
35  
36  
37

### 38 *Conclusions*

39  
40 The current research intended to examine often assumed, but largely untested idea  
41  
42 that effortful, proactive control is engaged prior to the presentation of stimulus in the  
43  
44 LWPC Stroop paradigm. This was done using pupillometry as a direct measure of  
45  
46 cognitive effort during the preparatory phase of the task. It is important to  
47  
48 acknowledge that the rich literature within the wider debate about the role of control  
49  
50 in the PC paradigm has involved more elaborate designs, the current research is an  
51  
52 initial exploration utilising pupillometry, and any conclusions are limited to the classic  
53  
54 version of the task. Therefore, it is hoped that the results from the two experiments  
55  
56 demonstrate a simple point: the nature of the processes that results in the LWPC  
57  
58  
59  
60

1  
2  
3 effect, at least for those observed in the classic PC paradigm, was not shown to  
4  
5 involve *effortful* control engaged *in anticipation* of conflict.  
6  
7  
8  
9

10 Supplementary Material

11  
12 The Supplementary Material is available at: [qjep.sagepub.com](http://qjep.sagepub.com)  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Peer Review Version

## References

- Abrahamse, E., Braem, S., Notebaert, W., & Verguts, T. (2016). Grounding cognitive control in associative learning. *Psychological Bulletin*, *142*(7), 693 – 728.
- Algom, D., & Chajut, E. (2019). Reclaiming the Stroop effect back from control to input-driven attention and perception. *Frontiers in Psychology*, *10*, 1683.
- Algom, D., Fitousi, D., & Chajut, E. (2022). Can the Stroop effect serve as the gold standard of conflict monitoring and control? A conceptual critique. *Memory & Cognition*, *50*(5), 883-897.
- Augustinova, M., Parris, B. A., & Ferrand, L. (2019) The Loci of Stroop Interference and Facilitation. Effects with Manual and Vocal Responses. *Frontiers in Psychology*, *10*, 1786
- Augustinova, M., Silvert, L., Spatola, N., & Ferrand, L. (2018). Further investigation of distinct components of Stroop interference and of their reduction by short response-stimulus intervals. *Acta Psychologica*, *189*, 54-62.
- Blais, C., & Bunge, S. (2010). Behavioral and neural evidence for item-specific performance monitoring. *Journal of Cognitive Neuroscience*, *22*(12), 2758-2767.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624 - 652.
- Braem, S., Bugg, J. M., Schmidt, J. R., Crump, M. J., Weissman, D. H., Notebaert, W., & Egner, T. (2019). Measuring adaptive control in conflict tasks. *Trends in Cognitive Sciences*, *23*(9), 769-783.

1  
2  
3 Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of  
4  
5 working memory variation: Dual mechanisms of cognitive control. *Variation in*  
6  
7 *Working Memory*, 76 - 106.

8  
9  
10 Bugg, J. M. (2014). Conflict-triggered top-down control: Default mode, last resort, or  
11  
12 no such thing?. *Journal of Experimental Psychology: Learning, Memory, and*  
13  
14 *Cognition*, 40(2), 567 - 587.

15  
16  
17 Bugg, J. M., DeLosh, E. L., Davalos, D. B., & Davis, H. P. (2007). Age differences in  
18  
19 Stroop interference: Contributions of general slowing and task-specific  
20  
21 deficits. *Aging, Neuropsychology, and Cognition*, 14(2), 155-167.

22  
23  
24 Bugg, J. M., & Hutchison, K. A. (2013). Converging evidence for control of color-  
25  
26 word Stroop interference at the item level. *Journal of Experimental*  
27  
28 *Psychology: Human Perception and Performance*, 39(2), 433 - 449.

29  
30  
31 Bugg, J. M., Jacoby, L. L., & Toth, J. P. (2008). Multiple levels of control in the  
32  
33 Stroop task. *Memory & Cognition*, 36(8), 1484-1494.

34  
35  
36 Burca, M., Chausse, P., Ferrand, L., Parris, B. A., & Augustinova, M. (2022). Some  
37  
38 further clarifications on age-related differences in the Stroop task: New  
39  
40 evidence from the two-to-one Stroop paradigm. *Psychonomic Bulletin &*  
41  
42 *Review*, 29(2), 492-500.

