



## Reliability of the emotional Stroop task: An investigation of patients with panic disorder

T. Dresler<sup>a,b,\*</sup>, A.-C. Ehlis<sup>a</sup>, C. Hindi Attar<sup>c</sup>, L.H. Ernst<sup>a</sup>, S.V. Tupak<sup>b</sup>, T. Hahn<sup>d</sup>, B. Warrings<sup>b</sup>, F. Markulin<sup>b</sup>, C. Spitzer<sup>e,f</sup>, B. Löwe<sup>e</sup>, J. Deckert<sup>b</sup>, A.J. Fallgatter<sup>a</sup>

<sup>a</sup> Department of Psychiatry and Psychotherapy, University of Tuebingen, Germany

<sup>b</sup> Department of Psychiatry, Psychosomatics and Psychotherapy, University Hospital Wuerzburg, Germany

<sup>c</sup> Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

<sup>d</sup> Department of Cognitive Psychology II, Johann Wolfgang Goethe University Frankfurt/Main, Germany

<sup>e</sup> Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

<sup>f</sup> Asklepios Fachklinikum Tiefenbrunn, Asklepios Psychiatry Niedersachsen, Germany

### ARTICLE INFO

#### Article history:

Received 3 April 2012

Received in revised form

22 May 2012

Accepted 13 June 2012

#### Keywords:

Emotional Stroop task

Reliability

Panic disorder

### ABSTRACT

Despite its popularity in clinical research, the emotional Stroop task's reliability in patient groups is unknown. Given the low reliability of interference scores in healthy subjects, correlations with other variables pose a problem, especially as reliability in clinical samples is unknown. To assess reliability in panic disorder for the first time, we used the split-half method in two independent samples of patients and controls. As expected, only patients showed the behavioral interference effect. Reliability of interference scores (i.e. mean response latency emotional *minus* neutral words) was insufficiently low for patient and control samples; however, reliability scores derived from the conditions' response latencies (i.e. mean response latency emotional *or* neutral words) were much higher. The assumption that reliability scores in patients might differ from controls was not supported. This finding questions the use of correlations with external variables and suggests the use of response latencies instead of interference scores.

© 2012 Elsevier Ltd. All rights reserved.

### 1. Introduction

The emotional Stroop paradigm is a widely used test, both in basic and clinical research (Williams et al., 1996). It assesses the impact of emotional stimuli on attentional processes, with prolonged response latencies for naming the ink color of emotional compared to neutral words indicating so-called emotional interference<sup>1</sup> (i.e. mean response latency emotional *minus* neutral words). Such emotional interference effects have repeatedly been

found in various patient groups and anxious healthy subjects (cf. Bar-Haim et al., 2007; Williams et al., 1996). Despite the emotional Stroop task's popularity in research, the reliability especially for the interference scores has sparsely been investigated in healthy subjects, whereas for clinical populations no data are available. Kindt et al. (1996) were among the first who tested the reliability of the emotional Stroop task in healthy subjects. The authors reported that the 3-month test–retest reliability was insufficiently low for the interference scores in a card and a single trial<sup>2</sup> spider Stroop task (~0.2) which led the authors to assume these scores to be a rather unstable phenomenon. A Stroop task with taboo and self-relevant words (Siegrist, 1997) also yielded low test–retest reliabilities (~–0.1 to 0.1) in healthy subjects. In both studies the response latencies for each condition showed much higher reliability across time. Such a pattern of low reliability of the interference score (~–0.2 to 0.2) and high reliability of the response latencies (i.e. mean response latency emotional *or* neutral words;

\* Corresponding author. Psychophysiology and Optical Imaging, Department of Psychiatry and Psychotherapy, University of Tuebingen, Calwerstr. 14, D-72076 Tuebingen, Germany. Tel.: +49 7071 29 82627; fax: +49 7071 29 5379.

