Failures of executive function when at a height: Negative height-related appraisals are associated with poor executive function during a virtual height stressor

Vera E. Newman[a], Belinda J. Liddell[a], Tom Beesley[b] & Steven B. Most[a]

a. School of Psychology

UNSW Sydney, NSW, 2052

Australia

b. Department of Psychology

Lancaster University, Lancaster, LA1 4YF

United Kingdom

Email:

Vera Newman: veraenewman@gmail.com

Belinda Liddell: b.liddell@unsw.edu.au

Tom Beesley: t.beesley@lancaster.ac.uk

Steven Most: s.most@unsw.edu.au

Please address correspondence to:

Vera Newman
Email: veraenewman@gmail.com

Or

Steven Most

School of Psychology
UNSW Sydney
Sydney, NSW, Australia
Email: s.most@unsw.edu.au

Word Count: 5383

Abstract

It is difficult to maintain cognitive functioning in threatening contexts, even when it is imperative to do so. Research indicates that precarious situations can impair subsequent executive functioning, depending on whether they are appraised as threatening. Here, we used virtual reality to place participants at ground level or at a virtual height in order to examine the impact of a threat-related context on concurrent executive function and whether this relationship was modulated by negative appraisals of heights. Executive function was assessed via the Go/NoGo and N-Back tasks, indexing response inhibition and working memory updating respectively. Participants with negative appraisals of heights exhibited impaired executive function on both tasks when performing at a virtual height (i.e., a threat-related context) but not at ground-level, demonstrating the importance of considering the cognitive consequences of individual differences in negative interpretations of emotionally-evocative situations. We suggest that a virtual reality approach holds practical benefits for understanding how individuals are able to maintain cognitive ability when embedded within threatening situations.

Keywords: executive function, stress, threat, appraisals, virtual reality

Threatening situations cause stress and anxiety, both of which can impair executive functioning (Lupien, Maheu, Tu, Fiocco, & Schramek, 2007; Shields, Sazma, & Yonelinas, 2016). An unfortunate consequence is that one's ability to maintain cognitive control can be compromised in precisely those situations where safety and well-being might rely on effective cognitive control. Notably, the degree to which any given situation produces such effects can vary across task or individual (Robinson, Bond, & Roiser, 2015; Roos et al., 2017; Schoofs, Pabst, Brand, & Wolf, 2013). Evidence suggests that different aspects of lab stress-inductions (e.g., psychological versus physiological stress; Dickerson & Kemeny, 2004) and within-person psychological factors (e.g., appraisals and pre-existing anxieties) modulate one's response to those threatening situations (Crişan, Vulturar, Miclea, & Miu, 2016; Gaab, Rohleder, Nater, & Ehlert, 2005; Villada, Hidalgo, Almela, & Salvador, 2016). The Biopsychosocial Model of Challenge and Threat also proposes that an individual's response to a situation is dictated by one's appraisals of that specific situation (Blascovich & Mendes, 2010; Seery, 2013). Similarly, biased interpretations of a situation contribute to the development and maintenance of anxiety (Clark & Beck, 2010), and cognitive models suggest that anxiety (e.g., which is often experienced in contexts that people appraise as threatening) can be detrimental to attention and cognitive control (Eysenck, Derakshan, Santos & Calvo, 2007). That is, coupled with anxious responding to threatening situations, specific appraisals may play a major role in how people perceive and/or respond to potentially threatening situations, and may thus influence the degree to which such situations impact cognition and executive functioning.

It is therefore important to consider how appraisals of potentially threatening situations might modulate the impact of such situations on various outcomes. Research has found that negative height-related appraisals were related to increased psychological and physiological responses (i.e. peak fear, anxiety-related cognitions and physiological responses) during a height exposure task even when controlling for differences in acrophobia (i.e., fear of heights; Steinman & Teachman, 2011). This is consistent with evidence that interpretation biases in anxiety may moderate the relationship between anxiety and physiological arousal during stress (Rozenman, Vreeland & Piacentini, 2016), and that altering interpretation biases in major depression can decrease subsequent stress reactivity (Joormann, Waugh, & Gotlib, 2015). However, a remaining question is whether negative appraisals of one's situation influence not only subjective affect and physiological responses, but also cognitive performance while experiencing a potentially threatening environment.

The impact of negative appraisals on executive function might depend on which aspect of executive function is being measured. Executive functions (EFs) are fundamental cognitive processes that guide behavior toward the achievement of goals (Banich, 2009; Diamond, 2013), and an influential model of EFs suggests that – rather than constituting a unitary construct – they can be broken down into dissociable (but correlated) core functions. These include *response inhibition* (inhibition of a prepotent response), *updating* (the updating and monitoring of working memory representations), and *switching* (shifting between mental sets/tasks; Miyake et al., 2000). Although they can be dissociated, these EFs appear to converge on their reliance on the prefrontal cortex (Miyake & Friedman, 2012), an area where functioning is negatively impacted by stress hormones, including those released via the activation of the HPA axis (Arnsten, 2009). A recent meta-analysis indicates that acute experiences of stress impair updating but enhance response inhibition (Shields et al., 2016). Such a finding highlights the importance – when considering the consequences of appraising a situation as threatening – of delineating the consequences for different aspects of EF.