43  
44  
45 Cheesman, J., Merikle, P. M. (1986). Distinguishing conscious from unconscious  
46  
47 perceptual processes. *Canadian Journal of Psychology / Revue Canadienne*  
48  
49 *de Psychologie*, 40(4), 343-367.

50  
51  
52 Chiew, K. S., & Braver, T. S. (2013). Temporal dynamics of motivation-cognitive  
53  
54 control interactions revealed by high-resolution pupillometry. *Frontiers in*  
55  
56 *Psychology*, 4, 15.



- 1  
2  
3 Crump, M. J., Gong, Z., & Milliken, B. (2006). The context-specific proportion  
4  
5 congruent Stroop effect: Location as a contextual cue. *Psychonomic Bulletin*  
6  
7 & *Review*, 13(2), 316-321.
- 8  
9  
10 De Jong, R., Berendsen, E., & Cools, R. (1999). Goal neglect and inhibitory  
11  
12 limitations: Dissociable causes of interference effects in conflict  
13  
14 situations. *Acta Psychologica*, 101(2-3), 379-394.
- 15  
16  
17 Diede, N. T., & Bugg, J. M. (2017). Cognitive effort is modulated outside of the  
18  
19 explicit awareness of conflict frequency: Evidence from pupillometry. *Journal*  
20  
21 *of Experimental Psychology: Learning, Memory, and Cognition*, 43(5), 824 -  
22  
23 835.
- 24  
25  
26 Dishon-Berkovits, M., & Algom, D. (2000). The Stroop effect: It is not the robust  
27  
28 phenomenon that you have thought it to be. *Memory & Cognition*, 28(8),  
29  
30 1437-1449.
- 31  
32  
33 Egner, T. (2008). Multiple conflict-driven control mechanisms in the human  
34  
35 brain. *Trends in cognitive sciences*, 12(10), 374-380.
- 36  
37  
38 Egner, T. (2014). Creatures of habit (and control): A multi-level learning perspective  
39  
40 on the modulation of congruency effects. *Frontiers in Psychology*, 5, 1247.
- 41  
42  
43 Egner, T. (2017). Past, present, and future of the congruency sequence effect as an  
44  
45 index of cognitive control. In *The Wiley handbook of cognitive control* (pp. 64-  
46  
47 78). John Wiley & Sons.
- 48  
49  
50 Gonthier, C., Braver, T. S., & Bugg, J. M. (2016). Dissociating proactive and reactive  
51  
52 control in the Stroop task. *Memory & Cognition*, 44(5), 778-788.
- 53  
54  
55 Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information:  
56  
57 strategic control of activation of responses. *Journal of Experimental*  
58  
59 *Psychology: General*, 121(4), 480 - 506.
- 60

- 1  
2  
3 Hasshim, N., & Parris, B. A. (2015). Assessing stimulus–stimulus (semantic) conflict  
4  
5 in the Stroop task using saccadic two-to-one color response mapping and  
6  
7 preresponse pupillary measures. *Attention, Perception, &*  
8  
9 *Psychophysics*, 77(8), 2601-2610.
- 10  
11  
12 Hasshim, N., & Parris, B. A. (2021). The role of contingency and correlation in the  
13  
14 Stroop task. *Quarterly Journal of Experimental Psychology*, 74(10), 1657-  
15  
16 1668.
- 17  
18  
19 Henik, A., Bugg, J. M., & Goldfarb, L. (2018). Inspired by the past and looking to the  
20  
21 future of the Stroop effect [Editorial]. *Acta Psychologica*, 189, 1–3.
- 22  
23  
24 Hershman, R., & Henik, A. (2019). Dissociation between reaction time and pupil  
25  
26 dilation in the Stroop task. *Journal of Experimental Psychology: Learning,*  
27  
28 *Memory, and Cognition*, 45(10), 1899 - 1909.
- 29  
30  
31 Hershman, R., Milshtein, D., & Henik, A. (2023). The contribution of temporal  
32  
33 analysis of pupillometry measurements to cognitive research. *Psychological*  
34  
35 *Research*, 87(1), 28-42.
- 36  
37  
38 Jacoby, L. L., Lindsay, D. S., & Hessels, S. (2003). Item-specific control of automatic  
39  
40 processes: Stroop process dissociations. *Psychonomic Bulletin &*  
41  
42 *Review*, 10(3), 638-644.
- 43  
44  
45 JASP Team (2022). JASP (Version 0.16.3)[Computer software]
- 46  
47  
48 Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on  
49  
50 memory. *Science*, 154(3756), 1583-1585.
- 51  
52  
53 Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of  
54  
55 attention: The contributions of goal neglect, response competition, and task  
56  
57 set to Stroop interference. *Journal of Experimental Psychology: General*,  
58  
59 132(1), 47–70.
- 60