E-mail address: [Thomas.Dresler@med.uni-tuebingen.de](mailto:Thomas.Dresler@med.uni-tuebingen.de) (T. Dresler).

<sup>1</sup> Although the term emotional interference is widely used and will be used throughout this manuscript, the emotional Stroop task actually does not assess interference per se. This stems from the classical Stroop task in which color words are semantically related to the ink color resulting in interference when word and ink color are incongruent. In the emotional Stroop task the word content is neither semantically related to the ink color, nor do the emotional words imply responses that directly compete with the correct response. Given this, the task measures in how far attention is withdrawn from the required task by the emotional stimuli.

<sup>2</sup> In the card version a card contains more than one word (i.e. here spider-specific words), whereas in the single trial version each trial is presented separately in a successive manner.

>0.80) has been replicated by more recent studies (Eide et al., 2002; Strauss et al., 2005). These reliability findings only apply to healthy subjects. However, the emotional Stroop task is frequently used in different clinical populations assuming stable interference effects and it is “important for researchers to determine how these interference reliability scores compare with scores in clinical populations, especially as the ES [emotional Stroop] task is considered an important tool in determining psychopathology in depressed and anxious individuals” (Eide et al., 2002, p. 519, see also Strauss et al., 2005, p. 336). From this perspective, it is necessary to evaluate the emotional Stroop task’s reliability in patient samples, especially when emotional interference effects are supposed to reflect stable subject characteristics that might reflect symptom severity or might be changed by different therapeutic interventions. Whenever the emotional Stroop task is used in applied research in patients as compared to the more basic research in healthy subjects, information on reliability of different task indices is essential. Disregarding reliability in such cases may lead to wrong conclusions.

Therefore, we investigated the reliability scores of emotional interference in a clinical population (i.e. panic disorder) using two independent samples. This study has the potential to indicate in how far findings from healthy samples can or cannot be transferred to patients with panic disorder. Since the emotional Stroop task was only applied once, we used split-half reliability scores. In a previous study (Dresler et al., 2009) split-half reliability has been found to be higher than the usually reported test–retest reliability scores. The present data have been collected within the context of two functional magnetic resonance imaging (fMRI) studies that will be published elsewhere.

## 2. Methods

### 2.1. Participants

Both patient samples were recruited from in- and outpatient centers at the University Hospital Wuerzburg (Department of Psychiatry, Psychosomatics and Psychotherapy,  $n = 17$ , sample 1) and the University of Hamburg (Department of Psychosomatic Medicine and Psychotherapy,  $n = 20$ , sample 2). Healthy controls ( $n = 26/23$ ) were recruited via advertisements and internal data bases in Wuerzburg and Hamburg. Patients had panic disorder with or without agoraphobia as confirmed by a structured clinical interview (SKID, Wittchen et al., 1997) or standardized clinical interview (DIA-X, Wittchen and Pfister, 1997). Comorbid disorders included depression, other anxiety disorders and specific somatoform disorders that are usually associated with panic disorder. Sample characteristics are given in Tables 1 and 2. In sample 2 four patients showed remitted or partially remitted panic disorder, two patients were non-native speakers, but were fluent in German language. All participants were of Caucasian origin. 11 patients in sample 1 and 12 patients in sample 2 were on – mostly antidepressant – medication. Medication profiles of sample 1 and sample 2 are given in Tables 1 and 2. Low-potency antipsychotics and benzodiazepines were not taken before the experiment. The study was approved by the ethics committees of the University of Wuerzburg and the Medical Board in Hamburg and was in compliance with the latest declaration of Helsinki from 2008.