Although predictions can be made based on existing experiments that employ stress inductions, generalizations from this work are limited by methodological considerations. For example, whereas people often need to maintain high levels of cognitive functioning while immersed in a dangerous situation, performance in lab-based environments is most frequently assessed *following* the offset of a stressor (Plessow, Fischer, Kirschbaum, & Goschke, 2011; Schoofs, Preuß, & Wolf, 2008). This is important, as appraisals of threat may have different impacts depending on when cognitive demands are introduced: investigations of how cognition is impacted during perceived threat may help distinguish shorter-term versus longer-lasting impacts of activation of the sympathetic nervous system (SNS) and the hypothalamic-pituitary-adrenal (HPA) axis, respectively (Hermans, Henckens, Joëls, & Fernández, 2014; Lupien et al., 2007; Wolf, 2003). Further, although lab-based stress induction paradigms trigger robust and measureable physiological reactions, and have led to significant advances in understanding the physiological mechanisms underlying such responses (Allen et al., 2017; Kirschbaum, Pirke, & Hellhammer, 1993), several such lab-based stressors bear limited resemblance to the kinds of threatening contexts people may encounter outside the lab (Zanstra & Johnston, 2011). Thus, a remaining challenge is to optimally test the cognitive impact of real-time immersion in a scenario that people appraise as threatening, while maintaining some semblance to situations that people might reasonably expect to encounter outside the lab. The current study thus investigated whether the impact of a threatening situation on executive function differed depending on whether the situation was appraised as threatening.

Towards this end, the current study immersed participants in a virtual reality (VR) environment associated with measurable variation in negative appraisals, and we assessed whether its impact on EFs could be predicted by such negative appraisals. The VR environment involved a cityscape height manipulation, selected to take advantage of widespread individual differences in negative appraisals of heights (Kapfhammer, Huppert, Grill, Fitz, & Brandt, 2015; Steinman & Teachman, 2011). In a *high* VR height scenario, participants sat on a pole high above a virtual city, and in a *low* VR scenario they sat at virtual ground level. In order to predict the impact of this manipulation on EF, participants completed the previously validated Heights Interpretation Questionnaire (Steinman & Teachman, 2011), which assesses the degree to which one tends to appraise heights as dangerous. Given that exposure to threatening contexts has been found to have a differing impact on EF measures of updating and response inhibition (Shields et al., 2016), we assessed these two EFs in a within-subjects design using the N-back and Go/NoGo tasks respectively. These tasks were chosen due to their frequent use in investigations of executive function (Miyake et al., 2000; Shields et al., 2016; Simmonds, Pekar, & Mostofsky, 2008), as well as the ability to adapt these tasks into the auditory modality so as not to interfere with visual processing of the VR environment. We also collected physiological responses (heart rate and heart rate variability) and subjective measures of affective reactivity to further discern the mechanisms underlying how negative appraisals of the VR height manipulation may modulate these responses.

We hypothesized that the VR height manipulation would negatively impact both response inhibition and updating functions, but that this effect would be moderated by participant appraisals of the situation, such that executive functioning would be more impaired for participants with more negative appraisals of heights than for those with less negative appraisals of heights. We also hypothesized that negative appraisals of heights would be associated with increased physiological responses and higher negative affect at the high VR height, but not the low VR height, compared to less negative height-related appraisals.

Method

Participants

Fifty-eight participants (33 females; Age_{years}: M=21, SD=5.6) gave informed consent in line with the University of New South Wales, Sydney Human Research Ethics Advisory Panel and either received course credit for their participation or were reimbursed $15 for their time. Participants were instructed to avoid any caffeinated beverages and excessive physical exercise during the four hours preceding the experiment. Participants were required to have normal or corrected-to-normal vision. Participants were excluded from participating if they were a smoker, used illicit drugs, consumed large amounts of alcohol regularly (>3 drinks/day), had serious medical conditions (i.e. any condition affecting the heart, any diagnosed psychological disorder), or were on prescribed medications (this included birth control). These exclusion criteria were adopted due to evidence suggesting that these factors influence cardiovascular recordings (Fatisson, Oswald, & Lalonde, 2016; Hayano et al., 1990; Quintana, Alvares, & Heathers, 2016).

Measures

Virtual reality scenario. VR scenarios were custom-created in the Unity video game engine by a professional video-game programmer utilizing Development Kit 2 for the Oculus Rift (California, US). Within the VR cityscape, participants were virtually placed in a chair either at ground height or atop a pole several stories above ground. At the High VR height, participants' virtual chair was tilted forward, facing towards the ground (see Figure 1). In a within-subjects design, participants performed two executive function tasks (response inhibition and updating) in both the low- and high- height scenarios in an alternating order (counterbalancing both scenario and task). See Table 1 for clarification of how conditions, VR height scenario, and task order were counterbalanced. Participants performed two "Experimental Stages", consisting of four counterbalanced task blocks.

<Insert Table 1 and Figure 1 about here>

Table 1.
Counterbalancing information, depicting how task and VR height were sequenced for each "Experimental Stage". All participants performed both tasks (response inhibition, inhibition) at both heights (high, low) for each Experimental Stage. Each participant was assigned to one of the four depicted orders for an Experimental Stage, and proceeded through their allocated sequence twice.

Assigned Order		Counterbalanced Sequence of Task and VR Height			
		1	**2**	**3**	**4**
1	**VR Height**	High	Low	High	Low
	Task	Inhibition	Inhibition	Updating	Updating
2	**VR Height**	Low	High	Low	High
	Task	Inhibition	Inhibition	Updating	Updating
3	**VR Height**	High	Low	High	Low
	Task	Updating	Updating	Inhibition	Inhibition
4	**VR Height**	Low	High	Low	High
	Task	Updating	Updating	Inhibition	Inhibition

