Electrophysiological correlates of the reverse Stroop effect: Results from a simulated handgun task

Carolina Diaz-Piedra [a,b,*,1], Evelyn Gianfranchi [a,1], Andrés Catena [c], Leandro L. Di Stasi [a,d,*]

[a] Mind, Brain, and Behavior Research Center-CIMCYC, University of Granada, Campus de Cartuja, 18071 Granada, Spain
[b] College of Nursing and Health Innovation, Arizona State University, 550 N 3rd St, Phoenix 85004, AZ, USA
[c] Faculty of Psychology, University of Granada, Campus de Cartuja, 18071 Granada, Spain
[d] Joint Center University of Granada - Spanish Army Training and Doctrine Command, Gran Via n° 48, 18071 Granada, Spain

ABSTRACT

The color-word reverse Stroop (RS) effect still represents an interesting puzzle for cognitive researchers as an interference between incongruent ink colors and the meaning of the words is not always found. Here, we examined whether an unfamiliar and complex visuomotor task would produce a RS effect. Forty inexperienced shooters carried out a simulated shooting task. To test if the RS effect is related to the stimuli processing or to a late processing of the color (early and late time-windows), electroencephalographic global field power (GFP) variations were recorded with a high-impedance system (32 channels configuration in a standard monopolar montage, referenced to FCz and grounded to FPz). The color-word RS effect was reflected in the performance of 32 participants, suggesting that the strength of the association between the target and the specific response requested might be central to the RS interference. This behavioral result was paralleled by GFP modulations in 20 participants. A significant increase of the GFP for the congruent trials (e.g., the word "red" written in red ink) was recorded after stimulus presentation (conflict detection), followed by an increase for the incongruent trials (e.g., the word "red" written in green ink) just before the shooting (conflict resolution). Despite the limitations of the study, such as the inclusion of a low number of channels in the GFP analyses, the results suggest that the RS interference is easily elicited in tasks requiring an unfamiliar response, which supports the strength of association hypothesis. Moreover, as implied by the GFP modulations, the interference might occur early in time, but also in a later stage, closer to the response.

1. Introduction

The reverse Stroop (RS) effect was first described by J. R. Stroop in his third experiment in which a new, unexpected interference was observed after eight days of practicing color naming of incompatible words (the classic Stroop task): the ink color interfered with word reading (Stroop, 1935). Since then, the RS effect is an interesting puzzle for cognitive researchers because failing to inhibit this interference would also deteriorate an automatic response (for experienced readers), such as perceiving and responding to the meaning of familiar words (Danziger et al., 2002; Washburn, 2016). However, whereas the well-known Stroop effect has been examined in hundreds of studies, much less (and inconsistent) research is available on the RS effect (for a review, see MacLeod, 1991). In a standard color-word RS task (see Fig. 1),

when an incongruent ink color is presented (e.g., the word "red" written in green ink), participants must respond to the word meaning (i.e., the relevant information, for example, reading aloud the word "red" or targeting a red patch) while ignoring the ink color (i.e., the distracting information, green, in this particular case). The expected RS effect derives from the interference produced by incongruent ink colors in responding to the meaning of the words (through either any verbal or motor response), relative to neutral (e.g., the word "red" written in black ink) or congruent ink colors (e.g., the word "red" written in red ink). Such interference would lead to a decrease in speed and accuracy for incongruent stimuli, which is the RS effect.

In the last 30 years, a variety of experiments have been carried out using different versions of the color-word RS task with samples from both patients and healthy individuals (for a detailed list of studies, see

---

Table 1). The results from these studies suggest that a RS effect may not always occur. Inconsistent findings could be due to differences in experimental protocols, particularly the kind of response requested by task instructions. Indeed, a translation between the stimuli and the response to the target seems to be the key. Therefore, as verbal (word meaning) and sensory (color) information is processed in separate systems with different codes (Blais and Besner, 2007), an interference would be more likely if participants need to carry out a translation to respond correctly (Durgin, 2000, 2003). In a RS task, a verbal, dominant, response would not be affected by the color of the stimulus (i.e., the distracting information), as the relevant stimulus information and the response are compatible. However, the limited empirical evidence (10 papers) is equivocal in this regard. The majority of the studies that requested a verbal response (i.e., reading aloud the word stimuli) did find the RS effect on speed (6 found an increase in reaction times, 4 did not), but not on accuracy for incongruent stimuli (5 did not find any change in the number of correct responses, 5 did not report it). For those studies, a hypothetical translation model would only account for accuracy results.

However, following this hypothetical translation model, a motor, non-dominant response would be affected by the color of the stimulus (i.e., the distracting information), since the relevant stimulus information (i.e., the color) and the motor response are incompatible, therefore producing an interference. The variety of methodologies among the RS studies that requested a motor response makes a straightforward analysis challenging. The simplest RS task that requires a motor response uses a paper and pencil form where participants are asked to choose the correct answer (color patch) among a set of alternatives. The 8 studies that used such methodology found a consistent decrease in accuracy for incongruent stimuli, which is coherent with a hypothetical translation model. However, speed assessment was not considered in the analysis. Most of the studies using motor responses requested pressing a specific key on a keyboard (either color-name or color labelled keys, or unlabeled keys, number of studies [n] = 19). They do allow tracking correctly both speed and accuracy, but failed to provide consistent results. All these studies found a RS effect on speed for incongruent stimuli (an increase in reaction times), even with color labelled keys, which would not require a translation. Moreover, the effect was modulated by other variables (e.g., stimulus onset asynchrony, type of cue, trials order; see Lu and Proctor, 2001; Shalev and Algom, 2000; Simon and Baker, 1995; Steinhauser and Hübner, 2009). The RS effect on accuracy for incongruent stimuli was not clear, as only 10 out of the total of 19 studies found a decrease in the number of correct responses. Notably 2 of the studies that did not find a decrease in the accuracy used color labelled keys (Blais and Besner, 2006; Davies et al., 1991), although they both reported increased RT for the congruent condition. In studies employing more complex – than just pressing a key – motor responses (e.g., pointing the correct answer with a PC mouse), the RS effect on speed and accuracy is also unclear. For instance, a speed decrease for incongruent stimuli was reported in 3 studies where participants used the mouse to choose the color corresponding to the meaning of the stimulus

word among a set of colored patches presented on the screen. However, the effect on accuracy was not confirmed in the two studies where color patches were substituted by words. Similarly to what observed before, a hypothetical translation model would only account for accuracy results.

As the procedures followed are not detailed in most RS studies, it is difficult to determine if a translation to respond correctly was needed. Trying to disentangle this issue, Blais and Besner (2006) designed a series of experiments to test the hypothetical translation model. The experiments required either a vocal response, pressing color-name labelled keys, or pointing color names with the PC mouse. As a RS effect was observed even in the absence of translation (Blais and Besner, 2006, 2007), the authors formulated the so-called "strength of association" hypothesis to explain their findings which emphasized the importance of the specific response and the amount of practice participants have developed in making the response (a lot of practice, in the case of reading words). With respect to motor – and therefore less practiced – responses, they proposed that the RS effect could be the result of a late processing of the color. This would mean that, during the extra time it takes to select a motor response, the color is processed producing an interference. Based on this, we hypothesized that a less practiced response (even less practiced than pressing a key) would produce a RS effect on response times and accuracy. If present, such effect would be related to the processing of the stimuli (stimulus conflict) or to a late processing of the color (response conflict).

Recently, Wood and collaborators (Wood et al., 2016) conducted a RS experiment using an unfamiliar visuomotor task: a RS simulated handgun task with a plastic gun, performed by participants with no experience in handgun shooting. The RS effect was expressed by a worse shot accuracy for incongruent trials. However, this result was modulated by participants' anxiety level, as induced by a "high-threat" study condition (induced pressure to perform). Therefore, the actual interference remains unclear. The authors also measured participants' eye movements. They found that the participants struggled to inhibit a sort of visual interference created by the irrelevant information (the color), as the process caused a significant increase in visual search time for incongruent stimuli (Wood et al., 2016). However, this interpretation regarding the potential RS underpinning processes does not provide any information whether a stimulus or a response conflict has happened. To study this, we adapted the procedure with the RS simulated handgun task followed by Wood and collaborators, but using electroencephalographic (EEG) metrics.

EEG has proved its utility to precisely track the processes underpinning cognitive tasks (Cacioppo et al., 2007) and it has been used, quite often, for the study of other interference tasks such as classic Stroop, Go/No-Go, or Simon tasks (see Heidlmayr et al., 2020 for a review; see Supplementary Table S1 for a summary). However, very few studies have tried to address the RS effect using EEG. For instance, significant modulations on power spectra and on various early and late event-related potentials (ERP) components have been found during an RS task requiring pressing a key (Appelbaum et al., 2014; Atkinson et al., 2003; Compton et al., 2018; Kray et al., 2005). Among the early
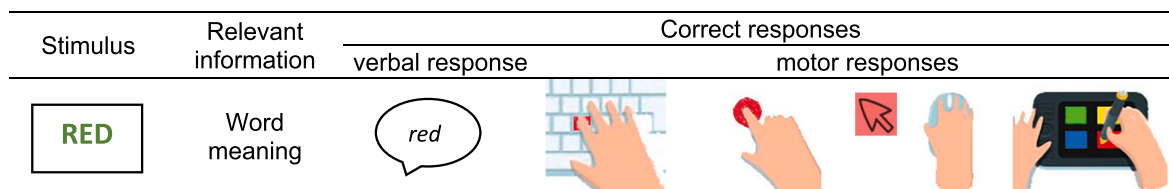
Fig. 1. In a standard color-word reverse Stroop task, when an incongruent ink color is presented (for example, the stimulus on the left), participants must respond to the word meaning, which is the relevant information. In the last 30 years, many modified versions of this task have been developed. While maintaining the critical structural elements, variants of the test differ in presentation modalities (e.g., computerized or paper and pencil, number of trials, items, colors, stimuli ratio), in scoring methods for speed (e.g., reading time, reaction times), and/or accuracy (e.g., errors, correct answers rates). Moreover, tasks differ in the requested responses, which can be either verbal (reading aloud) or motor responses (e.g., ticking the correct answer, pressing a key). (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

Table 1

Characteristics and results of the studies employing the color-word reverse Stroop task in the last 30 years (*n* = 47, since the review conducted by MacLeod, 1991).