### 2.2. Stimuli and task

The emotional Stroop task comprised one neutral (e.g., oat, coil, paper) and one panic-related (e.g., attack, catastrophe, emergency) set of words selected from previous related studies (e.g., Dresler et al., 2009; Compton et al., 2003; Isenberg et al., 1999; McKenna

**Table 1**  
Sample 1 characteristics.

| Variable                             | Patients ( $n = 17$ )   | Controls ( $n = 26$ ) | Statistics                          |
|--------------------------------------|---|-----------------------|-------------------------------------|
| age (years)                          | 40.00 ± 11.57   | 37.77 ± 9.00          | $t_{39} = 0.712$ ,<br>$p = 0.480$   |
| gender (female/male)                 | 10/7  | 12/14                 | $\chi^2_1 = 0.660$ ,<br>$p = 0.416$ |
| handedness (right/left/ambidextrous) | 13/2/2  | 24/2/–                | $p = 0.148^a$                       |
| medication profile <sup>b</sup>      | 7 SSRI<br>1 SNRI<br>4 TCA<br>1 low-potential antipsychotics<br>2 anticonvulsants<br>5 benzodiazepines | –                     |                                     |

<sup>a</sup> Freeman–Halton extension of the exact Fisher-test.

<sup>b</sup> SSRI = selective serotonin reuptake inhibitors, SNRI = serotonin–norepinephrine reuptake inhibitors, TCA = tricyclic or tetracyclic antidepressants.

and Sharma, 1995; McNally et al., 1990, 1994, 1992; Whalen et al., 1998). Each condition comprised 15 words; there were no differences between conditions regarding the number of letters ( $t_{28} = 0.442$ ,  $p = 0.662$ ), numbers of syllables ( $t_{28} = 0.487$ ,  $p = 0.630$ ), and frequency in spoken ( $t_{28} = 0.547$ ,  $p = 0.589$ ) and written language ( $t_{28} = 0.771$ ;  $p = 0.446$ ) according to the Celex database (Baayen et al., 1995). Panic-related words were rated more panic-relevant ( $t_{10} = 30.232$ ;  $p < 0.001$ ), more body-related ( $t_{10} = 25.662$ ;  $p < 0.001$ ), but similar in familiarity ( $t_{10} = 1.312$ ;  $p = 0.219$ ) according to an evaluation by 11 people working at the University Hospital Wuerzburg. Words were presented visually against a black background via goggles in one sample and a tilted mirror that was mounted to the MRI head coil in the other sample. The experiment was set up using customized experimental control software (Presentation, Neurobehavioral Systems Inc., Albany, CA, USA, <http://www.neurobs.com>). In one sample a block design was chosen (with an additional condition only showing white fixation crosses not requiring any response), in the other sample the design was event-related. Subjects had to indicate word ink color by pressing buttons using MRI-compatible button boxes. The assignment of colors to buttons was counterbalanced. Both experiments comprised one run. Design characteristics are given in Table 3.

In the block design (sample 1), words were separated by a black screen of the same duration (1500 ms). Blocks comprised ten words, in an additional condition a block comprised fixation crosses not requiring responses; there were no two blocks of the

**Table 2**  
Sample 2 characteristics.

| Variable                             | Patients ( $n = 20$ )   | Controls ( $n = 23$ ) | Statistics                          |
|--------------------------------------|---|-----------------------|-------------------------------------|
| age (years)                          | 31.74 ± 7.41  | 31.35 ± 7.99          | $t_{41} = 0.166$ ,<br>$p = 0.869$   |
| gender (female/male)                 | 11/9  | 12/11                 | $\chi^2_1 = 0.034$ ,<br>$p = 0.853$ |
| handedness (right/left/ambidextrous) | 20/–/–  | 23/–/–                | –                                   |
| medication profile <sup>a</sup>      | 6 SSRI<br>3 SNRI<br>3 TCA<br>7 low-potential antipsychotics<br>1 atypical antipsychotic<br>2 benzodiazepines<br>1 non-benzodiazepine hypnotic agent | –                     |                                     |

<sup>a</sup> SSRI = selective serotonin reuptake inhibitors, SNRI = serotonin–norepinephrine reuptake inhibitors, TCA = tricyclic or tetracyclic antidepressants.

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات