International Journal of Psychophysiology xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

International Journal of Psychophysiology

journal homepage: www.elsevier.com/locate/ijpsycho

Cardiac reactivity and preserved performance under stress: Two sides of the same coin?

Nathalie Pattyn [a,b,c,d,*], Olivier Mairesse [a,e], Aisha Cortoos [a,f], José Morais [c], Eric Soetens [b], Bart Roelands [d], Annick van den Nest [a], Régine Kolinsky [c]

[a] VIPER Research Unit, Royal Military Academy, Renaissancelaan, 30, 1000 Brussel, Belgium
[b] Dept of Biological Psychology, Vrije Universiteit Brussel, Pleinlaan, 2, 1050 Brussel, Belgium
[c] Unité de Recherches en Neurosciences Cognitives, Université Libre de Bruxelles, CP 191, avenue F.D. Roosevelt, 50, 1050 Bruxelles, Belgium
[d] Department of Human Physiology, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel, Belgium
[e] CHU Brugmann, Sleep Unit, Belgium
[f] UZ Brussel, Sleep Unit, Belgium

ARTICLE INFO

Article history:
Received 10 July 2012
Received in revised form 11 March 2013
Accepted 12 March 2013
Available online xxxx

Keywords:
Stress
Performance
Cardiac reactivity
Stroop

ABSTRACT

In the present experiment, cognitive control under stress was investigated using a real-life paradigm, namely an evaluation flight for military student pilots. The magnitude of cognitive interference on color–word, numerical and emotional Stroop paradigms was studied during a baseline recording and right before the test flight. Cardio-respiratory parameters were simultaneously assessed during rest and the performance of the Stroop tasks. Cognitive data suggested a different speed/accuracy trade-off under stress, and no modulation of the interference effect for color words or numerical stimuli. However, we observed a major increase in error rates for specific emotional stimuli related to the evaluation situation in the stress condition. The increase in cognitive interference from emotional stimuli, expressed as an increase in error rates, was correlated to the decreased cardiac reactivity to challenge in the stress situation. This relationship is discussed in the framework of Sanders' (1983) model of stress and performance. In terms of future research, this link warrants a fruitful lead to be followed for investigating the causal mechanism of performance decrements under the influence of stress.

1. Introduction

Ever since Yerkes and Dodson (1908), the effects of stress on cognitive performance have been investigated in both experimental and applied psychology. However, the concepts stress and performance have too often been stocked with idiosyncratic or unspecific connotations, hampering comparison of the available results. Operationalization of stress or arousal in research paradigms indeed is quite complex. While questionnaires rely on subjective evaluation, physiological measures like heart rate variability quantify a systemic outcome. Operational definitions based on response have been prominent in psychophysiological literature (Sanders, 1983), ever since Selye (1956) introduced the concept of stress as the response of the body to any demand made upon it. Intervening variable-definitions (Cox, 1978, in Sanders, 1983) on the other hand, have become very influential in clinical and coping literature, whereas this conceptualization has been surprisingly absent in psychophysiological research on stress and performance. The two most influential contemporary stress and performance models did nonetheless underscore an intervening variable concept, termed 'resource recruitment' in Hockey's (1997) compensatory control regulation and 'effort' in Sanders' (1983) cognitive-energetical model. In order to allow inferences about mechanisms at play, psychophysiological investigations of stress and performance should frame the design and data-analysis within such models, which happens surprisingly seldom in the most recent research.

With regard to performance, as mentioned by Kofman et al. (2006), very few studies examined the effects of stress on executive functions. Indeed, considering the importance of stress in applied research, for example on aviation or traffic safety (e.g. Matthews et al., 1998), and considering the involvement of higher order cognitive functions such as planning, monitoring and cognitive control for an adequate performance in the aforementioned settings, there is a remarkable lack of experimental results on the effects of stress on executive functions. According to Matthews et al. (1997), stress mainly affects performance in applied settings through cognitive interference. In experimental paradigms investigating executive functions, distinct variants of the Stroop task (Stroop, 1935) are known to trigger this type of cognitive interference.

* Corresponding author at: VIPER Research Unit, Royal Military Academy, Renaissancelaan, 30, 1000 Brussel, Belgium. Tel.: +32 2 629 15 94, +32 476 66 12 03 (mobile); fax: +32 2 629 24 89.
E-mail address: nathalie.pattyn@rma.ac.be (N. Pattyn).

0167-8760/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.ijpsycho.2013.03.008

In the classical color–word Stroop task, subjects are instructed to name the color in which a word appears. Congruent stimuli display coherence between semantics and appearance (e.g. the word *red* in red), whereas incongruent stimuli do not (e.g. the word *red* in blue). The numerical Stroop variant (Pavese and Umiltà, 1998) requires participants to respond to the amount of characters presented onscreen; these can be either congruent (three times 3) or incongruent (four times 3). The response conflict elicited by incongruent stimuli materializes as longer reaction times (RTs) and lower accuracy (for a review, see MacLeod, 1991). The emotional Stroop variant includes (among the color words) negatively valenced emotional words, related to a particular individual's area of concern, causing interference as a result of attentional bias toward threat related expressions (McKenna and Sharma, 1995).

The emotional Stroop version therefore allows inferences about attentional biases, reflected as longer response latencies to name the ink color of emotional words as compared to neutral words. As Williams et al. (1997) summed up, emotional Stroop interference could be attributed to variance in state or trait emotion, to variance in the particular situation in which the task is performed, or to the specific nature of the negatively valenced words used. Impact of negative valence requires the presented material to be accustomed to both the subject's current concerns, and the situation, as in Ray's (1979) pioneering emotional Stroop challenge.

In addition to aforementioned Stroop interferences, effects from negative priming and inverse negative priming have been investigated as well. Negative priming refers to a slowed response time to a target stimulus that has been previously ignored (correct response for stimulus S is the inhibited response for item S-1, e.g. *yellow* presented in blue after *blue* presented in red), and involves inhibition of a mechanism of selective attention (Tipper, 1985). Inverse negative priming further adds the reciprocal variation of relevant and irrelevant dimension (*yellow* presented in blue after *blue* presented in yellow). Examining negative priming effects thus provides more insight in the quality of cognitive control for selecting relevant information.

Moreover, the Stroop tasks are among the most widely applied paradigms to elicit stress in laboratory conditions, for the purpose of investigating autonomic reactivity to mental stress (e.g. Akerstedt et al., 1983; Kamarck et al., 1994; Heims et al., 2006; Wright et al., 2007). In this line of research, reactivity is conceptualized as "an acute and relatively rapid change in a cardiovascular parameter as a function of the presentation of a stressor" (Hughdahl, 1995). This concept of reactivity could be applied to quantify the "recruitment of resources" from Hockey's model, or the "effort" from Sander's model. Kofman et al. (2006) emphasized that, since both the regulation of the autonomic stress response and the inhibition of prepotent responses activate common prefrontal cortical regions, particularly the anterior cingulate cortex (ACC), an interaction between these processes could indeed be expected on this neural basis.