- 1  
2  
3 Laeng, B., Ørbo, M., Holmlund, T., & Miozzo, M. (2011). Pupillary stroop  
4  
5 effects. *Cognitive Processing*, 12(1), 13-21.  
6  
7  
8 Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the  
9  
10 preconscious?. *Perspectives on Psychological Science*, 7(1), 18-27.  
11  
12  
13 Lindsay, D. S., & Jacoby, L. L. (1994). Stroop process dissociations: the relationship  
14  
15 between facilitation and interference. *Journal of Experimental Psychology:*  
16  
17 *Human Perception and Performance*, 20(2), 219 - 234.  
18  
19  
20 Logan, G. D., & Zbrodoff, N. J. (1979). When it helps to be misled: Facilitative effects  
21  
22 of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory*  
23  
24 & *cognition*, 7(3), 166-174.  
25  
26  
27 Lowe, D. G., & Mitterer, J. O. (1982). Selective and divided attention in a Stroop  
28  
29 task. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*,  
30  
31 36(4), 684–700.  
32  
33  
34 MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an  
35  
36 integrative review. *Psychological bulletin*, 109(2), 163 - 203.  
37  
38  
39 Mathôt, S., Fabius, J., Van Heusden, E., & Van der Stigchel, S. (2018). Safe and  
40  
41 sensible preprocessing and baseline correction of pupil-size data. *Behavior*  
42  
43 *Research Methods*, 50(1), 94-106.  
44  
45  
46 Mathôt, S., & Vilotijević, A. (2022). Methods in cognitive pupillometry: Design,  
47  
48 preprocessing, and statistical analysis. *Behavior Research Methods*, 1-23.  
49  
50  
51 Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex  
52  
53 function. *Annual review of neuroscience*, 24(1), 167-202.  
54  
55  
56 Norman, D. A., & Shallice, T. (1986). Attention to action. In *Consciousness and self-*  
57  
58 *regulation* (pp. 1-18). Springer, Boston, MA.  
59  
60