Author(s)	Exp.[1]	Sample[2]	Instructions	Reference[3]	Ratio[4]	RS effect[5] RT	Accuracy
Verbal response							
Dulaney and Rogers, 1994	4	40 students (age: 22.6) 40 elders (age: 69.9)	Reading aloud	Neutral (white or black)	$33_i{:}33_c{:}33_n$	Students ↑ Elders ↑	n/a
Hepp et al., 1996	1	44 schizophrenia patients (age: 36.0) 50 controls (age: 36.9)	Reading aloud	Neutral (white)	$33_i{:}33_c{:}33_n$	Patients ∅ Controls ∅	n/a
Moriguchi and Morikawa, 1998	1	24 students (age: 21.0)	Reading aloud	Neutral (black)	$40_i{:}40_c{:}20_n$	↑	n/a
Woodward et al., 2002	1	30 Parkinson patients (age: 69.3) 34 controls (age: 69.9)	Reading aloud	Neutral (black)	$50_i{:}50_n$	Patients ↑ Controls ↑	n/a
Roelofs, 2005	3	56 students (age: n/a)	Reading aloud	Congruent	$33_i{:}33_c{:}33_n$	↑	∅
Peru et al., 2006	1	163 children (age: 6–10)	Reading aloud	Neutral (black)	$60_i{:}20_c{:}20_n$	↑	∅
Zimmer et al., 2016	1	121 students (age: 23.8)	Reading aloud	Neutral (black)	$50_i{:}50_n$	↑	n/a
Verbal/motor response							
Logan and Zbrodoff, 1998	3	90 students (age: n/a)	Reading aloud; Typewriting on a keyboard.	Congruent	$50_i{:}50_c$ $50_i{:}50_c$ $20_i{:}80_c;$ $80_i{:}20_c$	Verbal ∅ Manual ↑	Verbal ∅ Manual ↓
Ikeda et al., 2010	1	47 students (age: 20.8)	Reading aloud; Ticking the correct alternative.	Congruent	$50_i{:}50_c$	Verbal ∅ Manual ↑	Verbal ∅ Manual ↓
Motor response							
Color labelled keys							
Davies et al., 1991	3	61 adults (age 22–35)	Pressing one out of four keys	Congruent	$50_i{:}50_c$	↑	∅
Elvevåg et al., 2000	1	28 schizophrenia patients (age: 36.0) 48 controls (age: 43.0)	Pressing one out of two keys	Congruent	n/a	Patients ↑ Controls ↑	Patients ↓ Controls ↓
Atkinson et al., 2003	1	10 adults (age: 29.7)	Pressing one out of three keys	Congruent	n/a	↑	n/a
Appelbaum et al., 2014	2	78 adults (age: 18–38)	Pressing one out of four keys	Congruent	$50_i{:}50_c$ $20_i{:}80_c$	↑	n/a
Color(−name) labelled keys							
Sugg and McDonald, 1994	1	56 students (age: n/a)	Pressing one out of two keys (color or color name condition)	Congruent	$33_i{:}33_c{:}33_n$	Color ∅ Color name↑	n/a
Blais and Besner, 2006[a]	4	156 students (age: n/a)	(Exp 1) Reading aloud or pressing one out of four keys (color name); (Exp 2–4) Pressing one out of two keys (color name)	Congruent (Exp 4) Neutral (white)	$50_i{:}50_c$ $75_i{:}25_c$ $25_i{:}50_c{:}25_i$ $25_i{:}50_c{:}25_i$	Verbal ∅ Manual ↑ ↑ ↑ ↑	Verbal ∅ Manual ↓ ↓ ∅ ↓
Unlabelled keys							
Melara and Mounts, 1993	2	192 students (age: n/a)	Pressing one out of two keys	Congruent	$50_i{:}50_c$	↑	n/a
Simon and Baker, 1995	3	96 students (age: n/a)	Pressing one out of two keys	Congruent	$50_i{:}50_c$	↑	∅
Shalev and Algom, 2000	1	16 students (age: n/a)	Pressing one out of two keys	Congruent	n/a	↑	∅
Lu and Proctor, 2001	2	82 students (age: n/a)	Pressing one out of two keys	Congruent	$50_i{:}50_c$	↑	∅
Ruff et al., 2001	1	12 students (age: 23.1)	Pressing one out of four keys	Neutral (black)	$50_i{:}50_n$	↑	∅
Kray et al., 2005	1	14 students (age: 21.7) 14 elders (age: 62.9)	Pressing one out of four keys	Congruent	n/a	Students ↑ Elders ↑	n/a
Nishimura, 2006	2	–	Pressing a key	Congruent	–	↑	↓
Kertzman et al., 2006	1	62 gamblers (age: 40.5) 83 controls (age 40.5)	Pressing one out of two keys	Neutral (black)	$33_i{:}33_c{:}33_n$	Gamblers ↑ Controls↑	Gamblers ↓ Controls↓
Woodward et al., 2006	1	12 adults (age: 34.5)	Pressing one out of four keys	Neutral (black)	$50_i{:}50_n$	↑	n/a
Steinhauser and Hübner, 2009	2	48 adults (age: 23.4)	Pressing one out of two keys	Neutral (colored "X")	$40_i{:}30_c{:}30_n$	↑	↓
Song and Hakoda, 2015	1	20 students (age: 21.7)	Pressing one out of five keys	Congruent	n/a	↑	↓
Compton et al., 2018	1	30 students (age: n/a)	Pressing one out of six keys	Congruent	$50_i{:}50_c$	↑	↓
Sobel et al., 2020	5	84 students (age: 21.7)	Pressing one out of four keys	Congruent	$50_i{:}50_c$	↑	↓
Mouse							
Durgin, 2000	1	40 students (age: n/a)	Selecting a colored patch	Neutral (gray)	$50_i{:}50_n$	↑	↓
Durgin, 2003	2	67 students (age: n/a)	(Exp 1) Selecting a colored patch; (Exp 2) Selecting a colored patch (present) or go (absent)	Out-of-set incongruent	$50_i{:}50_i$	↑	↓

potentials, the N200 (~250 ms latency; Appelbaum et al., 2014) has been consistently reported as related to the RS conflict monitoring or detection processes (Kray et al., 2005; but see Atkinson et al., 2003). Among the late components (>400 ms latency), the so-called late positive component (LPC; Appelbaum et al., 2014), the frontal lateralized negativity (Atkinson et al., 2003), and the N400/N450 (Zurrón et al., 2009) seem associated to the RS conflict suppression or resolution (e.g., Heidlmayr et al., 2020; Kray et al., 2005) (see Table S1 in the Supplementary Material for more details). However, depending on the specific features of the experimental paradigms, the modulations were also explained as due to (1) the features of a spatial cue presented prior to the stimuli (Compton et al., 2018), and (2) to the manipulation of the proportion of incongruent trials (Appelbaum et al., 2014). Overall, these studies reported great inconsistencies in results, but also in methodologies (task manipulations, focusing on either early or late EEG modulations), which makes impossible reaching a comprehensive framework.

Here, we tried to clarify the processes underlying the RS effect, using an unfamiliar and complex visuomotor task with a good ecological validity: a RS simulated handgun task. We tested participants with no experience in shooting with a RS simulated handgun task using a plastic gun (Wood et al., 2016), while measuring their EEG Global Field Power (GFP). The GFP, the amount of activity of the potential field, reflects the general brain activity elicited by a stimulus and its processing (Lehmann and Skrandies, 1980; Skrandies, 1990). We calculated the GFP in two specific time-windows, a stimulus-locked window and a response-locked window. We expected to find a clear RS effect: an increase in response times and errors, as well as worse shot accuracy, when responding to the incongruent stimuli. Moreover, we expected a GFP modulation (i.e., amplitude) depending on the stimulus type, reflecting the interference at the basis of the RS effect. The modulation should be present in both time-windows, but specially in the response-locked window, as suggested by several studies focusing on late ERP components (e.g., Appelbaum et al., 2014; Kray et al., 2005). Therefore, the GFP modulation would reflect differences in the overall brain activity requested for processing the two types of stimuli (congruent vs. incongruent). On the other hand, differences in the GFP modulation in the two time-windows should reflect possible differences between early stimulus processing (conflict detection) and response preparation (conflict resolution).