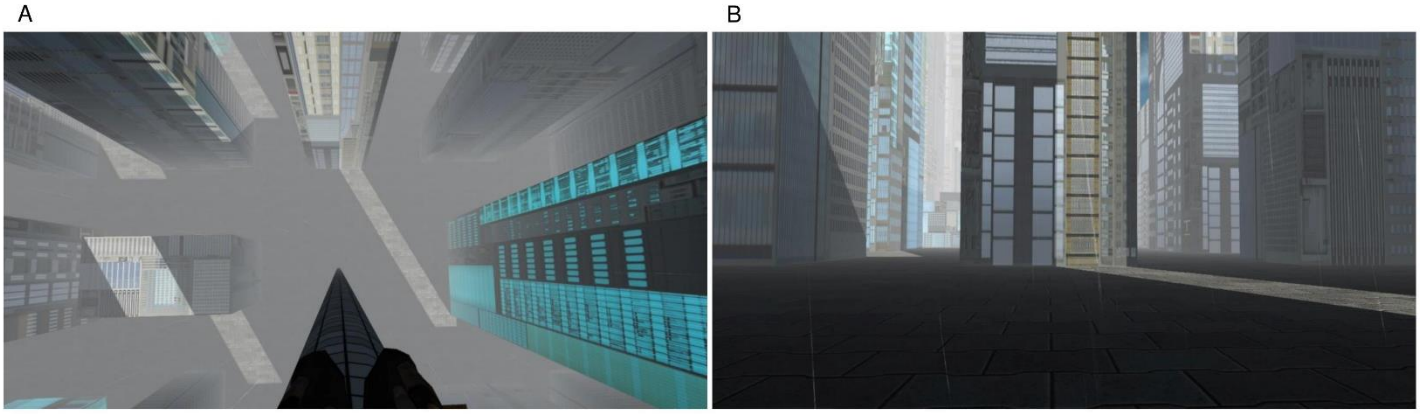

Figure 1. The virtual reality height manipulation scenarios, depicting the High VR height (Panel A) and the Low VR height (Panel B)

Behavioral measures. Behavioral measures were collected using Presentation software (Neurobehavioral Systems, Inc.), running on a separate computer, via a mouse that participants held on their lap.

Response inhibition task. Participants performed an auditory version of a Go/NoGo task. They were instructed to respond as fast as possible to the 'Go' stimulus (a 550-ms 900Hz sine tone) and withholding responses to the 'NoGo' stimulus (a 550-ms 1200Hz sine tone). The NoGo tone was perceptually higher in pitch compared to the Go stimulus, and all participants could reliably distinguish between the cues. Participants performed four ~3-minute blocks of the task within each experimental stage, with 200 trials per block (20% NoGo trials, 80% Go trials). Trial type and inter-stimulus interval (between 400-800-ms) were pseudorandomized, with trials lasting for 600-ms.

Updating task. Participants also performed an auditory 2-Back task to engage working memory updating processes. Stimuli were each presented for 1000-ms after which silence followed for another 1000-ms. Stimuli were spoken letters A-E, presented in a randomised order. Participants responded with a mouse-click whenever a letter was the same as that heard two previously ("2-back"). There were 125 trials per block, and targets occurred on approximately 25% of trials. Each trial lasted 2000-ms, and responses were counted if they occurred within 50-1800-ms of the beginning of the stimulus presentation which signaled the beginning of each trial. Each block of the task took approximately 3 minutes.

Physiological measures. Physiological measures comprising heart rate (HR) and heart rate variability (HRV) were collected to explore how exposure to the VR manipulation influenced physiological reactivity (i.e., whether the high height manipulation increased SNS activity by increased HR and decreased HRV). Heart rate and heart rate variability data were collected via a Polar Watch chest band (RS800CX), a method previously validated against standard ECG (Quintana, Heathers, & Kemp, 2012; Weippert et al., 2010). This method was used to facilitate cardiovascular recordings in the context of the VR set-up. Participants were instructed on attaching the chest band, and were able to do this in a private testing room. Sympathetic arousal (via HR) and parasympathetic modulation over heart rate (via HRV) were sampled at 1000Hz and were recorded continuously, with markers in the data indicating onsets/offsets of each VR scenario.

Self-report measures. To index participant perceptions of the VR heights, and to develop future experimental paradigms, participants used 1-9 Likert scales (1=Not at all, 5=Somewhat, 9=Very much so) to report their experience of the low and high VR conditions separately (e.g., reporting stressfulness, unpleasantness, and how much of a physical reaction they experienced). These measures indexed whether participants had any adverse reactions to our specific VR height manipulation, and thus a previously validated measure for this was unavailable. Participants responded to three questions indexing responses at the low (Cronbach's α=0.76) and the high VR height (Cronbach's α=0.91). For example, one question read "How **unpleasant** did you find being in the virtual reality *at the high level*?"

Heights Interpretation Questionnaire. To index participant negative appraisals of heights, participants completed the Heights Interpretation Questionnaire (Steinman & Teachman, 2011). This questionnaire asks participants to imagine two separate scenarios in which they are climbing up a high ladder alongside a house (8 items; Cronbach's α=0.89), or are on the balcony of the 15th floor of a building (8 items; Cronbach's α=0.92). Participants are asked to rate (on a Likert scale, 1=Not likely, 5=Very likely) how believable it is that they would perceive the situations as dangerous, and whether they would experience any physical or emotional consequences (i.e., one item on the questionnaire asks "How likely is it that you will panic and lose control". These ratings are then summed, such that higher scores indicate a higher negative appraisal of heights.

Depression, Anxiety and Stress Scales-21. To assess baseline levels of depression, anxiety and stress, the DASS-21 scale was utilized (Lovibond & Lovibond, 1995). The questionnaire has a subscale (7 items each) for the depression (Cronbach's α=0.77), anxiety (Cronbach's α=0.83) and stress (Cronbach's α=0.82), and has shown good reliability and internal consistency (Henry & Crawford, 2005; Lovibond & Lovibond, 1995).