- 1  
2  
3 Parris, B. A., Hasshim, N., & Dienes, Z. (2021). Look into my eyes: Pupillometry  
4 reveals that a post-hypnotic suggestion for word blindness reduces Stroop  
5 interference by marshalling greater effortful control. *European Journal of*  
6 *Neuroscience*, 53(8), 2819-2834.  
7  
8  
9  
10  
11  
12 Parris, B. A., Hasshim, N., Wadsley, M., Augustinova, M., & Ferrand, L. (2021). The  
13 loci of Stroop effects: a critical review of methods and evidence for levels of  
14 processing contributing to color-word Stroop effects and the implications for  
15 the loci of attentional selection. *Psychological Research*, 1-25.  
16  
17  
18  
19  
20  
21  
22 Posner, M. I. (2012). Attentional networks and consciousness. *Frontiers in*  
23 *Psychology*, 3, 64.  
24  
25  
26  
27 Schmidt, J. R. (2013). The Parallel Episodic Processing (PEP) model: Dissociating  
28 contingency and conflict adaptation in the item-specific proportion congruent  
29 paradigm. *Acta Psychologica*, 142(1), 119–126.  
30  
31  
32  
33 Schmidt, J. R. (2019). Evidence against conflict monitoring and adaptation: An  
34 updated review. *Psychonomic Bulletin & Review*, 26(3), 753–771.  
35  
36  
37  
38 Schmidt, J. R., & Besner, D. (2008). The Stroop effect: why proportion congruent has  
39 nothing to do with congruency and everything to do with contingency. *Journal*  
40 *of Experimental Psychology: Learning, Memory, and Cognition*, 34(3), 514 -  
41 523.  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 Spinelli, G., & Lupker, S. J. (2021). Proactive control in the Stroop task: A conflict-  
4  
5 frequency manipulation free of item-specific, contingency-learning, and color-  
6  
7 word correlation confounds. *Journal of Experimental Psychology: Learning,*  
8  
9 *Memory, and Cognition, 47(10), 1550-1562.*
- 10  
11  
12 Spinelli, G., & Lupker, S. J. (2022). Robust evidence for proactive conflict adaptation  
13  
14 in the proportion-congruent paradigm. *Journal of Experimental Psychology:*  
15  
16 *Learning, Memory, and Cognition, 49(5), 675-700.*  
17  
18 <https://doi.org/10.1037/xlm0001144>.
- 19  
20  
21 Spinelli, G., & Lupker, S. J. (in press). Target-distractor correlation does not imply  
22  
23 causation of the Stroop effect. *Quarterly Journal of Experimental Psychology.*
- 24  
25  
26 Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of*  
27  
28 *Experimental Psychology, 18(6), 643-662.*
- 29  
30  
31 Tzelgov, J., Henik, A., & Berger, J. (1992). Controlling Stroop effects by manipulating  
32  
33 expectations for color words. *Memory & Cognition, 20(6), 727-735.*
- 34  
35  
36 Verguts, T., & Notebaert, W. (2009). Adaptation by binding: A learning account of  
37  
38 cognitive control. *Trends in Cognitive Sciences, 13(6), 252-257.*
- 39  
40  
41 West, R., & Baylis, G. C. (1998). Effects of increased response dominance and  
42  
43 contextual disintegration on the Stroop interference effect in older  
44  
45 adults. *Psychology and Aging, 13(2), 206.*
- 46  
47  
48 West, R., & Alain, C. (2000). Effects of task context and fluctuations of attention on  
49  
50 neural activity supporting performance of the Stroop task. *Brain*  
51  
52 *research, 873(1), 102-111.*
- 53  
54  
55 White, B. W. (1969). Interference in identifying attributes and attribute  
56  
57 names. *Perception & Psychophysics, 6, 166-168.*
- 58  
59  
60

## Figure Captions

Figure 1. Series of events during a single trial

Note: Text in bold indicate the sampling time-windows used in the pupillometry analysis. Participants responded by pressing the 1,2, or 3 keys on a standard QWERTY keyboard.

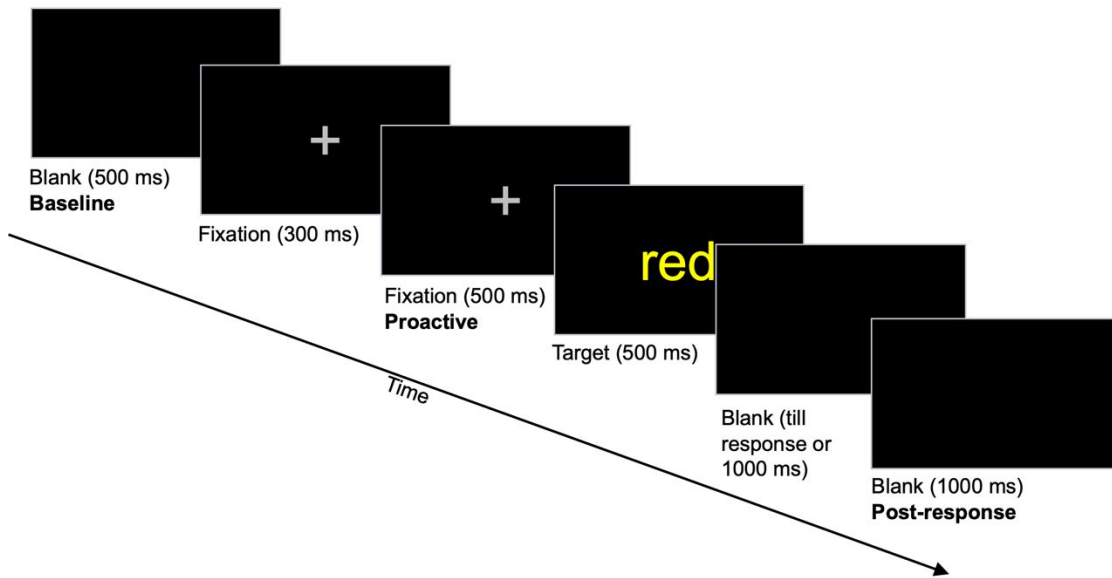
Figure 2. Mean response times for each condition in Experiment 1. Error bars represent 95% confidence intervals.

Figure 3. Mean increase in pupil size after response was made, for each condition in Experiment 1. Error bars represent 95% confidence intervals.

Figure 4. Pupil size (percentage change) relative to baseline Note: The data were downsampled to 100Hz as recommended by Mathôt (2022); and excludes samples not within 2.5 SD following Hershman & Henik (2019). The vertical axes represent the percentage change relative to baseline, while the horizontal axes show the time, in milliseconds, to target onset (0 representing onset).

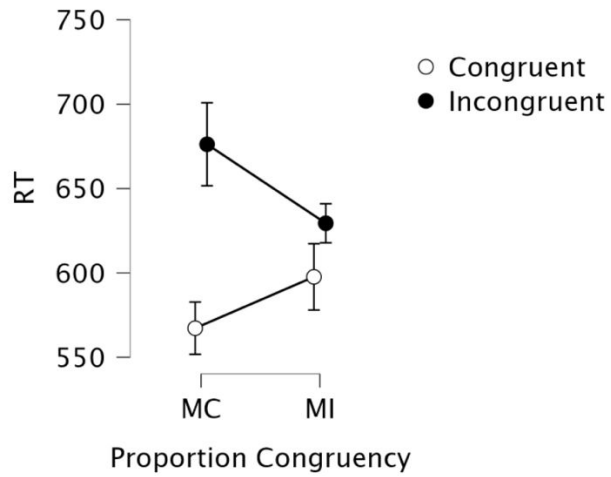
Figure 5. Mean response times for each condition in experiment 2. Error bars represent 95% confidence intervals.

Figure 6. Mean increase in pupil size in the exploratory post-response analysis in Experiment 2. Error bars represent 95% confidence intervals.



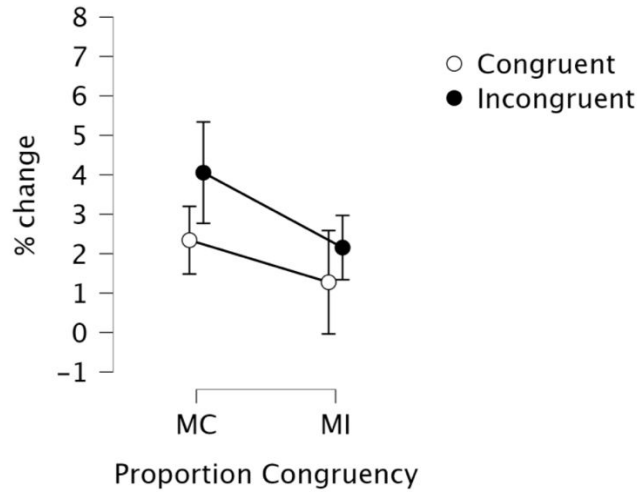
**Figure 1.** Series of events during a single trial

Note: Text in bold indicate the sampling time-windows used in the pupillometry analysis. Participants responded by pressing the 1,2, or 3 keys on a standard QWERTY keyboard.



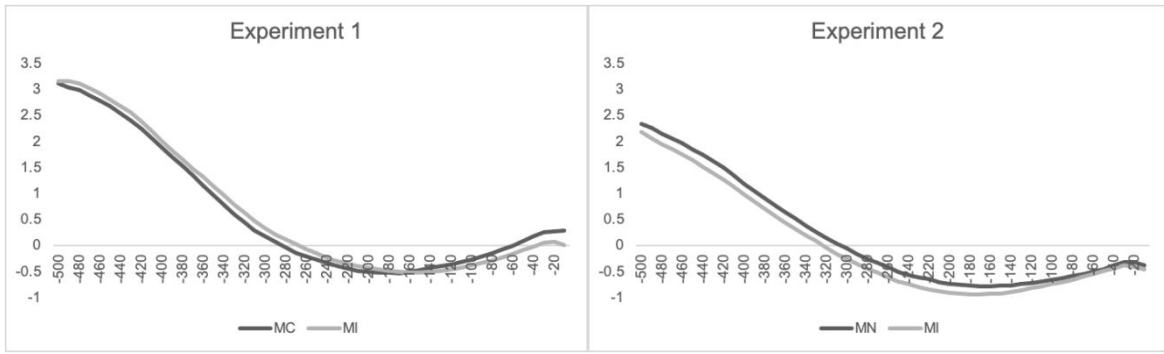
**Figure 2.** Mean response times for each condition in Experiment 1. Error bars represent 95% confidence intervals.