Table 1 (*continued*)

Author(s)	Exp.[1]	Sample[2]	Instructions	Reference[3]	Ratio[4]	RS effect[5]	
						RT	Accuracy
Blais and Besner, 2007	1	32 students (age: n/a)	Selecting a word	Neutral (white)	$60_i{:}20_c{:}20_n$	↑	∅
Durgin et al., 2008	1	87 adults (age: n/a)	Selecting a colored patch	Neutral (gray)	$66_i{:}33_n$	↑	↓
Yamamoto et al., 2016	1	20 students (age: 20.8)	Selecting a word	Neutral (white)	$33_i{:}33_c{:}33_n$	↑	∅
Rosenberg and Strohl, 2018	1	23 adults with sleep issues (age: n/a)	Selecting a colored patch	Congruent	n/a	↑	n/a
Pen							
Sasaki et al., 1993	1	–	Ticking the correct alternative	Congruent	–	–	–
Matsumoto et al., 2012	1	1945 participants (age: 7–86)	Ticking the correct alternative	Congruent	–	–	–
Takeuchi et al., 2012	1	118 students (age: 21.6)	Ticking the correct alternative	Congruent	n/a	n/a	↓
Ikeda et al., 2013a	1	20 elders (age: 70.8)	Ticking the correct alternative	Congruent	n/a	n/a	↓
Ikeda et al., 2013b	1	376 children and students (age: 5–24)	Ticking the correct alternative	Congruent	n/a	n/a	↓
Watanabe et al., 2013	1	216 students (age: 19.9)	Ticking the correct alternative	Congruent	–	–	–
Takeuchi et al., 2015	1	958 adults (age: 20.8)	Ticking the correct alternative	Congruent	n/a	n/a	↓
Nagamine, 2017	1	21 schizophrenia patients (age: 34–61)	Ticking the correct alternative	Congruent	n/a	n/a	↓
Other							
Smithson et al., 2006	2	13 adults (age: 20–62)	Selecting a colored patch (joystick)	Each color combination	$70_i{:}15_c{:}15_n$	↑	n/a
Yasumura et al., 2014	1	10 ADHD patients (age: 11.2) 11 Asperger patients (age: 10.5) 15 controls (age: 9.5)	Selecting a colored patch (touch screen)	Congruent	n/a	n/a	ADHD ↓ Asperger ↓ Controls ↓
Wood et al., 2016	1	24 students (age: 20.15)	Shooting at a colored target (plastic gun)	Congruent	$75_i{:}25_c$	n/a	↓
Miller et al., 2016	1	18 adults (age: 22.7)	Selecting a colored patch (cylinder on an interactive tabletop)	Congruent	$50_i{:}50_c$	↑	↓
Yasumura et al., 2019	1	67 ADHD patients (age: 9.8) 140 controls (age: 9.8)	Selecting a colored patch (touch screen)	Neutral (black)	n/a	n/a	Patients ↓ Controls ↓

Note. RT = reaction/response times. [1] = number of experiments employing the RS; [2] = for each study we report *n*, mean age or, in alternative, age-range, and groups; [3] = category used as a reference for the incongruent stimuli; [4] = stimuli presentation ratio (percentage), with $_i$, $_c$, and $_n$ indicating respectively incongruent, congruent, and neutral stimuli; [5] = behavioral dependent variables assessed to find an RS effect; ∅ = no effect; ↑ = significant RS effect, meaning an increase in the dependent variable values for incongruent stimuli; ↓ = significant RS effect, meaning a decrease in the dependent variable values for incongruent stimuli; n/a = not measured; – = details not available due to the language of the paper (Japanese).

When present, differences between experiments (see column *Exp.*) in terms of response modality, reference, or ratio are reported in the correspondent columns, one row for each experiment. Results are always reported according to group or response modality.

[a] Blais and Besner, 2006: this is the only study that, manipulating at the same time both modality response and ratio, found significant differences requiring a detailed results notation.

2. Method

2.1. Experimental design

The experiment consisted of a RS simulated handgun task (48 stimuli in total) in which the participants had to shoot with a plastic gun at the target stimulus colored with the ink color corresponding to the meaning of the word presented at the center of the screen. The study followed a within-subjects design with *congruency* as the independent variable (two task conditions, congruent vs. incongruent). To assess shooting performance, we considered three dependent variables: mean response times (RT; both overall and for hits only), hit rate (percentage of correct responses), and shooting accuracy. Regarding EEG metrics, we considered the GFP (Lehmann and Skrandies, 1980; Lehmann and Skrandies, 1984) measured in two 1-s time-windows as the dependent variables: a stimulus-locked time-window (0 to 1 s after the onset of the stimulus) and a response-locked time-window (from the response, i.e., pressure on the gun trigger, to −1 s, to capture EEG activity *before* the response).

2.2. Participants

A convenience sample of 40 students of the University of Granada (mean [M] age ± standard deviation [SD] = 19.8 ± 2.8 years; 25 females) took part in the study. Inclusion criteria were the following: (1) normal or corrected-to-normal (contact lenses) vision, and (2) appropriate levels of arousal before the experimental session, operationalized as a score lower than 3 on the Stanford Sleepiness Scale (SSS, Connor et al., 2002; Diaz-Piedra et al., 2021; Hoddes et al., 1973), which indicates no fatigue and/or sleepiness. Exclusion criteria were the following: (1) a medical history of significant head injuries or neurological disorders, (2) problems with color perception, and (3) a score equal or above 50 at the Cognitive Failures Questionnaire (CFQ, Broadbent et al., 1982). The CFQ assesses individual proneness to commit cognitive errors in everyday tasks. CFQ average scores was 27.9 (SD ± 9.2), indicating a homogeneous sample of participants with a low rate of self-report cognitive failures. We excluded one female participant after the arousal assessment, and two participants (1 male and 1 female) because they suffered from mild-deuteranomaly (red-green color blindness in which the green cones do not detect well the green color and, at the same time, are too sensitive to yellows, oranges, and reds; Clark, 1924).

2.3. Stimuli

The RS simulated handgun task was designed and presented using E-Prime software, version 3.0 (Psychology Software Tools Inc., Sharpsburg, PA, USA). Each stimulus included a central black circle (30 cm diameter), in which a word was displayed (RED, GREEN, BLUE or YELLOW, written in the native language of the participants [i.e., Spanish]). The word was written in either the congruent (36 stimuli) or an incongruent (12 stimuli) ink. The proportion between congruent and incongruent stimuli followed the literature recommendations to keep 75% of congruency rate in order to induce optimum levels of interference (Kane and Engle, 2003). The central circle was framed by four 30 cm-diameter targets, displayed in each corner of the projected image (see Section 2.4). Each target, one for each of the four colors, consisted of ten concentric rings with a 'bull's-eye' (i.e., the center) of 3 cm diameter. Hitting the bull's-eye awarded 1 point, whereas each ring emanating from the center (1.5 cm width) awarded ascending values up to 10 points for the most external circle, with a higher score indicating a less accurate shot. Participants had to shoot at the target that was colored as indicated by the meaning of the word in the center of the screen, ignoring its ink-color. Each stimulus was preceded by a fixation point (5 s) and then presented for 2 s. We used the stimuli developed by Wood et al. (2016), changing the language from English to Spanish and the background color from black to white because the wall in which we projected the stimuli was white (see Section 2.4).

2.4. Apparatus

We used a video projector (EB-X24, EPSON Pty Ltd., Australia) linked to an Acer PC (Microsoft Windows™ XP operating system) running *E*-Prime software, version 3.0, to display the RS stimuli on one wall of the laboratory. The size of the projected image was 134×100 cm, which was about 250 cm away from the participants (resulting in a view angle of ~30° horizontally and ~23° vertically). A small action camera (SJ5000X Elite, SJCAM Global Inc., China) was put in front of the projector (without interfering with the projection area) to record a video (120 fps and 720p resolution) of the performance in terms of hit rate and shooting accuracy. Participants shot foam ammunitions using a NERF Rival Kronos XVIII-500 plastic gun (Hasbro Inc., Pawtucket, RI, USA) with a capacity of 4 ammunitions. During the task, we tracked the exact time of each shot (i.e., RT) by means of a marker synchronized with the EEG recording. The marker was developed including a small button on the gun trigger that sent a signal to an Arduino UNO Board for each shot. Once the Arduino UNO Board received the marker, it sent a Transistor-Transistor Logic pulse of 20 ms to the EEG signal amplifier (see Section 2.5) to synchronize the recorded brain activity with the shot. Moreover, another marker was sent to the EEG amplifier both at the onset and at the offset of the stimuli.

We recorded brain activity (500 Hz) using the actiCAP Xpress Twist dry-electrode EEG system (BrainProducts GmbH, Gilching, Germany). The system included 32 active, high-impedance dry electrodes (10/20 International System; Jasper, 1958) and, additionally, one reference and one ground electrode. Electrodes were applied using a flexible neoprene cap (54–60 cm head circumference) with a chin belt to ensure the fastening on the participant's head. Three different gold-plated electrode typologies are included in the system to adjust the electrodes to participants' scalp through the so-called QuickBits: five T-shaped flat tips placed on the forehead (Fp1, Fp2, F7, F8, and passive Ground), seven 14 mm-length mushroom-head tips (TP9, P7, O1, Oz, O2, P8, TP10), and 22 electrodes of the 12 mm-length mushroom-head tips covering the rest of scalp positions. Before electrode placement, the areas of the skin corresponding to the electrodes were gently cleaned with alcohol. To increase the signal quality, we used SuperVisc High Viscosity Electrolyte-Gel for Active Electrodes (EASYCAP GmbH, Herrsching, Germany) on the forehead, Cz, and reference electrodes. The actiCap Twist array was connected to a LiveAmp EEG signal amplifier (BrainProducts GmbH, Gilching, Germany), which sent filtered data (low-pass third order sinc filtering with −3 dB 131 Hz frequency) through a Bluetooth dongle connected on a Lenovo Laptop (13 in. screen, 1920×1080 pixels resolution, Microsoft Windows™ 10 operating system). The laptop ran BrainVision Recorder version 1.21.0201 (BrainProducts GmbH, Gilching, Germany) to record the brain activity for each participant and to store it in a .vhdr file.