Emotion regulation questionnaire. To index whether individuals with high versus low negative appraisals of heights differed in their habitual use of reappraisal and suppression, we utilized the Emotion Regulation Questionnaire (ERQ; Gross & John, 2003). The ERQ indexes the degree to which individuals assert using reappraisal (6 items; Cronbach's α=0.81) or suppression (4 items; Cronbach's α=0.80) emotion regulation strategies in their daily lives. This measure gives an indication as to how individuals habitually employ emotion regulation strategies across a variety of contexts and domains.

Procedure

After providing informed consent, participants completed the Heights Interpretation Questionnaire (Steinman & Teachman, 2011) to index negative appraisals of heights, the DASS-21 and the ERQ. Participants were fitted with the PolarWatch chest band, and baseline heart rate was collected during an initial 5-minute rest period. Participants were then situated in front of the computer on a standard desk chair, and were given 20 practice trials on each executive function task (i.e. the Go/NoGo task indexing response inhibition and the N-Back task indexing Updating). Following the practice trials they were given an opportunity to seek clarification about the tasks. They were then fitted with the head-mounted VR display and performed four blocks of each of the Response Inhibition and Updating tasks within alternating low- and high-threat VR scenarios. These tasks were completed in a counterbalanced order assigned via random allocation as per Table 1, with approximately 20 seconds rest between tasks.

Following completion of the behavioral tasks, participants were asked to provide self-report measures of their experiences within the scenarios. Participants were then thanked for participation and fully debriefed.

Data Coding

Behavioral data. The dependent variable for both tasks was a sensitivity index calculated from hit and false-alarm rates: d' = z(correct responses) – z(false alarms). This measure was utilized to facilitate comparison in task performance across the two executive function tasks, thereby fulfilling an aim of the study in comparing the impact of the VR height manipulation on performance of two executive functions tasks within the same analysis. Further, the d' metric has been utilized in previous work examining Go/NoGo and N-Back task performance (Casey et al., 2007; Haatveit et al., 2010; Pelegrina et al., 2015; Schulz et al., 2007). Five participants were excluded from analyses due to d' scores that were 2 SDs below the mean d', as this indicates that participants did not understand the task instructions or did not engage with the EF task properly.[1] Full behavioral data was not available for an additional 3 participants due to a technical malfunction of the behavioral recording system, resulting in a final sample of 50.

Physiological response data. Time markers denoting the start and finish of experimental task blocks were imported into Kubios (analysis software; Tarvainen, Niskanen, Lipponen, Ranta-Aho, & Karjalainen, 2014). Average heart rate and Root Mean Square of Successive Differences (RMSSD) indices were obtained for the baseline and experimental blocks. We used RMSSD to represent HRV, as it is less subject to confounding effects of respiration and is more reliable than frequency measures over shorter recording periods compared to frequency-based indices of HRV (Quintana et al., 2016).

[1] Although the below results detail analyses excluding these participants due to poor performance, the key findings remain unaltered in analyses including these participants.

Classification of negative appraisals of heights. To examine the key questions as to whether negative appraisals of heights would influence the physiological, affective and cognitive responses to a VR height manipulation, a median split (median=33) of the Heights Interpretation Questionnaire (Steinman & Teachman, 2011) was created. Participants were classified as having either High (*n*=26) or Low (*n*=24) Negative Appraisal of Heights (NAH). We used a median split approach to separate the high and low negative appraisals groups due to the within-subjects nature of the performance data, which served to help us understand how VR-induced threat impacted on performance of the response inhibition and updating tasks within a single analysis. This analysis was conducted to develop an understanding of how individual differences in appraisals influence performance across threatening and non-threatening contexts alike. However, to further aid the interpretation of how negative appraisals of heights were related to executive function performance, we conducted additional analyses of the data. We conducted regressions with performance in the first block of each task, at each height separately, as the dependent variable. These results are reported in the Supplementary Material; note that these analyses lead to similar conclusions to the outcomes presented here.

Results

Demographics

Mean sample characteristics are reported in Table 2. The high and low negative appraisals of heights groups did not differ significantly on key characteristics including gender, χ^2 (1,N=50) = 0.02, *p*=.96, or age, *t*(48)=1.93, *p*=.059.[2] Scores on the questionnaire were similar to those reported in Steinman and Teachman (2011), with a range of 17-62,

[2] Note that while groups do not differ significantly on age, the groups trend towards showing a difference in age. However, if age is factored into the main analyses as a covariate, the behavioral findings remain significant.

suggesting a good range of negative interpretations of heights within our randomly recruited sample.

Table 2

Mean participant characteristics

	High Negative Appraisals of Heights	Low Negative Appraisals of Heights	Sig.
Final Sample size	*n* = 26	*n* = 24	
Heights Interpretation	39.59 (7.07)	24.88 (5.08)	$p<.001$
Gender	15 F/11 M	14 F/10 M	$p=.96$[a]
Age	21.73 (5.75)	19.38 (1.66)	$p=.06$
Depression	2.92 (3.60)	2.33 (1.93)	$p=.48$
Anxiety	2.88 (4.07)	1.92 (1.93)	$p=.29$
Stress	4.77 (3.88)	4.42 (3.40)	$p=.74$
Trait Reappraisal	4.88 (0.85)	4.79 (1.11)	$p=.74$
Trait Suppression	3.95 (1.54)	3.80 (1.36)	$p=.72$

Note: Standard deviations appear in parentheses. F = Female; M = Male.
* Denotes a statistically significant outcome.
[a] Denotes a χ squared test.

Executive Function

Overall effects of VR Height Manipulation. To assess the overall impact of the VR Height Manipulation on EF, performance measures were submitted to an ANCOVA, with Negative Appraisals of Heights (High NAH versus Low NAH) as a between subjects-factor, and VR Height (High VR versus Low VR), Experimental Stage (Stage 1 versus Stage 2) and

Task (Response Inhibition versus Updating) as within-subjects factors. Further, covariates to control for order effects of the EF task and VR Height counterbalancing were included[3].

A significant Negative Appraisals of Heights x VR Height interaction emerged, $F(1,46)=6.30$, $p=.016$, $\eta_p^2=0.12$. Follow-up tests on the interaction demonstrated that, collapsing across executive function task, participants with high NAH demonstrated significantly impaired executive function compared to those with low NAH at the high VR height, $F(1,46)=9.09$, $p=.004$, $\eta_p^2=0.17$, but not at the low VR height, $F(1,46)=1.71$, $p=.20$, $\eta_p^2=0.04$. This interaction is displayed in Figure 2. Further, if intra-group comparisons are considered (rather than inter-group comparisons as above) on the interaction, an interesting picture emerged. Individuals with high NAH had numerically worse performance at the high VR height compared to the low VR height, $F(1,46)=2.94$, $p=.09$, $\eta_p^2=0.06$, but this did not reach the threshold for significance. However, the opposite pattern emerged for those individuals reporting low NAH, as they demonstrated marginally improved executive functions at the high VR height compared to the low VR height, $F(1,46)=3.72$, $p=.06$, $\eta_p^2=0.08$. This result might suggest a potential facilitatory effect of the high VR height for those with low negative appraisals of the situation, but does not reach the threshold for statistical significance and therefore must be interpreted with caution.

[3] This approach was taken to control for any unintended effects of the VR height and Task counterbalancing. However, if these covariates are removed from the analysis, the significant effects described remain unchanged. In addition, without covariates, a significant Experimental Stage x Task interaction emerges, $F(1,48)=8.99$, $p=.004$, $\eta_p^2=0.16$. Follow-up comparisons indicate that performance improves from Stage 1 to Stage 2 for both response inhibition, MD= -0.18, 95% CI [-0.30, -0.06], $p=.003$, and updating, MD= -0.51, 95% CI [-0.68, -0.33], $p=<001$, measures, with the significant interaction indicating that the magnitude of the improvement is larger for updating than response inhibition.

This interaction clarified a significant main effect of Negative Appraisals of Heights, $F(1,46)=5.23$, $p=.027$, $\eta_p^2=0.10$, whereby – overall – participants with high NAH (M=2.59, S=0.11) performed significantly worse than participants with low NAH (M=2.97, SE=0.12). A significant main effect of Experimental Stage, $F(1,46)=15.43$, $p<.001$, $\eta_p^2=0.25$, revealed a significant improvement in performance from Stage 1 (M=2.61, SE=0.08) to Stage 2 (M=2.95, SE=0.09). The main effect of Task or any interactions involving Task did not reach significance (all *p*'s>.16), nor did any three-way interactions (all *p*'s>.35) or the four-way interaction, $F(1,46)=2.59$, $p=.11$, $\eta_p^2=0.05$. Given there were no significant effects of Task, the effects of Task were not broken down in additional primary analyses, so further analysis of this factor is restricted to presentation within the Supplementary Information. Further, analyses breaking down the impact of the VR height manipulation on executive function by stage, thereby investigating whether informal participant self-reports of habituation to the scenarios were reflected in executive performance, are reported in the

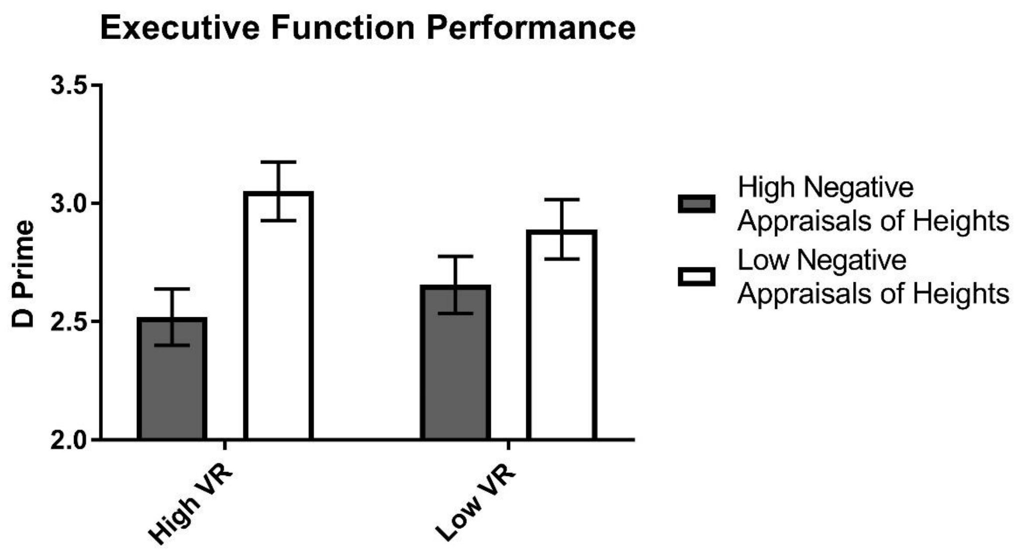

Figure 2. Performance in d' on response inhibition and updating tasks combined across the entire experiment, based on the adjusted means controlling for covariates. Error bars are standard error of the mean.