In order to further explore the interplay between real-life stress and cognitive interference, within a psychophysiological frame linking reactivity and the quality of performance, student pilots were subjected to color–word, numerical and emotional Stroop tasks, once in a baseline recording and once right before their Progress Test General Flying (PTGF). PTGF is the most feared examination flight in the basic flight training, as failure to pass might force the student pilot out of the training. To assess stress induction efficiency, prior to experimental measurements, subjects were administered the State Trait Anxiety Inventory (STAI) (Spielberger, 1983). Cardio-respiratory parameters were recorded both during rest and execution of the cognitive tests. We expected faster responses on both the color–word and the numerical Stroop task under stress, but increased interference effects as a result of a decline in cognitive control. Furthermore, we expected these effects to be related to the magnitude of stress reactivity, as quantified by the cardio-respiratory parameters, fitting within the resource recruitment or effort notion.

2. Method

2.1. Subjects

Student pilots (N = 12) from the Belgian Air Force in their basic flight training, aged 19 to 25 years (mean = 22.5), all medically fit to fly and free of significant medical antecedents, with normal vision, participated.

2.2. Procedure

The total duration of the procedure approximated 40 min. Prior to cognitive testing, participants were equipped with the LifeShirt system (VivoMetrics, Inc.). After a rest recording period of 5 min, participants completed the STAI questionnaire. The cognitive battery was computer driven and lasted for approximately 20 min. Onscreen instructions were followed by a series of 7 cognitive tasks in the following sequence: a Stroop color–word task with neutral words (S1-N) among the color words on a white background, a Stroop task including emotional words among the color words (S1-E) on a white background, a similar Stroop task with neutral words (S2-N) on a black background and a Stroop task with emotional words (S2-E) on a black background. Subsequently, two recognition tasks (Rec1 and Rec2) were presented, each including neutral and emotional words from the lists presented in the four previous tasks, as well as new words. The last Stroop paradigm was a numerical Stroop task (Num). Task presentation (lists and stimuli) was counterbalanced, to control for potential order effects. The procedure was applied in a repeated measure design: the baseline recording took place after approximately one third of the flight training, the stress-condition recording was planned just before the PTGF, a major stress-inducing flight evaluation. This evaluation flight would always take place as the first flight of the day, therefore, all recordings started around 09.00 AM (thus avoiding circadian interference), ended around 09.40 AM, after which the student pilot started with the briefing for his PTGF. Recording sessions were separated by minimum 2 and maximum 5 months. The memory tasks will not be discussed in the present paper. In the description and discussion of results, task will refer to the cognitive tasks, and test will be used to qualify the pre-test condition (i.e. before the evaluation flight).

2.3. Task description

In the color–word Stroop task, words were presented in the middle of the screen, in bold Courier New font, 14 points, under a vertical visual angle of 2° 03′. The response–stimulus interval (RSI) was 32 ms; response times and error rates were recorded. After detailed and standardized onscreen instructions, a 60 trial practice block was inserted, providing subjects with performance feedback. Subjects were instructed to respond to the color in which a word appeared as quickly and accurately as possible, using color-labeled keys on the keyboard. The Stroop task consisted of two lists with stimuli presented on either a white or a black background, counter-balanced between participants. Half of the participants got to see the list on a white background first, whereas the other half received first the list on a black background. Furthermore, the contents of each list were again counterbalanced, meaning each list was presented on the white background half of the time, and on the black background the other half. Each list contained 14 general emotional, 7 pilot specific emotional, 7 student pilot specific emotional, and 28 neutral words, as well as 30 congruent, 30 incongruent, 5 negative priming and 15 inverse negative priming trials. Stimuli appeared in red, blue, yellow or green, in a pseudo-randomized order, limiting consecutive appearance of same color to 2.

For the numerical Stroop task, as described by Pavese and Umiltà (1998), participants had to respond to the amount of stimuli present on the screen. Stimuli were either numbers (2, 3, 4 or 5) or crosses

Please cite this article as: Pattyn, N., et al., Cardiac reactivity and preserved performance under stress: Two sides of the same coin? International Journal of Psychophysiology (2013), http://dx.doi.org/10.1016/j.ijpsycho.2013.03.008

(X), grouped by two, three, four or five. Each stimulus string appeared at every possible location in the horizontal visual range. The experimental list comprised 250 trials, with 20% congruent, 20% neutral, and 60% incongruent trials, presented in a randomized order. The response-stimulus interval (RSI) was 32 ms as well. Recorded variables were response times and error rates. The experimental block was preceded by a 60 trial practice block, during which feedback was provided to participants. Subjects responded through the numerically labeled keys on the keyboard.

2.4. Apparatus

All experimental sessions were run at the same location and the same time of day, starting between 08.30 and 10.30. During the testing, subjects wore a headset to minimize possible noise disturbance. Stimulus presentation, timing and data recording were controlled using E-Prime software (Schneider et al., 2003) on a Sony VAIO laptop computer, with a 15.4′ monitor, at 40 cm viewing distance. Cardiorespiratory parameters were recorded non-invasively through the LifeShirt system (VivoMetrics, Inc.). A standard single lead ECG was recorded with a sampling frequency of 200 Hz and re-sampled at 1000 Hz for R-wave detection. Respiratory movements were measured by respiratory inductive plethysmography; abdominal and ribcage excursions were recorded at 50 Hz. All data were visually inspected for artifacts; ectopic beats or erroneous R-wave detections were manually corrected (removal of erroneous detection/artifact followed by a cubic spline interpolation; corrections < 1%).

Through a derivative based algorithm R-waves were detected, and RR intervals calculated. Other computed outcome variables were respiratory frequency (F_resp), tidal volume (TV), and respiratory sinus arrhythmia (RSA), which was calculated with the peak-valley method. As variation in TV across experimental conditions was more relevant than absolute volume values per se, the Qualitative Diagnostic Calibration (a proprietary relative calibration of volume in arbitrary units implemented in the software Vivologic 2.9.3) was applied to individual data-files.

2.5. Stimuli

Color words used in the classical Stroop task (yellow, green, red and blue), were supplemented with neutral and emotional words, matched for word length and frequency across all conditions and lists with the use of the Dutch CELEX database (Baayen et al., 1995). Wherever possible, stimuli were also matched for familiarity according to the Hermans and De Houwer rating (1994). As for emotional content, there were 4 conditions of words: neutral, general emotional, pilot specific emotional, and student pilot specific emotional. As for congruence, there were four conditions of color words: congruent, incongruent, negative priming and inverse negative priming. All words are listed in Appendix 1. General emotional words were selected from the database of Hermans and De Houwer (1994), pilot specific and student pilot specific emotional words were generated by three independent pilots who did not participate in the experiment. Neutral words were matched to these selected stimuli, on word length and chosen from average frequency.

2.6. Cognitive data-analysis

Color Stroop and emotional Stroop effects were analyzed separately. We aimed at investigating differences between distinct levels of interference, and therefore performed ANOVA on the color Stroop data with conflict as a four level factor (levels: congruent, incongruent, negative priming and inverse negative priming). Congruence and priming effects were explored separately so as to permit investigation of incremental conflict levels. ANOVA on emotional Stroop data was performed with emotion as a four level factor (levels: neutral, general emotional, pilot specific emotional and student pilot specific emotional words). Except when explicitly mentioned, all testing was two-sided testing. Post-hoc investigation of significant differences was done with the ANOVA contrasts, and all statistical testing applied the significance level of 0.05.

2.7. Physiological data-analysis

Considering the known trade-off in psychophysiology between the need for longer recordings from a statistical point of view versus the need to analyze a period with a signal as close as possible to stationarity, 2-minute sequences are recommended as an acceptable compromise (Berntson et al., 1997). Hence, the subtests were programmed to approximate this duration. These 2-minute sequences were thus extracted from the recordings of the different subtests, starting after the practice trials when applicable, or at the beginning of the task. A repeated measures 8 (Test) $*$ 2 (Session) MANOVA was first performed on all physiological variables to map interdependency patterns before performing separate ANOVA's, if necessary complemented with ANCOVA's. Post-hoc investigation of significant differences was done with the ANOVA contrasts, and all statistical testing applied the significance level of 0.05.

3. Results

3.1. Anxiety scores

The baseline session yielded mean scores for trait and state anxiety of 42.64 (SEM = 1.42) and 43.92 (SEM = 1.23) respectively, whereas pre-exam recording returned mean scores of 40.64 (SEM = 1.23) and 50.15 (SEM = 2.79) for trait and state. Under the a priori assumption of elevated anxiety scores in the pre-exam recording, a one-tailed *t*-test confirmed the state anxiety increase to be significant [$t(11) = 2.18$; $p = 0.021$; $\eta^2 = 0.30$].

3.2. Color–word Stroop task

A 4 (condition) $*$ 2 (session) ANOVA on RTs yielded a significant effect of condition [$F(3,33) = 3.13$; $p = 0.039$; $\eta^2 = 0.221$], but no significant effect of session [$F(1,11) = 2.17$; $p > 0.1$], nor any interaction between session and condition [$F(3,33) < 1$]. As depicted in Fig. 1, the pre-exam condition does not modulate the interference effect. With respect to accuracy, the stress-level does not seem to modulate the interference effect in congruent and incongruent word conditions, as depicted in Fig. 2. A 4 $*$ 2 ANOVA on error rates after arcsine transformation returned none of the observed variations significant: condition [$F(3,33) = 1.99$; $p > 0.1$], session [$F(1,11) < 1$] and their interaction [$F(3,33) < 1$] all failed to reach significance.

3.3. Emotional Stroop task

Session did not impact on the emotional content interference effects, as visualized in Fig. 3. RT variation did not reach significance: the 4 (emotion) $*$ 2 (session) ANOVA did not yield any significant effect. The accuracy data, on the other hand, did show the expected interference pattern, as depicted in Fig. 4. Pre-exam recordings showed an increase in error rates as compared to the baseline condition, with highest increase in errors in the ‘emo-stud’ condition, containing words related to student pilot evaluation. A 4 $*$ 2 ANOVA after arcsine transformation returned a significant effect of session [$F(1,11) = 7.18$; $p = 0.021$; $\eta^2 = 0.395$] and of emotion [$F(3,33) = 55.61$; $p < 0.001$; $\eta^2 = 0{,}835$], as well as a significant interaction between both [$F(3,33) = 3.78$; $p = 0.02$; $\eta^2 = 0.256$]. Only the contrast analysis on student pilot specific emotional stimuli [$F(1,11) = 13.82$; $p = 0.003$] showed these to distinguish significantly between sessions.

Please cite this article as: Pattyn, N., et al., Cardiac reactivity and preserved performance under stress: Two sides of the same coin? International Journal of Psychophysiology (2013), http://dx.doi.org/10.1016/j.ijpsycho.2013.03.008

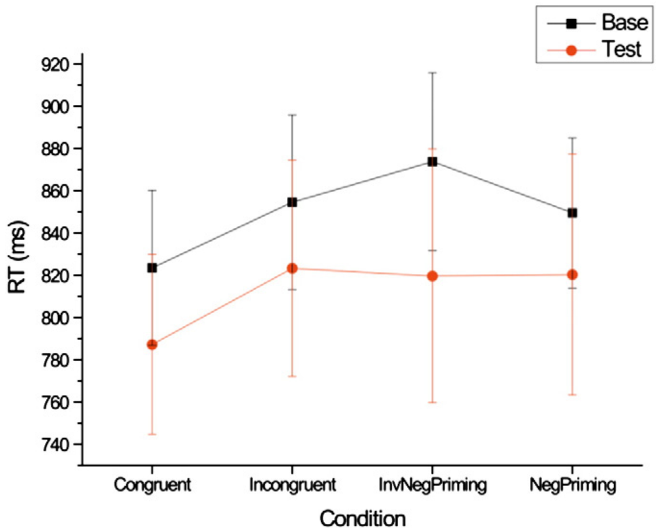

Fig. 1. RTs (ms) for the different conditions of the Stroop color–word test for student pilots on baseline and pre-test recordings.

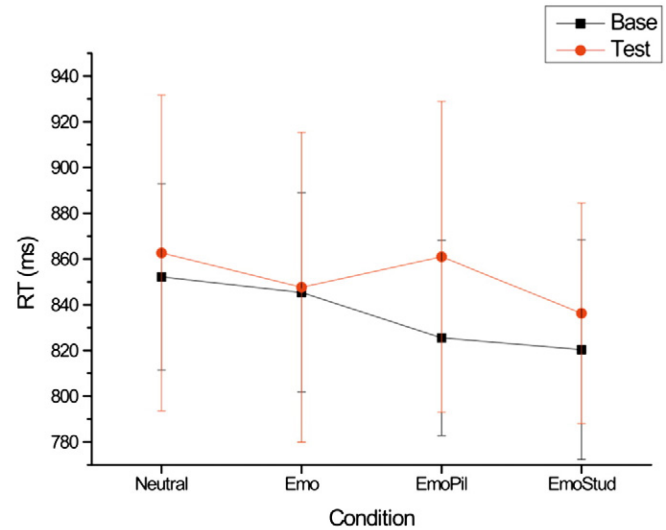

Fig. 3. RTs (ms) for the different conditions of the emotional Stroop task for student pilots on baseline and pre-test recordings.

3.4. Numerical Stroop test

As shown in Fig. 5, session did not seem to modulate the interference effect. A 2 (condition) ∗ 2 (session) ANOVA on RTs confirmed the significant effect of condition [$F(1,11) = 60.2$; $p < 0.001$; $\eta^2 = 0.858$], but the absence of a significant session effect [$F(1,11) = 1.57$; $p > 0.1$], and a lack of interaction between both [$F(1,11) = 1.32$; $p > 0.1$]. As for accuracy, as shown on Fig. 6, session did not impact on error rate, but interference in the incongruent condition seemed substantially higher in the pre-exam test session. A 2 (condition) ∗ 2 (session) ANOVA after arcsine transformation confirmed the significant main effect of condition [$F(1,11) = 38.62$; $p < 0.001$; $\eta^2 = 0.794$], the lack of significant main effect of session [$F < 1$], and the absence of a significant interaction between both [$F < 1$].

3.5. Physiological data

Mean values and respective standard deviations for the recorded and computed physiological parameters for both baseline and pre-exam recordings are reported in Table 1, according to the different subtest rest, S1-N, S1-E, S2-N, S2-E, Rec1, Rec2 and Num. These physiological data have been reported in details elsewhere (Pattyn et al., 2010). To summarize, two different effects showed when comparing baseline and pre-exam recordings. Firstly, the mean rest values differ, most markedly for RRI (772 ms for the pre-exam vs 851 ms for the baseline recording) and for RSA (133 ms for the pre-exam vs 189 ms for the baseline recording), indicating the expected higher activation during the pre-exam session. Secondly, the size of the initial reactivity – the difference between rest values and the first test presentation – due to cognitive test presentation decreases in the pre-exam condition, most markedly for RRI.

A repeated measures 8 (Test) ∗ 2 (Session) MANOVA on the variables described in Table 1 showed a significant effect of both Test [$F(7,143) = 28.73$; $p < 0.001$; $\eta^2 = 0.584$], and Session [$F(1,143) = 3.8$; $p = 0.003$; $\eta^2 = 0.12$], but no interaction [$F(7,143) < 1$]. The univariate ANOVAs for Test showed a significant effect for F_resp [$F(7,143) = 17.59$; $p < 0.001$; $\eta^2 = 0.463$] and for RSA [$F(7,143) = 9.01$; $p < 0.001$; $\eta^2 = 0.306$]. The univariate ANOVAs for Session showed a significant effect for RRI [$F(1,143) = 11.17$; $p = 0.001$; $\eta^2 = 0.072$] and for RSA [$F(1,143) = 10.35$; $p = 0.002$; $\eta^2 = 0.067$]. Separate MANOVAs for the baseline and pre-exam recordings showed similar effects of Test, respectively

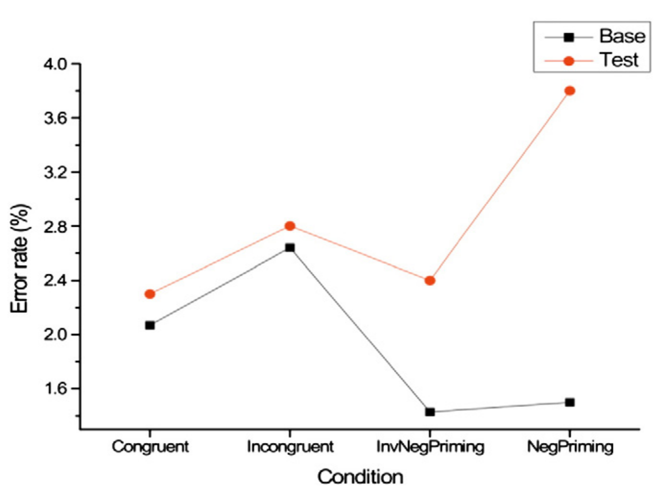

Fig. 2. Error rates (%) for the different conditions of the Stroop color–word task for student pilots on baseline and pre-test recordings.

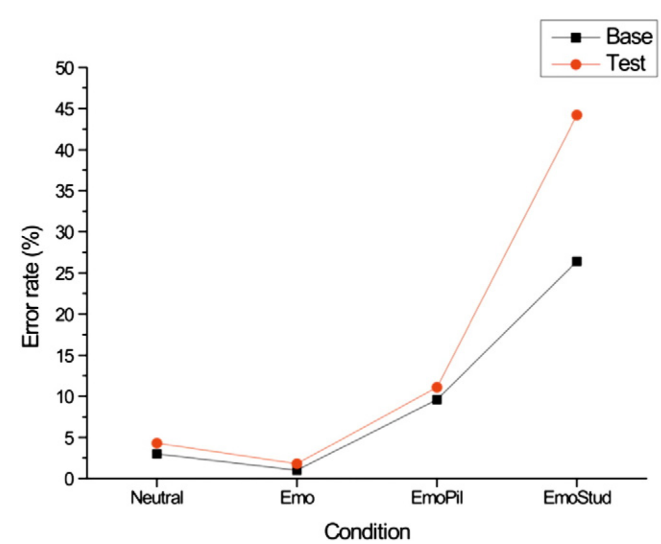

Fig. 4. Error rates (%) for the different conditions of the emotional Stroop task for student pilots on baseline and pre-test recordings.

Please cite this article as: Pattyn, N., et al., Cardiac reactivity and preserved performance under stress: Two sides of the same coin? International Journal of Psychophysiology (2013), http://dx.doi.org/10.1016/j.ijpsycho.2013.03.008

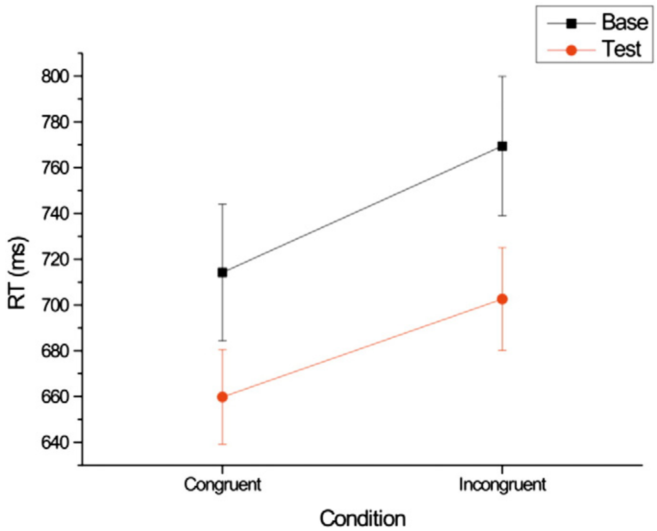

Fig. 5. RTs (ms) for the numerical Stroop task for student pilots on baseline and pre-test recordings.

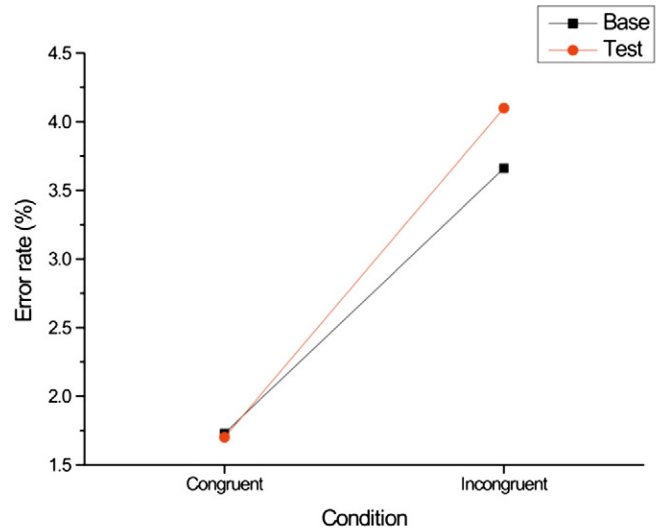

Fig. 6. Error rates (%) for the numerical Stroop task for student pilots on baseline and pre-test recordings.

[$F(7,71) = 16.97$; $p < 0.001$; $\eta^2 = 0.626$] and [$F(7,71) = 14.91$; $p < 0.001$; $\eta^2 = 0.592$], with univariate ANOVAs for both showing significant variations for F_resp and RSA. As we previously discussed in details (Pattyn et al., 2010), the analysis strategy of using MANOVAs first aimed at mapping the internal dependencies of the different outcome variables. Indeed, as we have shown, the RSA reactivity is mainly dependent on respiration, and these two can thus not be treated as independent variables.

Considering the fact that the data showed no effect of presentation of emotional words, no effect of task difficulty nor of task switching, neither during the baseline nor during the pre-exam session, the main effect in the physiological data is thus the initial reactivity between rest recordings and recordings during presentation of the first task, with a tendency to return to rest values along time-on-task.

Despite the overall MANOVA showing no significant interaction between session and test, the initial reactivity in RRI differs between baseline and pre-exam recordings, and RRI is the parameter for which this difference appears most clearly. The difference in reactivity is most noticeable when the RRI values described in Table 1 are expressed in proportional changes when compared to rest values. In baseline recordings, the initial reactivity in RRI is expressed as a mean decrease of 10%, whereas in pre-exam recordings, this initial reactivity represents only a decrease of 2.3%. This interesting effect might be blurred by interindividual difference, and would thus require a within-subject standardization to show significance, which is a repeatedly acknowledged issue in psychophysiological research (e.g., Bush et al., 1993). The range-correction procedure (Lykken and Venables, 1971, in Grossman and Kollai, 1993) was thus applied to the RRI for initial reactivity data. The result of this range-correction procedure is depicted in Fig. 7.

This effect was examined through a 2 (Session) $*$ 2 (Test) repeated measures ANOVA, which showed a significant effect of test [$F(1,10) = 22.38$; $p = 0.001$; $\eta^2 = 0.72$], no significant effect of session [$F(1,10) = 2.29$; $p > 0.1$] and a significant interaction between session and test [$F(1,10) = 10.62$; $p = 0.01$; $\eta^2 = 0.54$].

3.6. Linking the cognitive and physiological data

Results in the stress condition are characterized by higher error rates, an increased resting activation (RRI & RSA) and a decreased

Table 1
Summary of means and standard deviations for cardio-respiratory parameters throughout the sequence of cognitive tests, for baseline and pre-exam recordings. All data were computed based on 2 minute segments.

	Rest	S1-N	S1-E	S2-N	S2-E	Rec1	Rec2	Num
RRI (ms)	851	765	803	795	807	841	877	828
baseline	111	127	110	109	96	110	115	120
RRI (ms)	772	753	747	754	758	779	774	770
exam	111	107	98	95	98	120	103	99
F_resp (min^{-1})	9.5	19.1	17.6	17.9	17.9	15.3	16.0	17.6
baseline	2.5	3.1	3.4	4.1	3.1	2.5	2.9	2.9
F_resp (min^{-1})	10.4	19.8	18.1	17.1	17.4	16.0	15.4	17.5
exam	2.5	2.5	2.2	4.1	3.3	2.1	2.8	4.3
TV (a.u.)	19242	19628	19622	19594	19552	19884	20258	19670
baseline	4759	4949	5005	5009	5023	5018	5140	4989
TV (a.u.)	17864	18066	18040	18089	18063	18177	18230	18256
exam	5 734	5 868	5 879	5 856	5 879	5 947	5977	5973
Ti/Ttot	0.36	0.40	0.38	0.39	0.38	0.37	0.37	0.39
baseline	0.05	0.04	0.05	0.06	0.04	0.04	0.04	0.04
Ti/Ttot	0.36	0.41	0.39	0.38	0.39	0.37	0.38	0.39
exam	0.05	0.03	0.02	0.05	0.03	0.04	0.07	0.05
RSA (ms)	189	63	77	67	69	82	95	74
baseline	110	35	42	35	33	49	59	46
RSA (ms)	133	41	44	57	55	56	62	55
exam	100	21	21	43	43	28	35	40

Please cite this article as: Pattyn, N., et al., Cardiac reactivity and preserved performance under stress: Two sides of the same coin? International Journal of Psychophysiology (2013), http://dx.doi.org/10.1016/j.ijpsycho.2013.03.008

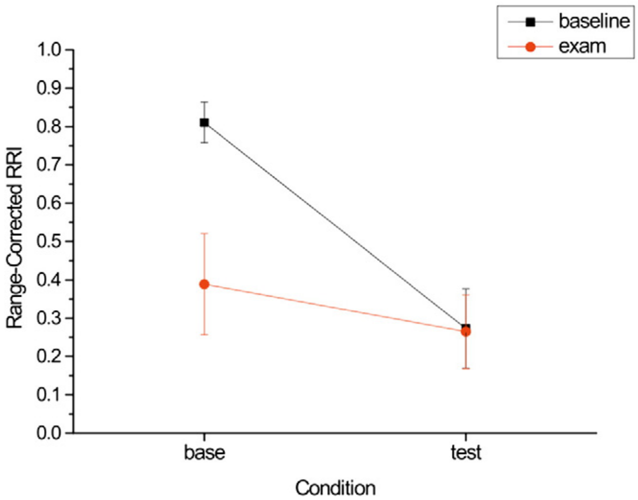

Fig. 7. Range-corrected RRI variation between rest and test recordings, for baseline and pre-exam sessions.

reactivity. The question remains whether these changes are merely co-occurring or whether there is a relationship between them. To investigate this, we performed a correlation analysis on the significant changes from baseline to stress, being the increase in error rate for specific emotional stimuli, the decrease in resting RRI and resting RSA, and the reactivity decrease in RRI. This yielded a significant correlation between the increase in error rate and the reactivity decrease ($r = 0.42$; $p = 0.03$), meaning that the more reactivity decreases in the stress condition, the more the error rate on specific stimuli increases. The increase in error rate did not correlate significantly with neither the decrease in resting RRI nor the decrease in resting RSA.

4. Discussion

The aim of the present study was to investigate the effect of a naturalistic stress paradigm on Stroop interferences, and whether this effect was related to the autonomic response to stress. We hypothesized that the effect of stress would fasten the responses, but show an increased interference effect due to a decline in cognitive control. Furthermore, we expected this decline in cognitive control to be related to stress reactivity, as measured by the cardio-respiratory parameters.

The STAI results showed that the chosen paradigm indeed elicited situational stress, and was thus suited to the purpose of our investigation.

The apparent speed–accuracy trade-off in the stress condition was not significant. Exam-related stress did not significantly alter performance on the classical color–word Stroop, nor was the size of the interference modulated. Insertion of emotional words did not significantly alter RTs either. However, and contrary to earlier studies where mainly RT alterations were found (Ray, 1979; Rutherford et al., 2004), we observed a remarkable increase in error rates on the student-pilot specific emotional stimuli, which lead us to conclude that a specific subject-related emotional interference effect might be sensitive to particular stress contexts. In summary, we did not replicate Kofman et al.'s (2006) facilitatory effect of stress on executive function, as shorter RTs did not reach significance, and co-occurred with increased error rates. In fact, the more specific, self-referent, and situationally relevant the emotional stimuli were, the larger the interference, expressed as error rates, under stressful testing conditions.

Several limitations in the present study should be addressed. The fact that there was no control for repetition of testing does not allow to rule out practice and carry-over effects between baseline and pre-exam testing. However, the minimal period of two months between two test sessions seems, in our opinion, sufficient to minimize such effects. Furthermore, these effects alone could not explain our findings, as responses to emotional words were slower and less accurate during retest. The idea that the use of similar word material might have boosted the emotional effect is not in line with previous research (for a review, see Williams et al., 1997), which tends to show that repetition might attenuate emotional responses through habituation. Another limitation of the present study is the small number of participants. Whereas this is a common drawback of applied research targeting very specific populations, the levels of significance obtained for the emotional effect, as well as the important effect sizes, warrant the robustness of the findings. This small sample size may however account for the lack of significance of the speed/accuracy trade-off in the non-emotionally related stimuli. With regard to the range-correction procedure this is in our opinion necessary and justified because, as

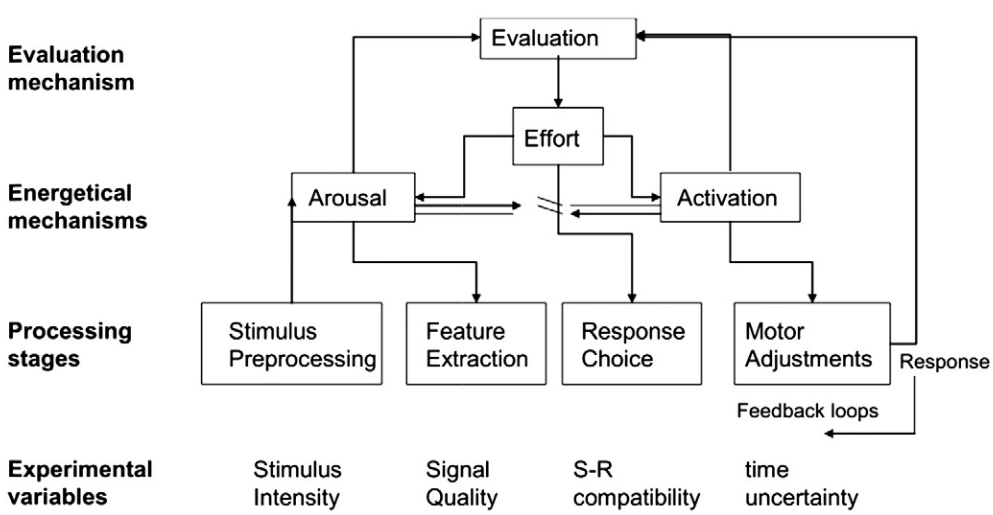

Fig. 8. "A cognitive-energetical linear stage model of information processing and stress. The cognitive level consists of computational processing stages derived by means of the additive factor method (Sternberg, 1969). There are three energetical supply mechanisms, two of which are basal (arousal and activation) and coupled to respectively input and output processing stages. The basal mechanisms are coordinated and supervised by effort, which is also directly linked to the central stage of response choice. Apart from direct energetical supply to this stage, effort serves the function of keeping the basal mechanisms at optimal value. Information about the state of the basal mechanisms is mediated by an evaluation mechanism."
From Sanders (1983).

Please cite this article as: Pattyn, N., et al., Cardiac reactivity and preserved performance under stress: Two sides of the same coin? International Journal of Psychophysiology (2013), http://dx.doi.org/10.1016/j.ijpsycho.2013.03.008

has been pointed out by several other authors in psychophysiology (e.g. Ohman et al., 2000), the range of variation we target with subtle changes in mental activity or overall activation are only representative of a minor portion of the full scale of potential cardiac reactivity. Hence, these subtle variations might be blurred by the known interindividual variability, or by larger effects, such as these of a real-life stress situation. Interindividual standardization procedures, like the range-correction for the RRI, allow for a more precise investigation of these effects.

The high error rates for the SP-specific stimuli, which are already higher in the baseline condition, are probably related to the specificity of the material. One could argue that the fact these words are in English makes them stand out when compared to the neutral and emotional stimuli, however, the pilot-specific stimuli are also presented in English (indeed, for these categories, there is no available translation, since during their briefings, flight, debriefings, classes and general instruction student pilots work in English) and do not show the same pattern of results. Previous preliminary results on university students (Pattyn, 2007) allow us to conclude that it is the specificity of the presented material that is the discriminant feature in eliciting the increase in error rates.

Overall, these results suggest a modulation of cognitive control related to emotional material under the influence of stress: our subjects seem less able to suppress irrelevant stimulus dimensions, only when these are related to their current concern.

We had expected that the effect of stress on performance would be related to the physiological reactivity. Indeed, physiological reactivity range appeared to be related to the quality of cognitive performance. This might seem surprising at first, however, it is in line with previous psychophysiological findings. In Kennedy and Scholey's (2000) serial subtraction task, higher increases in heart rate correlated with a better performance (higher accuracy and throughput), and lower heart rate baselines coupled to higher reactivity in response to cognitive processing, manifested as important performance mediators. These authors thus concluded that an individual's physiological efficiency, both in terms of increase in heart rate and glucose utilization during task, is predictive of cognitive performance. Hansen et al. (2003) investigated vagal influences on working memory and attention, and found high HRV individuals to perform better than low HRV subjects on both a working memory and a sustained attention task. Carter and Pasqualini (2004) observed a positive correlation between success on a gambling task and anticipatory autonomic response. These results not only confirm the association between performance quality and elicited physiological response, but also suggest a causal link between both.

Despite the fact that correlation may not be mistaken for causality, the observed relationship between the performance decrement (increased error rates) and the decrease in heart rate reactivity does fit within Sanders' model of performance. Indeed, one of our research questions was whether the performance decrement observed in the stress-condition, would be related to the physiological response to the task. Sanders (1983) applied his model, depicted in Fig. 8, to explain raised error rates under the influence of "stress", that is, a rise in arousal will trigger a signal to the activation system, enhancing response readiness and response execution, which causes an increase in error rates in the case of stronger demands on response choice. In order to avoid unacceptable error rates, immediate arousal in conditions of already increased activation should be suppressed and this may be accomplished by uncoupling both systems through effort (Sanders, 1983).

Although our results only allow to suggest this as a hypothesis for further research, the decreased arousal (decreased RRI-reactivity) in the pre-exam session might be a suppression of the arousal response on the background of a higher activation (higher resting heart rate and lower resting parasympathetic tone). These results challenge the traditional view on cardiac reactivity, stemming from stress research, as a dispositional marker of pathology. In the present results, when the system functions in a "nominal" state, reactivity, i.e. arousal, is high. However, in the "stress" condition, arousal due to mental challenge is decreased. As has been stated by various authors (e.g. Pool, 1989; Thayer and Lane, 2000; Van Diest et al., 2006) healthy systems are characterized by variability and flexibility, which allow to adapt to varying environmental demands.

The interaction between cognitive control, emotion regulation and autonomic reactivity suggested in the present findings warrants this lead to be fruitful for future research to further unveil the relationship between stress and performance.

Acknowledgments

This work was supported by a Prodex grant 90030 (European Space Agency/Belgian Federal Government) and by grant ERM-HF10 (Belgian Department of Defense). Dr Pattyn's work is supported by

Appendix 1

Table 2
Dutch neutral and general emotional words, along with affectivity and familiarity ratings (Hermans and De Houwer, 1994) and the word frequency, as identified through the CELEX database.

Condition	Stimulus	Affective load	Familiarity	Frequency
Emo	Aids	127	45	–
Emo	Alcoholisme	142	413	52
Emo	Bedreiging	173	378	787
Emo	Begrafenis	175	413	906
Emo	Bommen	134	303	433
Emo	Braaksel	17	343	100
Emo	Coma	158	314	103
Emo	Drugs	169	431	483
Emo	Executie	133	291	240
Emo	Gangster	17	306	60
Emo	Gezwel	144	331	162
Emo	Gijzelaar	171	298	85
Emo	Haat	136	45	1558
Emo	Incest	12	341	77
Emo	Kanker	136	445	750
Emo	lijk	175	364	993
Emo	Marteling	128	288	99
Emo	Misdaad	152	369	674
Emo	Moord	116	378	1609
Emo	Ongeluk	153	503	1783
Emo	Oorlog	123	43	7810
Emo	Pedofiel	148	284	7
Emo	Pijn	18	55	6313
Emo	Sadist	172	363	42
Emo	Slachting	148	313	84
Emo	Stank	167	422	664
Emo	Tandpijn	172	417	12
Emo	Tiran	17	278	133
Emo	Tumor	131	334	200
Neutral	Absorptie	375	261	40
Neutral	Accent	4	453	1145
Neutral	Adelaar	452	294	142
Neutral	Advocaat	367	388	1170
Neutral	Agentschap	389	339	112
Neutral	Ansjovis	338	313	81
Neutral	Antilope	452	258	15
Neutral	Autobus	386	469	94
Neutral	Basketbal	433	445	12
Neutral	Bladzijde	422	592	490
Neutral	Boog	4	353	1071
Neutral	Bord	408	588	1622
Neutral	Circus	477	389	274
Neutral	Cirkel	419	445	831
Neutral	Condoom	456	473	63
Neutral	Disco	409	436	45
Neutral	Eicel	447	394	110
Neutral	Gebeente	33	363	84
Neutral	Geur	452	48	2452
Neutral	Gist	397	352	97
Neutral	Haring	345	317	180

(continued on next page)

Please cite this article as: Pattyn, N., et al., Cardiac reactivity and preserved performance under stress: Two sides of the same coin? International Journal of Psychophysiology (2013), http://dx.doi.org/10.1016/j.ijpsycho.2013.03.008

Table 2 (*continued*)

Condition	Stimulus	Affective load	Familiarity	Frequency
Neutral	Hazelnoot	458	419	20
Neutral	Herfst	453	544	944
Neutral	Hoed	433	442	1314
Neutral	Honing	469	461	524
Neutral	Hoofdsteun	453	395	6
Neutral	Inkt	427	459	333
Neutral	Inspanning	444	58	1186
Neutral	Kaas	469	584	1837
Neutral	Kapper	445	494	379
Neutral	Klei	417	348	339
Neutral	Krant	434	577	3041
Neutral	Kruid	444	406	175
Neutral	Magazine	442	511	49
Neutral	Microscoop	4	336	136
Neutral	Muren	342	527	2169
Neutral	Naaimachine	4	355	77
Neutral	Piano	514	464	643
Neutral	Rok	425	497	903
Neutral	Saffier	481	242	10
Neutral	Sap	459	503	366
Neutral	Schaar	389	527	218
Neutral	Sigaar	323	359	751
Neutral	Slaapzaal	361	331	148
Neutral	Slager	37	478	303
Neutral	SPIEREN	467	506	1035
Neutral	Spinazie	448	428	130
Neutral	Staal	339	33	327
Neutral	Stoep	402	461	658
Neutral	Straling	377	366	428
Neutral	Streep	4	503	556
Neutral	Tapijt	445	481	485
Neutral	Tekstverwerker	413	413	49
Neutral	Televisie	466	603	1954

Table 3
Specific emotional words presented to the student pilots (SPs) in the experiment.

Pilot specific	SP specific
Alert	Test
Defect	Rev
Emergency	Evaluatie
Failure	Commissie
Fire	Buis
Mayday	Schrapping
Alarm	Onvoldoende
Birdstrike	Instructeur
Abort	Mislukken
Crash	Retest
Stall	Les
Checks	Syllabus
Drill	Falen
Incident	Fout

a Euro Space Society grant through the Fonds voor Wetenschappelijk Onderzoek. Dr Bart Roelands is a postdoctoral research fellow from the FWO (Fonds voor Wetenschappelijk Onderzoek). Dr Kolinsky is a Research Director from the FNRS (Fonds National de la Recherche Scientifique). The authors wish to thank the 5th Squadron from the Belgian Air Component Flying School (1 Wing), as well as the student pilots who agreed to participate, for their support and collaboration.

References

Akerstedt, T., Gillberg, M., Hjemdahl, P., Sigurdson, K., Gustavsson, I., Daleskog, M., et al., 1983. Comparison of urinary and plasma-catecholamine responses to mental stress. Acta Physiologica Scandinavica 117, 19–26.

Baayen, R.H., Piepenbrock, R., Gulikers, L., 1995. The CELEX Lexical Database [CD-ROM]. Linguistic Data Consortium, University of Pennsylvania, Philadelphia.

Berntson, G.G., Bigger, J.T., Eckberg, D.L., Grossman, P., Kaufmann, P.G., Malik, M., et al., 1997. Heart rate variability: Origins, methods, and interpretive caveats. Psychophysiology 34, 623–648.

Bush, L.K., Hess, U., Wolford, G., 1993. Transformations for within-subject designs — a Monte-Carlo investigation. Psychological Bulletin 113, 566–579.

Carter, S., Pasqualini, M.C.S., 2004. Stronger autonomic response accompanies better learning: a test of Damasio's somatic marker hypothesis. Cognition & Emotion 18, 901–911.

Grossman, P., Kollai, M., 1993. Respiratory sinus arrhythmia, cardiac vagal tone, and respiration — within-individual and between-individual relations. Psychophysiology 30, 486–495.

Hansen, A.L., Johnsen, B.H., Thayer, J.E., 2003. Vagal influence on working memory and attention. International Journal of Psychophysiology 48 (3), 263–274.

Heims, H.C., Critchley, H.D., Martin, N.H., Jager, H.R., Mathias, C.J., Cipolotti, L., 2006. Cognitive functioning in orthostatic hypotension due to pure autonomic failure. Clinical Autonomic Research 16, 113–120.

Hermans, D., De Houwer, J., 1994. Affective and Subjective familiarity ratings of 740 Dutch words. Psychologica Belgica 34, 115–139.

Hockey, G.R.J., 1997. Compensatory control in the regulation of human performance under stress and high workload: a cognitive energetical framework. Biological Psychology 45, 73–93.

Hughdahl, K., 1995. Psychophysiology: the mind–body perspective, 1st ed. Harvard University Press, Cambridge, MA.

Kamarck, T.W., Jennings, J.R., Pogue-Geile, M., Manuck, S.B., 1994. A multidimensional model for cardiovascular reactivity. Stability and cross-validation in 2 adult samples. Health Psychology 13 (6), 471–478.

Kennedy, D.O., Scholey, A.B., 2000. Glucose administration, heart rate and cognitive performance: effects of increasing mental effort. Psychopharmacology 149, 63–71.

Kofman, O., Meiran, N., Greenberg, E., Balas, M., Cohen, H., 2006. Enhanced performance on executive function tasks associated with examination stress: evidence from task switching and Stroop paradigms. Cognition & Emotion 20, 577–595.

MacLeod, C.M., 1991. Half a century of research on the Stroop effect — an integrative review. Psychological Bulletin 109, 163–203.

Matthews, G., Desmond, P.A., Joyner, L.A., Carcary, W., Gilliland, K., 1997. A comprehensive questionnaire measure of driver stress and affect. In: Rothengatter, T., Vaya, E.C. (Eds.), Traffic and Transport Psychology. Elsevier, Amsterdam.

Matthews, G., Dorn, L., Hoyes, T.W., Davies, D.R., Glendon, A.I., Taylor, R.G., 1998. Driver stress and performance on a driving simulator. Human Factors 40 (1), 136–149.

McKenna, F.P., Sharma, D., 1995. Intrusive cognitions: an investigation of the emotional Stroop task. Journal of Experimental Psychology: Learning and Cognition 21, 1595–1607.

Ohman, A., Hamm, A., Hugdahl, K., 2000. Cognition and the autonomic nervous system. In: Cacioppo, J.T., Tassinary, L.G., Berntson, G.G. (Eds.), Handbook of Psychophysiology. Cambridge University Press, Cambridge, MA, pp. 533–575.

Pattyn, N., 2007. Psychophysiological measures of cognitive performance in operational conditions: applications to aviation and space environments. Doctoral Dissertation. Vrije Universiteit Brussel and Royal Military Academy, Belgium.

Pattyn, N., Migeotte, P.F., Neyt, X., Van den Nest, A., Cluydts, R., 2010. Comparing real-life and laboratory induced stress reactivity on cardio-respiratory parameters: differentiation of a tonic and a phasic component. Physiology & Behavior 101, 218–223.

Pavese, A., Umiltà, C., 1998. Symbolic distance between numerosity and identity modulates Stroop interference. Journal of Experimental Psychology. Human Perception and Performance 24, 1535–1545.

Pool, R., 1989. Cardiac chaos. Science 243 (4897), 1419.

Ray, C., 1979. Examination stress and performance on a color–word interference test. Perceptual and Motor Skills 49, 400–402.

Rutherford, E.M., MacLeod, C., Campbell, L.W., 2004. Negative selectivity effects and emotional selectivity effects in anxiety: differential attentional correlates of state and trait variables. Cognition & Emotion 18, 711–720.

Sanders, A.F., 1983. Towards a model of stress and human-performance. Acta Psychologica 53, 61–97.

Schneider, W., Eschman, A., Zuccolotto, A., 2003. E-Prime Version 1.1 [Computer software]. Psychology Software Tools, Inc., Pittsburgh.

Selye, H., 1956. The Stress of Life. McGraw-Hill, Toronto.

Spielberger, C.D., 1983. State-Trait Anxiety Inventory. Consulting Psychologists Press Inc.

Sternberg, S., 1969. The discovery of processing stages: Extensions of Donders' method. In: Koster, W.G. (Ed.), Attention and performance II: Acta Psychologica, 30, 276–315.

Stroop, J.R., 1935. Studies of interference in serial verbal reactions. Journal of Experimental Psychology 18, 643–662.

Thayer, J.F., Lane, R.D., 2000. A model of neurovisceral integration in emotion regulation and dysregulation. Journal of Affective Disorders 61, 201–216.

Tipper, S.P., 1985. The negative priming effect: Inhibitory priming by ignored objects. Quarterly Journal of Experimental Psychology 37A, 571–590.

Van Diest, I., Thayer, J.F., Vandeputte, B., Van De Woestijne, K.P., Van den Bergh, O., 2006. Anxiety and respiratory variability. Physiology & Behavior 89, 189–195.

Williams, J.M.G., Watts, F.N., MacLeod, C., Mathews, A., 1997. Cognitive Psychology and Emotional Disorders, 2nd ed. Wiley, Chichester, UK.

Wright, C.E., O'Donnell, K., Brydon, L., Wardle, J., Steptoe, A., 2007. Family history of cardiovascular disease is associated with cardiovascular responses to stress in healthy young men and women. International Journal of Psychophysiology 63, 275–282.

Yerkes, R.M., Dodson, J.D., 1908. The relation of strength of stimulus to rapidity of habit-formation. Journal of Comparative Neurology and Psychology 18, 459–482.