2.5. EEG analyses

The .vhdr files were preprocessed and analyzed using EEGLab Matlab toolbox (Delorme & Makeig, 2004). The recording signal reference was FCz, also kept for preprocessing and analysis. First, we cleaned the signal by removing line fluctuations, scanning for frequencies ±1 Hz around the 50, 100, 150, and 200 Hz through the EEGlab Cleanline function (Mullen, 2012). After that, a band pass 0.1–30 Hz, 36 dB/octave filter, was used. Then, channels with a flatline duration of more than 5 s or with excessive signal-to-noise ratio (i.e., ±4 SD) were identified, removed, and not included in the successive analyses. Therefore, no periods of data saturation (flatline) were included in the time-windows (see below) employed for EEG analysis. Afterwards, a notch filter (45–55 Hz) was applied to improve the quality of the signal by removing powerline noise. The remaining channels were manually inspected for residual noise, and a part of them was interpolated if needed (spherical

spline interpolation method included in EEGLab; see Supplementary Table S2 for more information about the EEG channels included in the GFP analysis for each participant). Two time-windows (epochs) were selected for both task conditions (i.e., congruent vs. incongruent): a stimulus-locked (from 0 up to 1 s *after* the stimulus onset), and a response-locked (from 0 up to 1 s *before* the response, i.e., the precise moment of the shot). All epochs were baseline corrected and visually inspected for artifacts and eye movements (e.g., blinks; Kappenman and Luck, 2010). As a final step, all these epochs, as well as those exceeding ±100 μV, were discarded.

The number of clean, suitable channels varied from 5 channels for participant #2 to 19 channels for participants #19 and #27, before interpolation. The number of interpolated channels varied from 0 channels for participants #18, #25, #27, #35, and #39 to 10 channels for participant #38. Only signals with at least 9 final clear channels after interpolation, and at least 70% of overall artifact-free epochs (i.e., 48 epochs for each stimulus-locked and response-locked window and participant) were included in the EEG analysis. The high discarding rate was due to both the task features (for each shot, the recoil of the gun tended to induce movements in participants' head and shoulders, leading in some cases to a high number of movement artifacts), and the characteristics of the high-impedance active-dry actiCAP Xpress Twist system itself. Indeed, it tends to exhibit increased noise levels, and to be prone to electrode detachment in the case of hairy scalps, although reliably measuring EEG-related metrics (Mathewson et al., 2017).

A script was then employed to calculate the GFP (reference-free) (Lehmann and Skrandies, 1980, 1984), which is a way to compute the potential field strength by means of the spatial SD of average-referenced electrode voltages (Files et al., 2016; Lehmann and Skrandies, 1980, 1984). The GFP is thus considered a summary of multi-channel EEG data (Files et al., 2016). Here, we employed the following formula, as implemented in EEGLab, to calculate the GFP over the two time-windows for each condition in each participant:

$$GFP\ (reference\text{-}free) = \operatorname{sqrt}\left(\left(\frac{1}{2n}\right) * \operatorname{sum}_{i}\left(\operatorname{sum}_{j(ui-uj)}^{2}\right)\right)$$

where *n* is the number of electrodes which measure the potentials, *ei* and *ej*; *i, j* = 1 … n; the observed voltages are *ui* = *ei* - common reference.

2.6. Procedure

The study was conducted in conformity with the Code of Ethics of the World Medical Association (Declaration of Helsinki (W.M.A., 1964)), following the guidelines of the University of Granada's Institutional Review Board (IRB approval #985/CEIH/2019). Written informed consent was obtained from each participant prior to the experimental procedure.

After signing the consent form, participants filled in a sociodemographic data form, the SSS, and the CFQ. Then, they carried out the Color Blind Test (a modified version of the Ishihara plates [Clark, 1924]; Color Blind Test, EnChroma Inc., Berkeley, CA, USA).

After that, the experimenter provided participants with a demonstration on how to use and reload the plastic gun. Participants also completed a familiarization phase with the gun, shooting three times at the wall and reloading the gun. Then, after cleaning the scalp, the EEG cap was positioned on the participant's head. In addition, participants wore impact resistant plastic safety glasses for the entire duration of the procedure.

At this point, participants faced the RS simulated handgun task, preceded by an 8-stimuli practice. During practice, participants were presented with two different types of stimuli (each of them appeared 4 times, in random order, always at the center of the screen): (1) a white 30 cm-diameter target, to which the participants had to shoot as quickly and accurately possible, and (2) a full color circle (red, green, blue, or

yellow; 30 cm of diameter), surrounded by four color circles. In this last case, the participants had to shoot at the target colored with the same color of the stimulus. During the RS simulated handgun task (see Fig. 2), participants were asked to ignore the ink color of the color word presented at the center of the screen, shooting instead, as quickly and accurately as possible, at the target of the color corresponding to the meaning of the word. For the entire task, after 4 shots, a white screen displayed an instruction inviting the participants to reload the gun and then press the space bar to continue with the task once they were ready.

Five participants (all females) had not enough strength to reload the gun by pulling back the slide. They could not complete the practice and were then discarded. The final number of participants that completed the task was 32 (M age ± SD = 19.8 ± 2.8 years; 18 females), 20 of whom had also a complete EEG recording (M age ± SD = 19.9 ± 3.3 years; 11 females) (see Table 2 for the demographics of the two final samples).

2.7. Statistical analyses

For both congruent and incongruent conditions, we calculated the mean RT (both overall and for hits only), the hit rate (percentage of hits = [*n* of hits for condition / *n* of stimuli for condition] * 100), and the mean shooting accuracy. The mean RT per condition (both overall and for hits only) was automatically computed via a script run on Matlab software using the signals sent to the EEG by the gun trigger. To measure both the hit rate and the shooting accuracy, two independent evaluators inspected frame-by-frame the videos that recorded participants' performance. Any shot hitting the correct target (i.e., the target colored with the color corresponding to the word meaning) was defined as a hit, a correct answer. The shooting accuracy was measured, only for the hits, in terms of distance from the center of the target (score). Intraclass correlation coefficient (ICC) estimates and their 95% confident intervals were computed for both hit rate and shooting accuracy, based on a mean-rating (*k* = 2), using a 2-way mixed effects model for absolute agreement between evaluators.

We conducted three paired *t*-tests to compare the RTs, the hit rate (arcsine-transformed), and the shooting accuracy at the congruent and the incongruent conditions. For significant differences, we calculated Cohen's *d* effect size for within-subjects designs. Behavioral analyses included all the 32 participants that completed the task. Due to technical problems with the video recording, recorded performance could not be assessed for one participant. In this case, we interpolated the hit rate and shooting accuracy values using the group means for both conditions. The interpolation of these two values did not affect the significance of the performed tests (see section below).

Concerning the GFP, a 5000 random two-tailed permutation test (see the Matlab function "mult_comp_perm_t1", Files et al., 2016; Groppe et al., 2011) was employed to check whether the average GFP values significantly differed between conditions (i.e., GFP elicited by congruent vs. incongruent stimuli) in each time-point of the two time-windows of interest (response-locked and stimulus-locked). The obtained *p*-values were adjusted to control for family-wise error by the "tmax" method for multiple comparisons (Blair and Karniski, 1994; Westfall and Young, 1993). Internal consistency of the GFP data was confirmed by calculating the dependability estimate and the Interclass Correlation Coefficients (ICC) (ERP Reliability Analysis [ERA] toolbox; Clayson and Miller, 2017), as reported in the Supplementary Material (section "Global field power [GFP] internal consistency analysis").