<Insert Figure 2 around here>

Manipulation Check: Virtual Reality-Induced Threat

Self-report. To assess whether the VR height manipulation impacted high and low NAH participants' subjective report of distress differently, subjective responses regarding experienced stress, unpleasantness, and physical discomfort were summed to create an aggregate score for ratings of the high and low VR heights separately. The approach was also undertaken due to the high internal consistency demonstrated by the items assessing responses to the high and low VR heights, as reported in the Methods (low VR height: Cronbach's α=0.76; high VR height: Cronbach's α=0.91). Surprisingly — given that an interaction between NAH and VR height emerged in behavioral performance — a 2 (NAH) X 2 (VR height) ANOVA revealed no such interaction for subjective report of distress, $F(1,48)=1.76$, $p=.19$, $\eta_p^2=0.04$. This may indicate that threat-related appraisals impact cognitive performance even in the absence of differences in subjective distress, or when any differences in affective state are below the threshold for explicit awareness. Such an interpretation must remain tentative, however: when comparing high and low NAH participants at each VR height separately, those with high NAH rated the high VR height as significantly more distressing than did participants with low NAH, $t(48)=2.09$, $p=.042$, $d=0.60$, (high NAH: M= 12.35, SE=1.11; low NAH: M=9.00 , SE=1.15). However, at the low VR height no such difference in distress ratings emerged (high NAH: M= 6.35, SE=0.68; low NAH: M=4.92, SE=0.70; $t(40.07)=1.49$, $p=.14$, $d=0.42$. Thus, it may simply be that the current experiment did not have enough power to reveal differential subjective distress between high and low NAH participants as a function of VR height.

Physiological responses. To control for individual differences in HR and HRV, we entered the HR and HRV indices from the baseline as co-variates in an ANCOVA. Because the VR height manipulation was counterbalanced (i.e., low-high vs. high-low), the high-low/low-high order of conditions was entered into the ANCOVA as an additional covariate. Thus, mixed ANCOVAs were conducted for HR and HRV measures with between-subjects factor of Negative Appraisals of Heights (High NAH versus Low NAH), and VR Height (High VR versus Low VR) and Experimental Stage (Stage 1 versus Stage 2) as within-subjects factors.

For HR during the experimental blocks, there were no significant main effects or interactions: the Negative Appraisals of Heights x VR Height interaction did not reach significance, $F(1,46)=2.70$, $p=.11$, $\eta_p^2=0.06$, nor did any other main or interaction effects involving NAH (all other p's>.25). No further main or interaction effects involving VR Height or Experimental Stage were detected (all p's>.58).

For HRV, a significant main effect of Experimental Stage, $F(1,46)=11.34$, $p=.002$, $\eta_p^2=0.20$, indicated that, regardless of participant NAH, HRV decreased from Stage 1 (M=33.64, SE=1.94) to Stage 2 (M=32.22, SE=1.53). A significant VR Height x Experimental Stage interaction, $F(1,46)=15.87$, $p<.001$, $\eta_p^2=0.26$, was detected, but follow-up comparisons on this interaction were not significant (all p's>.66). There were no significant main or interaction effects involving NAH (all p's>.23).

Physiological responding during initial exposure to the VR. An additional question pertains to whether physiological responding was specifically altered during the *first* exposure to the high VR height manipulation. It may be that the physiological effects of the VR were washed out when analyzing across both Stage 1 and 2 in the mixed ANCOVAs above. To additionally probe if the initial high VR height manipulation resulted in increased HR or decreased HRV relative to baseline measures, a mixed ANOVA was conducted with Negative Appraisals of Heights (High NAH versus Low NAH) as a between-subjects variable and Time (Baseline Measure versus First High VR Height Measure) on indices of heart rate and heart rate variability.

When considering heart rate, a trend effect of Time was detected, $F(1,48)=3.52$, $p=.067$, $\eta_p^2=0.07$, suggesting that regardless of participant negative appraisals of heights, initial exposure to the high VR height (M=85.83, SE=1.98) resulted in marginally higher heart rate compared to baseline (M=83.30, SE=1.83). There was no significant Negative Appraisals of Heights main effect, $F(1,48)=1.65$, $p=.21$, $\eta_p^2=0.03$, or Negative Appraisals of Heights x Time interaction, $F(1,48)<1$.

In analyses pertaining to heart rate variability, a significant main effect of Time was detected, $F(1,48)=4.14$, $p=.048$, $\eta_p^2=0.08$. This effect indicated that overall, participant HRV decreased significantly from baseline (M=39.99, SE=4.43) to initial exposure to the high VR height (M=33.70, SE=2.84). No significant Negative Appraisals of Heights main effect, $F(1,48)=2.68$, $p=.11$, $\eta_p^2=0.05$, or Negative Appraisals of Heights x Time interaction, $F(1,48)<1$, was detected. These analyses suggest that the high VR height prompted small alterations in parasympathetic engagement during the first block of VR height exposure.

Discussion

The current study employed virtual reality (VR) to assess how immersion within an environment that one appraises to be dangerous impacts concurrent executive function performance. Results revealed that executive function (EF) was impaired when participants were placed at a virtual height, but only for those with self-reported negative appraisals of heights. Specifically, EF was impaired during exposure to the virtual height for those with elevated negative appraisals of heights, but EF was not impaired during a comparable low-threat ground-height manipulation. Additionally, analyses including both of the tested EF tasks (updating and response inhibition) suggested that the VR-induced performance deficits were not specific to one of these EFs over the other. Self-report data also indicated that individuals high in negative appraisals of heights found the VR height manipulation to be more distressing compared to individuals low in negative appraisals of heights. Importantly, the current study demonstrated that VR can be used to understand real-time performance in life-like situations, providing a foundation on which to build understanding of strategies that may be used to protect cognitive performance during threatening- or anxiety-eliciting situations.