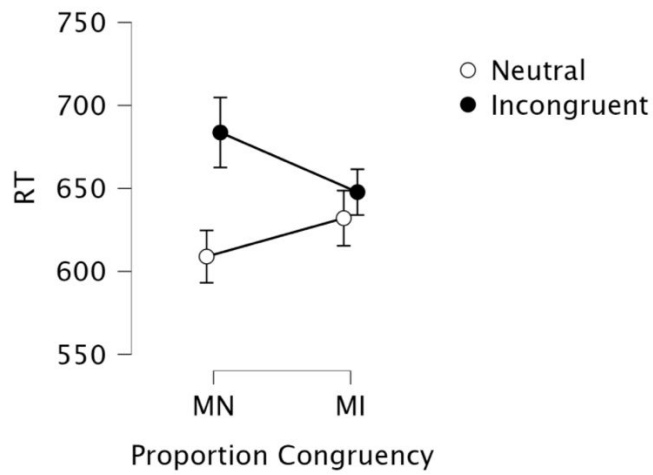


**Figure 3.** Mean increase in pupil size after response was made, for each condition in Experiment 1. Error bars represent 95% confidence intervals.

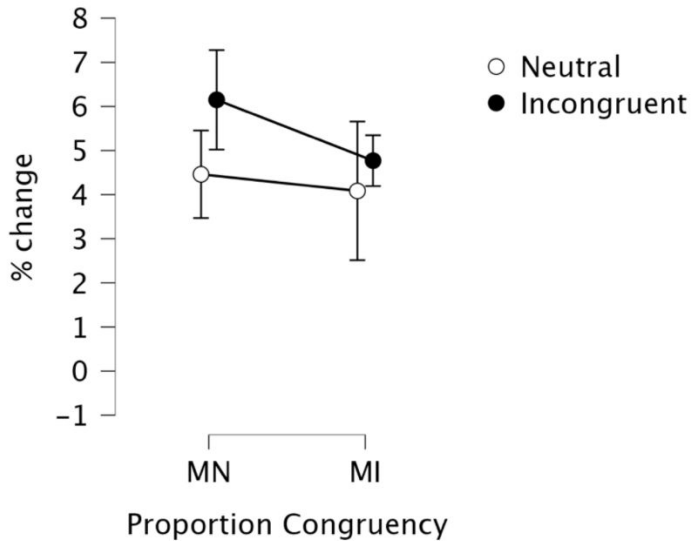
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



**Figure 4.** Pupil size (percentage change) relative to baseline Note: The data were downsampled to 100Hz as recommended by Mathôt (2022); and excludes samples not within 2.5 SD following Hershman & Henik (2019). The vertical axes represent the percentage change relative to baseline, while the horizontal axes show the time, in milliseconds, to target onset (0 representing onset).



**Figure 5.** Mean response times for each condition in experiment 2. Error bars represent 95% confidence intervals.



**Figure 6.** Mean increase in pupil size in the exploratory post-response analysis in Experiment 2. Error bars represent 95% confidence intervals.

**Table 1**  
*Means (SD) of response times and error rates across both experiments*

Experiment 1	MC		MI	
	Congruent	Incongruent	Congruent	Incongruent
RT (ms)	567.24 (94.40)	676.22 (146.52)	597.67 (129.01)	629.43 (122.05)
error rate (%)	2.70 (2.34)	7.78 (7.67)	3.78 (3.87)	5.18 (3.07)
Experiment 2	MN		MI	
	Neutral	Incongruent	Neutral	Incongruent
RT (ms)	608.91 (120.44)	683.66 (138.01)	632.05 (144.69)	647.75 (122.23)
error rate (%)	3.17 (2.19)	4.80 (5.72)	4.49 (5.56)	3.86 (3.40)

Peer Review Version

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60