3. Results

3.1. Shooting performance

Mean overall RT (±SD) for incongruent stimuli (1225 ± 190 ms) was higher than for congruent ones (1004 ± 185 ms), *t*(31) = −9.56, *p* < 0.001, *d* = 1.79. This significant trend was exhibited by 31 out of 32
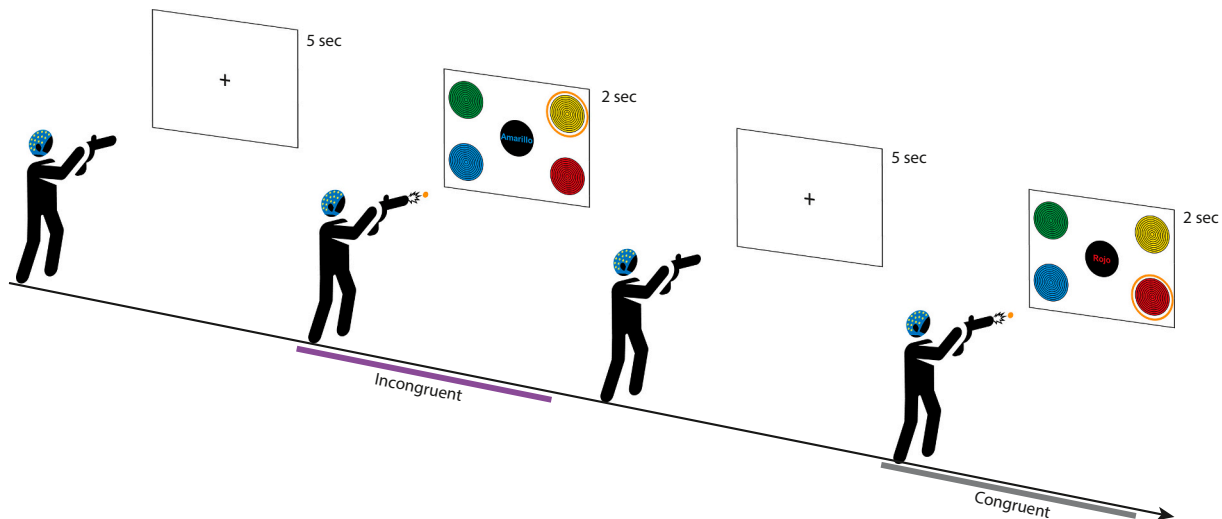
Fig. 2. Representation of the color-word reverse Stroop (RS) simulated handgun task. All the participants wore an EEG cap (represented in blue in the figure). A fixation cross was presented for 5 s, followed by the RS stimulus (2 s). The participants had to ignore the ink color of the color word stimuli and to shoot at the target of the color corresponding to the meaning of the word. The incongruent stimuli (left panels) presented a color word such as "yellow" (i.e., the Spanish word "amarillo") written in an incongruent ink color (e.g., blue). The congruent stimuli (right panels) presented a color word such as "red" (i.e., the Spanish word "rojo"), written in a congruent ink color (i.e., red). See also the video detailing the experimental procedure (see Video 1). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2

Summary of the demographic characteristics of the whole sample that performed the shooting task (left side of the table) and the subsample with an EEG recording (right side of the table). Means and standard deviations (in brackets) for the age (years), as well as the number and percentage (in brackets) of right-handed participants are separately reported for males and females in each sample.

Shooting task	Age M (SD)	Right-handed n (%)	EEG	Age M (SD)	Right-handed n (%)
Males n = 14	20.5 (3.3)	13 (92.9%)	Males n = 9	20.9 (3.8)	8 (88.9%)
Females n = 18	19.2 (2.4)	16 (88.9%)	Females n = 11	19.1 (2.7)	9 (81.8%)
Total n = 32	19.8 (2.8)	29 (90.6%)	Total n = 20	19.9 (3.3)	17 (85.0%)

participants. Individual differences between conditions (incongruent − congruent) ranged from − 19 to 522 milliseconds (ms). Mean RT for hits only (±SD) for incongruent stimuli (1213 ms ± 180 ms) was again higher than for congruent ones (1003 ms ± 183 ms), $t(31) = -9.17$, $p < 0.001$, $d = 1.76$. Again, this trend was exhibited by 31 out of 32 participants. Individual differences between conditions (incongruent − congruent) ranged from − 43 ms to 520 ms.

The ICC estimates for the hit rate and shooting accuracy were 0.969 [95% C.I. 0.965 − 0.972] and 0.937 [95% C.I. 0.930 − 0.943] respectively, which indicate excellent reliability between evaluators. Therefore, we employed the mean value of the two evaluators in the subsequent analyses over these two dependent variables.

Participants showed a significantly higher hit rate (arcsine-transformed) for the congruent (mean ± SD: 86.78% ± 40.38%) than for the incongruent stimuli (mean ± SD: 70.04% ± 41.69%), $t(31) = 3.83$, $p = 0.001$, $d = 0.69$. This significant trend was exhibited by 29 out of 32 participants. Individual differences between conditions (incongruent − congruent) ranged from − 0.75 to 0.11 percentage points. Mean shooting accuracy was better for the congruent (mean ± SD: 7.27 ± 1.98 score) than for the incongruent stimuli (mean ± SD: 7.99 ± 2.12 score), $t(31) = -3.962$, $p < 0.001$, $d = 0.71$. This significant trend was exhibited by 22 out of 32 participants. Individual differences between conditions

(incongruent − congruent) ranged from − 0.68 to 3.54 score. Shooting performance results are summarized in Table 3.

3.2. EEG global field power

Concerning the stimulus-locked time-window, two subsequent parts of the signal showed significant differences between conditions (290–306 ms and 352–358 ms *after* the stimulus presentation), with higher values for the congruent stimuli (GFP amplitude ≈ 37 μV, $p < 0.001$ for the significant time-points) as compared to the incongruent ones (≈32 μV).

As for the response-locked time-window, a significant difference between conditions emerged in three subsequent parts of the signal (384–342, 332–323, and 315–305 ms *before* the shot), with higher values for the incongruent stimuli (GFP amplitude ≈ 40 μV, $p = 0.01$ for the significant time-points) as compared to the congruent ones (≈35 μV) (Fig. 3).

Overall results were confirmed by parametric analysis (2 × 2 repeated measures ANOVA on the GFP mean amplitude values in the two time-windows; see Supplementary Material, Table S5 and Fig. S5).

Table 3

Summary of the shooting performance results for the 32 participants. Means and standard deviations (in brackets) for overall and hits only response times in milliseconds (ms), hit rate (percentage of hits, not transformed), and shooting accuracy (distance from the center of the target, ranging from 1 [best accuracy] to 10 [worst accuracy], calculated for the hits only) for both the congruent condition (n = 36 stimuli) and the incongruent condition (n = 12 stimuli).

	Congruent stimuli M (SD)	Incongruent stimuli M (SD)
Overall response times (ms)**	1004 (185)	1225 (190)
Hits only response times (ms)**	1003 (183)	1213 (180)
Hit rate (%)*	70.33 (26.16)	59.14 (31.28)
Shooting accuracy (score)**	7.27 (1.98)	7.99 (2.12)

* $p < 0.05$.
** $p < 0.001$.
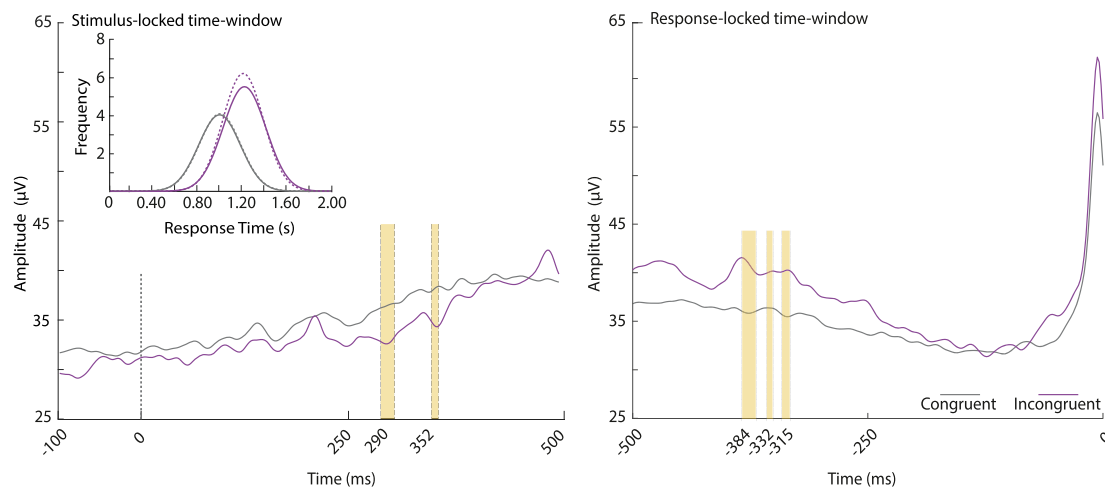
Fig. 3. The global field power (GFP) time course in the stimulus-locked (left panel) and response-locked (right panel) time-windows for congruent (gray lines) and incongruent (purple lines) stimuli. In the left panel, the 0 value represents the stimulus onset, whereas in the right panel it represents the moment of the participants' response (i.e., the shot). In the left panel, the −100 value represents the beginning of the baseline. The shaded yellowish areas in the GFP plots highlight the parts of the signal in which significant differences between conditions emerged ($n = 20$). Inset: the mean overall response time (RT) distributions for the two conditions ($n = 32$). The solid lines represent the overall RT distributions; the dotted lines represent the RT distributions for hits only. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4. Discussion

The color-word RS effect and its underlying processes are still an uncertain puzzle for cognitive researchers. Since there is evidence that the RS effect can be found even when a translation between the target and the specific response modality may not be needed (see Table 1 for a review), we hypothesized that target and the specific response modality requested of the participants might be central to the interference causing the RS effect (see Blais and Besner, 2006, 2007). Therefore, the response modality would be important in terms of the amount of practice the participants have developed in making the response. If a RS task requires a verbal response, as participants have a lot of practice reading words, a RS effect would not be expected. However, an unfamiliar (less practiced) and complex motor response would produce an interference, and such interference would be related to a late processing of the color (response conflict). We replicated the experimental setup by Wood and colleagues who used a visuomotor task with good ecological validity: a RS simulated handgun task, with participants with no shooting experience (Wood et al., 2016). They introduced a "high-threat" experimental condition (induced pressure to perform) and therefore the RS effect might be influenced by the participants' state anxiety. This condition was not present in our experimental design and therefore a comparison of results is not straightforward. We expected to find a clear RS effect with higher mean RTs, a lower mean hit rate, and worse mean shooting accuracy for the incongruent stimuli. Results from our experiment confirmed these expectations. For the first time, we reported a clear RS effect when participants were engaged in a task requiring a complex and unfamiliar visuomotor response (shooting at a target). These findings can, in part, reconcile the inconsistencies in previous results reported by studies employing motor, but relatively familiar, response modalities, such as pressing a specific key on a keyboard, or selecting a colored patch with a mouse (e.g., Blais and Besner, 2006; Durgin, 2000, 2003; Durgin et al., 2008; but see Blais and Besner, 2007; Yamamoto et al., 2016). It is possible that a hypothetical translation model might not be sufficient to explain those results. On the other hand, a model that puts the emphasis on the specific response modality and the amount of practice would indeed account for them, as unfamiliarity with the response modality would give color an opportunity to interfere with the response to a word (Blais and Besner, 2007).

Trying to ascertain whether the interference is related to a late processing of the color (response conflict, as proposed by Blais and

Besner, 2006, 2007), we calculated the GFP in two specific time-windows, a stimulus-locked window and a response-locked window. We expected a GFP modulation (i.e., amplitude) depending on the stimulus' congruency, reflecting this interference. According to our working hypothesis, the modulation should have been present in both time-windows, but specially in the response-locked window, reflecting differences not only between the processing requested by the two types of stimuli, but also between early stimulus processing (conflict detection) and response preparation (conflict resolution).

Regarding the stimulus-locked window, we found a significant GFP amplitude difference between congruent and incongruent stimuli, in two close time-windows between 290 and 358 ms after the stimulus (68 ms in total). In both time-windows, the GFP amplitude was significantly lower for the incongruent stimuli. This timing corresponds to that of the P3 ERP component family, generally related to early attentional stimulus processing and occasionally assessed in Stroop-based tasks (e.g., Atkinson et al., 2003; Kray et al., 2005; Szűcs et al., 2009; Warren and Marsh, 1979; Zurrón et al., 2009). Whereas the RS studies have reported unchanged P3 amplitudes between congruent and incongruent stimuli (Atkinson et al., 2003; Kray et al., 2005; Warren and Marsh, 1979), other Stroop studies have reported higher (Szűcs et al., 2009) and lower (Zurrón et al., 2009) amplitudes for the incongruent stimuli. Our results indicate the presence of a significant GFP amplitude modulation for incongruent stimuli in a time-window corresponding to the P3 ERP component during a RS task. Since a reduced amplitude in this time window is usually interpreted as negatively correlated with the stimulus processing (Zurrón et al., 2009), the reduced amplitude in the GFP may reflect the difficulty in stimulus processing for the incongruent stimuli. The absence of this effect in previous studies employing RS tasks (Atkinson et al., 2003; Kray et al., 2005; Warren and Marsh, 1979) may be explained by response modality issues (key pressure is a more familiar activity). This could lead to a reduction in the interference produced by incongruent stimuli, thus simplifying their early attentional processing.

Regarding the response-locked window, our results indicate that the incongruent stimuli elicit a higher GFP amplitude in three adjacent time-windows between 384 and 305 ms *before* the shot (79 ms in total). This timing may correspond to the N400/N450 components or to the so-called Late Sustained Potential (LSP), as investigated in both Stroop and RS tasks (Atkinson et al., 2003; Kray et al., 2005; Szűcs et al., 2009; Szűcs and Soltész, 2010; Zurrón et al., 2009) that can show a latency up

to more than 600 ms after the stimulus (Atkinson et al., 2003). These components, related to the interference elicited by the incongruent stimuli (Atkinson et al., 2003; Badzakova-Trajkov et al., 2009; Kray et al., 2005; Szűcs et al., 2009; Szűcs and Soltész, 2010; Zurrón et al., 2009), are thought to reflect the activity of the Anterior Cingulate Cortex (ACC). Among other processes, the ACC might be involved in the solution of this interference (Szűcs et al., 2009), as fMRI studies have demonstrated (see, among others, Huang et al., 2020; Song and Hakoda, 2015). Therefore, the increase in the GFP amplitude could reflect an increase in the ACC activation, indicating that a higher number of resources (i.e., attentional control; Durgin, 2000; Kane and Engle, 2003; Kane et al., 2007; Wood et al., 2016) is needed to solve the interference elicited by incongruent stimuli.

Overall, our results allow starting drawing a consistent framework for the processes underpinning the RS effect. First, they show the presence of a strong RS effect, in terms of both RTs and accuracy, when a complex and unfamiliar response modality is requested. Second, the GFP analysis in the two time-windows allows assessing, with a good time resolution, the electrophysiological correlates of the interference produced by the incongruent stimuli over time, giving us the chance to start unveiling the underlying processes. Indeed, the initial reduction in the GFP amplitude for the incongruent stimuli may reflect participants' struggle to process them. This result is in line with the so-called translational explanation (Durgin, 2000) by which a greater effort in processing incongruent stimuli is expected when a visuomotor response is requested due to the need to translate a word meaning into a visual code. The subsequent GFP amplitude increase for incongruent stimuli may mirror the effort of participants' attentional control (Durgin, 2000; Kane and Engle, 2003; Kane et al., 2007; Wood et al., 2016) to manage the interference raised by incongruent stimuli up to its solution. Our results seem to suggest that this interference is the result of both conflict detection and conflict resolution processes, and therefore both translation and strength of association hypotheses would be compatible. The absence of a comparison with a more familiar task (e.g., pressing a key) or of a direct assessment of the possible underlying association processes (e.g., through practice or thanks to experience) does not allow to clearly disentangle the two hypotheses, nor to clearly assess the role of the response modality. Therefore, any interpretation regarding the role of response modality and of the two hypotheses (translation and strength of association) should be considered as speculative. An important contribution may come from future studies analyzing experts' performance and GFP to verify the presence of modulations in conflict resolution (response-locked window) and unveil the pertinence of a strength of association hypothesis to explain motor-familiar responses. Notably, an increase in attentional control before the shot has been reported in experienced shooters, independently from the features of the task (Doppelmayr et al., 2008; Lawton et al., 1998; Loze et al., 2001; Pereira et al., 2018).

One may wonder if the results we present here are attributable to the unbalanced number of congruent and incongruent stimuli. It is worth noting the imbalance primary relies on the structure of the RS task itself, which requires a difference in the proportion of congruent and incongruent stimuli to induce an optimal degree of interference (Kane and Engle, 2003). One criticism regarding the use of this proportion can be the impossibility to rule out that the slower and less accurate responses on incongruent trials are attributable to their infrequent presentation, due to the use of the congruent trials as a baseline for measuring the RS effect. However, previous studies (Kane and Engle, 2003) showed that this can be regarded as true only when (1) the participants have reported a high rate of cognitive failures (which we ruled out with the CFQ questionnaire) and (2) they have previously faced task conditions with different proportions of stimuli. In any case, future studies may try to replicate the present results using different proportions of stimuli and including also a neutral condition in order to rule out any potential effect of stimuli imbalance. Also, the unbalanced number of clean EEG epochs between congruent and incongruent stimuli included in the

analysis may be thought to affect the GFP results. This imbalance also depends on a higher number of epochs discarded due to artifacts in the incongruent stimuli. This may be the effect of a reorientation movement of the plastic gun toward the correct target, elicited by the interference produced by the incongruent stimuli, (as suggested by the eye movements' results reported by Wood et al., 2016). However, this possibility seems unlikely in light of the results obtained from an additional analysis based on a random resampling of the congruent trials (see Supplementary Fig. S6 and Table S6). Once the number of congruent trials was balanced with that of the incongruent ones, the results were completely comparable to those reported above, both in stimulus-locked and response-locked time-window. Independently from the reliability of our results, one solution to prevent any possible effect of this imbalance may be the overall increase in the number of trials (which, in the present work, are 48), but this would also cause a fatigue effect due to the participants' shooting inexperience (more prone to physical and mental fatigue). For the same reason, a more classical EEG metric, such as ERPs and associated ERP topographic maps, would be hardly ideal in this context due to the high number of trials they require (hundreds; see Luck, 2014). In any case, it is worth noting that the GFP metric and waveform reflects the underlying ERPs (see Szűcs and Soltész, 2012 for an applied example in a Stroop task), indicating how strong a potential is being recorded on average across electrodes montage (Murray et al., 2008). On the other hand, there is no clear consensus over the reliability of other common EEG metrics, such as power spectra, for tasks involving epochs lasting less than 20 s (e.g., Gasser et al., 1985; Gudmundsson et al., 2007; but see Salinsky et al., 1991) and, to date, no reliable evidence has been reported regarding their appropriateness for measuring control-related aspects in motor tasks (see Parr et al., 2021).

Notwithstanding the above, our results should be viewed in the context of three main shortcomings related to the experimental procedure we designed and the instruments we used. First, the high variation in the number of clear and interpolated channels, as well as of the location of removed channels across participants may have had an effect on the GFP results. This should be accounted by the specific experimental design (i.e., within-participants design), and the overall common spatial distribution of the channels included in the GFP calculation (i.e., 70% of the channels corresponding to frontal and central scalp positions, see Supplementary Table S2).

The second concern is that the unfamiliarity with the task would prevent the distinction between shooting errors (i.e., a miss due to the participants not being accurate in shooting intended targets, because of shooting inexperience) and selection errors (i.e., a miss due to participants' choosing the incorrect response). However, since all the participants were inexperienced, the shooting errors, if present, should have been equally distributed in both congruent and incongruent stimuli. Consequently, the effect we are reporting can be regarded as a result of shooting errors.

The last concern is that we had to exclude 12 participants from the EEG analysis. This was because their signals were affected by the limitations of the high-impedance active-dry actiCAP Xpress Twist system. Indeed, the system proved to be particularly prone to electrode detachment in case of hairy scalps, and to increased signal noise during slight movements (such as those induced by the recoil of the gun). Overall, future studies should aim at deepening the present findings to build a more general explanation model, by increasing the sample size and employing a system able to reliably identify the EEG source localization of the effects.

4.1. Conclusions

The present study shows that the color-word RS effect is easily elicited in both RTs and shooting performance (hit rate and accuracy) when it requires an unfamiliar response, which supports the strength of association hypothesis. Moreover, the effect is paralleled by GFP modulations in both *early* and *late* time-windows. These results, although

with the abovementioned limitations, suggest that the interference producing the RS effect would occur early in time, at the stimulus processing stage, but also late in time, closer to the response.

Overall, our findings represent an important first step toward the systematization of the knowledge regarding color-word RS effect, supporting the idea that a classical effect, such as the RS, can be used in simple but innovative paradigms to disclose cognitive and electrophysiological processes, also yielding to a new research path for a variety of experimental and applied contexts, such as military and sport.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijpsycho.2022.02.006.

Data availability

Data will be made available on request.

Acknowledgements

References

Appelbaum, L.G., Boehler, C.N., Davis, L.A., Won, R.J., Woldorff, M.G., 2014. The dynamics of proactive and reactive cognitive control processes in the human brain. J. Cogn. Neurosci. 26, 1021–1038.

Atkinson, C.M., Drysdale, K.A., Fulham, W.R., 2003. Event-related potentials to stroop and reverse stroop stimuli. Int. J. Psychophysiol. 47, 1–21.

Badzakova-Trajkov, G., Barnett, K.J., Waldie, K.E., Kirk, I.J., 2009. An ERP investigation of the stroop task: the role of the cingulate in attentional allocation and conflict resolution. Brain Res. 1253, 139–148.

Blair, R.C., Karniski, W., 1994. Distribution-free statistical analyses of surface and volumetric maps. In: Thatcher, R.W., Hallett, M., John, E.R., Huerta, M. (Eds.), Functional Neuroimaging: Technical Foundations. Academic Press, pp. 19–28.

Blais, C., Besner, D., 2006. Reverse stroop effects with untranslated responses. J. Exp. Psychol. Hum. Percept. Perform. 32, 1345–1353.

Blais, C., Besner, D., 2007. A reverse stroop effect without translation or reading difficulty. Psychon. Bull. Rev. 14, 466–469.

Broadbent, D.E., Cooper, P.F., FitzGerald, P., Parkes, K.R., 1982. The cognitive failures questionnaire (CFQ) and its correlates. Br. J. Clin. Psychol. 21, 1–16.

Cacioppo, J.T., Tassinary, L.G., Berntson, G. (Eds.), 2007. Handbook of Psychophysiology. Cambridge University Press.

Clark, J.H., 1924. The ishihara test for color blindness. Am. J. Physiol. Opt. 5, 269–276.

Clayson, P.E., Miller, G.A., 2017. ERP reliability analysis (ERA) toolbox: an open-source toolbox for analyzing the reliability of event-related brain potentials. Int. J. Psychophysiol. 111, 68–79.

Compton, R.J., Heaton, E.C., Gaines, A., 2018. Is attention enhanced following performance errors? Testing the adaptive control hypothesis. Psychophysiology 55, e13022.

Connor, J., Norton, R., Ameratunga, S., Robinson, E., Civil, I., Dunn, R., Bailey, J., Jackson, R., 2002. Driver sleepiness and risk of serious injury to car occupants: a population-based case control study. Br. Med. J. 324, 1125.

Danziger, S., Estévez, A.F., Marí-Beffa, P., 2002. Stroop interference effects in partially colored stroop words. Psychon. Bull. Rev. 9, 536–541.

Davies, I.R.L., Corbett, G.G., Laws, G., McGurk, H., Moss, A.S.G., Smith, M.W., 1991. Linguistic basicness and colour information processing. Int. J. Psychol. 26, 311–327.

Delorme, A., Makeig, S., 2004. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21.

Diaz-Piedra, C., Rieiro, H., Di Stasi, L.L., 2021. Monitoring army drivers' workload during off-road missions: an experimental controlled field study. Saf. Sci. 134, 105092.

Doppelmayr, M., Finkenzeller, T., Sauseng, P., 2008. Frontal midline theta in the pre-shot phase of rifle shooting: differences between experts and novices. Neuropsychologia 46, 1463–1467.

Dulaney, C.L., Rogers, W.A., 1994. Mechanisms underlying reduction in stroop interference with practice for young and old adults. J. Exp. Psychol. Learn. Mem. Cogn. 20, 470–484.

Durgin, F.H., 2000. The reverse stroop effect. Psychon. Bull. Rev. 7, 121–125.

Durgin, F.H., 2003. Translation and competition among internal representations in a reverse stroop effect. Percept. Psychophys. 65, 367–378.

Durgin, F.H., Doyle, E., Egan, L., 2008. Upper-left gaze bias reveals competing search strategies in a reverse stroop task. Acta Psychol. 127, 428–448.

Elvevåg, B., Duncan, J., McKenna, P.J., 2000. The use of cognitive context in schizophrenia: an investigation. Psychol. Med. 30, 885–897.

Files, B.T., Lawhern, V.J., Ries, A.J., Marathe, A.R., 2016. A permutation test for unbalanced paired comparisons of global field power. Brain Topogr. 29, 345–357.

Gasser, T., Bächer, P., Steinberg, H., 1985. Test-retest reliability of spectral parameters of the EEG. Electroencephalogr. Clin. Neurophysiol. 60 (4), 312–319.

Groppe, D.M., Urbach, T.P., Kutas, M., 2011. Mass univariate analysis of event-related brain potentials/fields I: a critical tutorial review. Psychophysiology 48, 1711–1725.

Gudmundsson, S., Runarsson, T.P., Sigurdsson, S., Eiriksdottir, G., Johnsen, K., 2007. Reliability of quantitative EEG features. Clin. Neurophysiol. 118 (10), 2162–2171.

Heidlmayr, K., Kihlstedt, M., Isel, F., 2020. A review on the electroencephalography markers of stroop executive control processes. Brain Cogn. 146, 105637.

Hepp, H.H., Maier, S., Hermle, L., Spitzer, M., 1996. The stroop effect in schizophrenic patients. Schizophr. Res. 22, 187–195.

Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., Dement, W.C., 1973. Quantification of sleepiness: a new approach. Psychophysiology 10, 431–436.

Huang, Y., Su, L., Ma, Q., 2020. The stroop effect: an activation likelihood estimation meta-analysis in healthy young adults. Neurosci. Lett. 716, 134683.

Ikeda, Y., Hirata, S., Okuzumi, H., Kokubun, M., 2010. Features of stroop and reverse-stroop interference: analysis by response modality and evaluation. Percept. Mot. Skills 110, 654–660.

Ikeda, Y., Hirata, S., Maeda, W., Okuzumi, H., Kokubun, M., Ambe, H., Fujieda, Y., Suijo, K., Hayashi, J., Iteya, M., 2013. Relationship between stroop/reverse-stroop interference and postural sway when standing in elderly people. Compr. Psychol. 2, 22–25.

Ikeda, Y., Okuzumi, H., Kokubun, M., 2013b. stroop/reverse-stroop interference in typical development and its relation to symptoms of ADHD. Res. Dev. Disabil. 34, 2391–2398.

Jasper, H.H., 1958. The ten-twenty electrode system of the international federation. Electroencephalogr. Clin. Neurophysiol. 10, 370–375.

Kane, M.J., Brown, L.E., Little, J.C., Silvia, P.J., Myin-Germeys, I., Kwapil, T.R., 2007. For whom the mind wanders, and when: an experience-sampling study of working memory and executive control in daily life. Psychol. Sci. 18, 614–621.

Kane, M.J., Engle, R.W., 2003. Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to stroop interference. J. Exp. Psychol. Gen. 132, 47–70.

Kappenman, E.S., Luck, S.J., 2010. The effects of electrode impedance on data quality and statistical significance in ERP recordings. Psychophysiology 47, 888–904.

Kertzman, S., Lowengrub, K., Aizer, A., Nahum, Z.B., Kotler, M., Dannon, P.N., 2006. Stroop performance in pathological gamblers. Psychiatry Res. 142, 1–10.

Kray, J., Eppinger, B., Mecklinger, A., 2005. Age differences in attentional control: an event-related potential approach. Psychophysiology 42, 407–416.

Lawton, G.W., Hung, T.M., Saarela, P., Hatfield, B.D., 1998. Electroencephalography and mental states associated with elite performance. J. Sport Exerc. Psychol. 20, 35–53.

Lehmann, D., Skrandies, W., 1980. Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroencephalogr. Clin. Neurophysiol. 48, 609–621.

Lehmann, D., Skrandies, W., 1984. Spatial analysis of evoked potentials in man: a review. Prog. Neurobiol. 23, 227–250.

Logan, G.D., Zbrodoff, N.J., 1998. Stroop-type interference: congruity effects in color naming with typewritten responses. J. Exp. Psychol. Hum. Percept. Perform. 24, 978–992.