Overall, the findings indicate that individual differences in the propensity to appraise a situation as threatening or stressful are important to consider in experimental inductions, when examining the effects of a stressor or threat on cognitive performance. In particular, whereas negative appraisals of the scenario were associated with impaired executive function relative to those who did not appraise it negatively, performance among those who did not appraise heights negatively was suggestive of a facilitation effect: these individuals demonstrated marginally better executive function at the high- relative to the low- height, although given this finding is marginal, it should be interpreted with caution. This is intriguing because it highlights the possibility that an arousing situation might either impair or enhance cognitive performance depending on whether the situation appraised to be negative or not. The finding that performance was modulated by individual differences in negative appraisals of heights converges with work demonstrating a relationship between performance decrements following a stressor and stress-induced cortisol increases. In these prior studies, greater sensitivity to a stress induction (reflected in greater cortisol increases) was associated with greater cognitive impairments (Plessow, Schade, Kirschbaum, & Fischer, 2012; Schoofs et al., 2008).

The findings reported here accord with an abundance of evidence suggesting that executive function is impaired *following* acute stress (i.e., see Shields et al., 2016), and further this literature in indicating that executive function can be impaired *during* arousing negative experiences. However, our finding that threat does not differentially impair response inhibition and updating are at odds with a recent meta-analysis indicating differential effects of stress on these two components of executive function (Shields et al., 2016). In that meta-analysis, stress was found to significantly impair updating, but enhance response inhibition. This divergence in findings was unexpected but may be attributable to methodological factors, such as the assessment of EF during, as opposed to after, threat exposure. In line with this, research suggests that HPA axis activation, in tandem with activation of the SNS, is needed to observe stress-induced alterations in executive functions (Alexander, Hillier, Smith, Tivarus, & Beversdorf, 2007; Elzinga & Roelofs, 2005; Henderson, Snyder, Gupta, & Banich, 2012). Because the assessment of executive functions took place during exposure to the threat in the current study, there may not have been sufficient time for any activation of the HPA axis to take place.

Although we did not find significant effects of the VR manipulation on changes in HR and HRV, this may have been due to averaging across longer periods of whole experimental task blocks (taken over the entire VR exposure), thereby obscuring any potential initial impact of the VR scenario on physiological reactivity. Our results provide some indication that initial exposure to the VR altered physiological arousal in some participants relative to baseline (as reflected in numerical increases in HR and decreases in HRV). However, given that these effects are on the edge of significance with a small effect size, they should be interpreted with caution. Nonetheless, this aligns with the content of informal discussions with participants following a full debriefing on the task, who tended to report that they initially experienced physiological arousal or distress, but that this quickly dissipated. Given that participants were aware of the VR height manipulation prior to baseline due to the informed consent process, this knowledge may have induced an anticipatory threat response, thereby impacting physiological reactivity and executive function prior to any formal measurements. However, an anticipatory threat response due to the informed consent process is unlikely to solely account for these findings, particularly given the specificity of the influence of the VR manipulation on performance in the high VR height versus low VR height for those with high negative appraisals of heights. Further, the marginal increases in heart rate from baseline (which was taken following informed consent and a habituation period) to initial exposure to the high VR height suggest that any anticipatory responses would have been limited in their effects or duration.

Some limitations of the current investigation must be noted. We used a median split analysis to probe the relationship between appraisals of heights and our key cognitive, physiological and affective outcomes. However, we do note the limitations of this approach, most notably a loss of statistical power and an increase in the risk of false positive results (Altman & Royston, 2006). Future work investigating the impact of threat on concurrent performance may wish to adopt more advanced statistical or modelling techniques. In addition, participants reported habituating to the VR height manipulation across the experiment, which may have impacted participant's executive functions differently across the experiment. This could be rectified in future work by incorporating more varied or intense VR scenarios. The apparently rapid habituation to the height manipulation meant that its impact on executive performance (and its modulation by negative appraisals of height) was most prominent in the first stage of the experiment. Additionally, the executive tasks in the current study were presented in the auditory modality so as not to disrupt visual processing of the VR scenarios. While evidence suggests that performance in auditory and visual executive tasks are positively correlated (Roberts, Summerfield, & Hall, 2006; Spagna, Mackie, & Fan, 2015), our results warrant replication in other modalities.

Although real-time cognitive performance in a threatening context is a common challenge in the real world, this characteristic is not well represented in commonly used laboratory threat- or stress-inductions. VR poses an empirical opportunity for researchers to examine attention, memory and affective phenomena in realistic settings, without comprising the rigour of controlled lab environments. In revealing impaired cognitive ability during a real-time situation that is appraised as threatening, the current study underscores the importance of psychological factors (e.g., appraisals) in determining when and why certain contexts lead to cognitive failures.

Declarations

Ethics Approval and Consent to Participate

Studies reported in this manuscript were approved by the UNSW Sydney Human Research Ethics Advisory Panel, Approval Number 2474. All participants gave informed consent in line with the Declaration of Helsinki.

Availability of Data and Materials

The complete dataset supporting the conclusions of this article here is available in the OSF repository, available online at https://osf.io/gf7qe.

Declarations of Interest

None.

Funding

This work was supported by an Australian Government Research Training Program Scholarship to VEN and Australian Research Council grant FT120100707 to SBM. These funding bodies had no role in the study design, data collection, analysis and interpretation of data, in the writing of the report or decision to submit the article for publication.

Authors Contributions

All authors approved the final manuscript.

Acknowledgements

This work was supported by an Australian Government Research Training Program Scholarship to VEN and Australian Research Council grant FT120100707 to SBM. These funding bodies had no role in the study design, data collection, analysis and interpretation of data. We would like to thank Adam Halley-Prinable for creating our virtual scenarios. The experiments reported here were conducted as part of the first author's PhD thesis at UNSW

Sydney. The complete dataset supporting the conclusions of this article here is available in the OSF repository, available online at https://osf.io/gf7qe.