Loze, G.M., Collins, D., Holmes, P.S., 2001. Pre-shot EEG alpha-power reactivity during expert air-pistol shooting: a comparison of best and worst shots. J. Sports Sci. 19, 727–733.

Lu, C.H., Proctor, R.W., 2001. Influence of irrelevant information on human performance: effects of SR association strength and relative timing. Q. J. Exp. Psychol. A Hum. Exp. Psychol. 54, 95–136.

Luck, S.J., 2014. An Introduction to the Event-related Potential Technique. MIT press, Cambridge: Massachusetts.

MacLeod, C.M., 1991. Half a century of research on the stroop effect: an integrative review. Psychol. Bull. 109, 163–203.

Mathewson, K.E., Harrison, T.J., Kizuk, S.A., 2017. High and dry? Comparing active dry EEG electrodes to active and passive wet electrodes. Psychophysiology 54, 74–82.

Matsumoto, A., Hakoda, Y., Watanabe, M., 2012. Life-span development of stroop and reverse-stroop interference measured using matching responses. Jpn. J. Psychol. 83, 337–346.

Melara, R.D., Mounts, J.R., 1993. Selective attention to stroop dimensions: effects of baseline discriminability, response mode, and practice. Mem. Cogn. 21 (5), 627–645.

Miller, H.C., Kubicki, S., Caffier, D., Kolski, C., Naveteur, J., 2016. The stroop and reverse stroop effects as measured by an interactive tabletop. Int. J. Hum. Comput. Interact. 32, 363–372.

Moriguchi, K., Morikawa, Y., 1998. Time course analysis of the reverse-stroop effect in japanese kanji. Percept. Mot. Skills 87, 163–174.

Mullen, T., 2012. CleanLine EEGLAB plugin. In: Neuroimaging Informatics Tools and Resources Clearinghouse (NITRC). San Diego, CA.

Murray, M.M., Brunet, D., Michel, C.M., 2008. Topographic ERP analyses: a step-by-step tutorial review. Brain Topogr. 20, 249–264.

Nagamine, T., 2017. The reverse-stroop test predicts self-discipline of the money of schizophrenia patients. Int. Med. J. 24, 428–429.

Nishimura, R., 2006. The asymmetry of interhemispheric interaction in reverse-stroop task. Jpn. J. Psychol. 77, 163–169.

Pereira, M., Argelaguet, F., Millán, J.D.R., Lécuyer, A., 2018. Novice shooters with lower pre-shooting alpha power have better performance during competition in a virtual reality scenario. Front. Psychol. 9, 527.

Parr, J.V., Gallicchio, G., Wood, G., 2021. EEG correlates of verbal and conscious processing of motor control in sport and human movement: a systematic review. Int. Rev. Sport Exerc. Psychol. 1–32 https://doi.org/10.1080/1750984X.2021.1878548.

Peru, A., Faccioli, C., Tassinari, G., 2006. Stroop effects from 3 to 10 years: the critical role of reading acquisition. Arch. Ital. Biol. 144, 45–62.

Roelofs, A., 2005. The visual-auditory color-word stroop asymmetry and its time course. Mem. Cogn. 33, 1325–1336.

Rosenberg, C., Strohl, K., 2018. The reverse stroop test applied in the sleep clinic. Sleep 41, A226.

Ruff, C.C., Woodward, T.S., Laurens, K.R., Liddle, P.F., 2001. The role of the anterior cingulate cortex in conflict processing: evidence from reverse stroop interference. NeuroImage 14, 1150–1158.

Salinsky, M.C., Oken, B.S., Morehead, L., 1991. Test-retest reliability in EEG frequency analysis. Electroencephalogr. Clin. Neurophysiol. 79 (5), 382–392.

Sasaki, M., Hakoda, Y., Yamagami, R., 1993. Schizophrenia and reverse-stroop interference in the group version of the stroop and reverse-stroop test. Jpn. J. Psychol. 64, 43–50.

Shalev, L., Algom, D., 2000. Stroop and Garner effects in and out of Posner's beam: reconciling two conceptions of selective attention. J. Exp. Psychol. Hum. Percept. Perform. 26, 997–1017.

Simon, J.R., Baker, K.L., 1995. Effect of irrelevant information on the time to enter and retrieve relevant information in a stroop-type task. J. Exp. Psychol. Hum. Percept. Perform. 21, 1028–1043.

Skrandies, W., 1990. Global field power and topographic similarity. Brain Topogr. 3, 137–141.

Smithson, H.E., Khan, S.S., Sharpe, L.T., Stockman, A., 2006. Transitions between color categories mapped with a reverse stroop task. Vis. Neurosci. 23, 453–460.

Sobel, K.V., Puri, A.M., York, A.K., 2020. Visual search inverts the classic stroop asymmetry. Acta Psychol. 205, 103054.

Song, Y., Hakoda, Y., 2015. An fMRI study of the functional mechanisms of stroop/reverse-stroop effects. Behav. Brain Res. 290, 187–196.

Steinhauser, M., Hübner, R., 2009. Distinguishing response conflict and task conflict in the stroop task: evidence from ex-gaussian distribution analysis. J. Exp. Psychol. Hum. Percept. Perform. 35, 1398.

Stroop, J.R., 1935. Studies of interference in serial verbal reactions. J. Exp. Psychol. 18, 643–662.

Sugg, M.J., McDonald, J.E., 1994. Time course of inhibition in color-response and word-response versions of the stroop task. J. Exp. Psychol. Hum. Percept. Perform. 20, 647–675.

Szűcs, D., Soltész, F., White, S., 2009. Motor conflict in stroop tasks: direct evidence from single-trial electro-myography and electro-encephalography. NeuroImage 47, 1960–1973.

Szűcs, D., Soltész, F., 2010. Stimulus and response conflict in the color–word stroop task: a combined electro-myography and event-related potential study. Brain Res. 1325, 63–76.

Szűcs, D., Soltész, F., 2012. Functional definition of the N450 event-related brain potential marker of conflict processing: a numerical stroop study. BMC Neurosci. 13, 1–14.

Takeuchi, H., Taki, Y., Sassa, Y., Hashizume, H., Sekiguchi, A., Nagase, T., Nouchi, R., Fukushima, A., Kawashima, R., 2012. Regional gray and white matter volume associated with stroop interference: evidence from voxel-based morphometry. NeuroImage 59, 2899–2907.

Takeuchi, H., Taki, Y., Nouchi, R., Sekiguchi, A., Hashizume, H., Sassa, Y., Kotozaki, Y., Miyauchi, C.M., Yokoyama, R., Iizuka, K., Nakagawa, S., Nagase, T., Kunitoki, K., Kawashima, R., 2015. Degree centrality and fractional amplitude of low-frequency oscillations associated with stroop interference. NeuroImage 119, 197–209.

Warren, L.R., Marsh, G.R., 1979. Changes in event related potentials during processing of stroop stimuli. Int. J. Neurosci. 9, 217–223.

Washburn, D.A., 2016. The stroop effect at 80: the competition between stimulus control and cognitive control. J. Exp. Anal. Behav. 105 (1), 3–13.

Watanabe, M., Hakoda, Y., Matsumoto, A., 2013. Attentional control and task-switching costs investigated by the group version of the stroop tasks. Jpn. J. Psychol. 84, 64–68.

Westfall, P.H., Young, S.S., 1993. Resampling-based Multiple Testing: Examples and Methods for p-value Adjustment. John Wiley & Sons.

World Medical Association – W.M.A., 1964. Human experimentation. Code of ethics of the W.M.A. Br. Med. J. 2, 177.

Wood, G., Vine, S.J., Wilson, M.R., 2016. Working memory capacity, controlled attention and aiming performance under pressure. Psychol. Res. 80, 510–517.

Woodward, T.S., Bub, D.N., Hunter, M.A., 2002. Task switching deficits associated with Parkinson's disease reflect depleted attentional resources. Neuropsychologia 40, 1948–1955.

Woodward, T.S., Ruff, C.C., Ngan, E.T., 2006. Short-and long-term changes in anterior cingulate activation during resolution of task-set competition. Brain Res. 1068, 161–169.

Yamamoto, N., Incera, S., McLennan, C.T., 2016. A reverse stroop task with mouse tracking. Front. Psychol. 7, 670.

Yasumura, A., Kokubo, N., Yamamoto, H., Yasumura, Y., Nakagawa, E., Kaga, M., Hiraki, K., Inagaki, M., 2014. Neurobehavioral and hemodynamic evaluation of stroop and reverse stroop interference in children with attention-deficit/hyperactivity disorder. Brain Dev. 36, 97–106.

Yasumura, A., Omori, M., Fukuda, A., Takahashi, J., Yasumura, Y., Nakagawa, E., Koike, T., Yamashita, Y., Miyajima, T., Koeda, T., Aihara, M., Inagaki, M., 2019. Age-related differences in frontal lobe function in children with ADHD. Brain Dev. 41, 577–586.

Zimmer, P., Stritt, C., Bloch, W., Schmidt, F.P., Hübner, S.T., Binnebößel, S., Schenk, A., Oberste, M., 2016. The effects of different aerobic exercise intensities on serum serotonin concentrations and their association with stroop task performance: a randomized controlled trial. Eur. J. Appl. Physiol. 116, 2025–2034.

Zurrón, M., Pouso, M., Lindín, M., Galdo, S., Díaz, F., 2009. Event-related potentials with the stroop colour-word task: timing of semantic conflict. Int. J. Psychophysiol. 72, 246–252.