Figure Captions

Figure 1. The virtual reality height manipulation scenarios, depicting the High VR height (Panel A) and the Low VR height (Ground height, Panel B).

Figure 2. Performance in d' on response inhibition and updating tasks combined across the entire experiment, based on the adjusted means controlling for covariates. Error bars are standard error of the mean.

References

Alexander, J. K., Hillier, A., Smith, R. M., Tivarus, M. E., & Beversdorf, D. Q. (2007). Beta-adrenergic Modulation of Cognitive Flexibility during Stress. *Journal of Cognitive Neuroscience, 19*(3), 468-478. doi: 10.1162/jocn.2007.19.3.468

Allen, A. P., Kennedy, P. J., Dockray, S., Cryan, J. F., Dinan, T. G., & Clarke, G. (2017). The Trier Social Stress Test: Principles and practice. *Neurobiology of Stress, 6*(Supplement C), 113-126. doi: https://doi.org/10.1016/j.ynstr.2016.11.001

Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *BMJ : British Medical Journal, 332*(7549), 1080-1080.

Arnsten, A. F. T. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nat Rev Neurosci, 10*(6), 410-422.

Banich, M. T. (2009). Executive function: The search for an integrated account. *Current Directions in Psychological Science, 18*(2), 89-94. doi: 10.1111/j.1467-8721.2009.01615.x

Blascovich, J., & Mendes, W. B. (2010). Social psychophysiology and embodiment *Handbook of social psychology, Vol. 1, 5th ed.* (pp. 194-227). Hoboken, NJ, US: John Wiley & Sons Inc.

Casey, B. J., Epstein, J. N., Buhle, J., Liston, C., Davidson, M. C., Tonev, S. T., . . . Glover, G. (2007). Frontostriatal Connectivity and Its Role in Cognitive Control in Parent-Child Dyads With ADHD. *American Journal of Psychiatry, 164*(11), 1729-1736. doi: 10.1176/appi.ajp.2007.06101754

Crişan, L. G., Vulturar, R., Miclea, M., & Miu, A. C. (2016). Reactivity to Social Stress in Subclinical Social Anxiety: Emotional Experience, Cognitive Appraisals, Behavior, and Physiology. *Frontiers in Psychiatry, 7*, 5. doi: 10.3389/fpsyt.2016.00005

Clark, D. A., & Beck, A. T. (2010). Cognitive therapy of anxiety disorders: Science and practice. New York: Guildford Press.

Diamond, A. (2013). Executive functions. *Annual review of psychology, 64*, 135-168. doi: 10.1146/annurev-psych-113011-143750

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin, 130*(3), 355-391. doi: 10.1037/0033-2909.130.3.355

Elzinga, B. M., & Roelofs, K. (2005). Cortisol-Induced Impairments of Working Memory Require Acute Sympathetic Activation. *Behavioral neuroscience, 119*(1), 98-103. doi: 10.1037/0735-7044.119.1.98

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion. 7*(2), 336-353.

Fatisson, J., Oswald, V., & Lalonde, F. (2016). Influence diagram of physiological and environmental factors affecting heart rate variability: an extended literature overview. *Heart Int, 11*(1), e32-e40. doi: 10.5301/heartint.5000232

Gaab, J., Rohleder, N., Nater, U. M., & Ehlert, U. (2005). Psychological determinants of the cortisol stress response: the role of anticipatory cognitive appraisal. *Psychoneuroendocrinology, 30*(6), 599-610. doi: http://dx.doi.org/10.1016/j.psyneuen.2005.02.001

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*, 348-362.

Haatveit, B. C., Sundet, K., Hugdahl, K., Ueland, T., Melle, I., & Andreassen, O. A. (2010). The validity of d prime as a working memory index: results from the "Bergen n-back task". *J Clin Exp Neuropsychol, 32*(8), 871-880. doi: 10.1080/13803391003596421

Hayano, J., Yamada, M., Sakakibara, Y., Fujinami, T., Yokoyama, K., Watanabe, Y., & Takata, K. (1990). Short- and long-term effects of cigarette smoking on heart rate variability. *The American Journal of Cardiology, 65*(1), 84-88. doi: https://doi.org/10.1016/0002-9149(90)90030-5

Henderson, R. K., Snyder, H. R., Gupta, T., & Banich, M. T. (2012). When does stress help or harm? The effects of stress controllability and subjective stress response on Stroop performance. *Frontiers in Psychology, 3*, 179. doi: 10.3389/fpsyg.2012.00179

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): construct validity and normative data in a large non-clinical sample. *Br J Clin Psychol, 44*(Pt 2), 227-239. doi: 10.1348/014466505x29657

Hermans, E. J., Henckens, M. J. A. G., Joëls, M., & Fernández, G. (2014). Dynamic adaptation of large-scale brain networks in response to acute stressors. *Trends in Neurosciences, 37*(6), 304-314. doi: http://dx.doi.org/10.1016/j.tins.2014.03.006

Joormann, J., Waugh, C. E., & Gotlib, I. H. (2015). Cognitive Bias Modification for Interpretation in Major Depression: Effects on Memory and Stress Reactivity. *Clin Psychol Sci, 3*(1), 126-139. doi: 10.1177/2167702614560748

Kapfhammer, H.-P., Huppert, D., Grill, E., Fitz, W., & Brandt, T. (2015). Visual height intolerance and acrophobia: clinical characteristics and comorbidity patterns. *European Archives of Psychiatry and Clinical Neuroscience, 265*(5), 375-385. doi: 10.1007/s00406-014-0548-y

Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test' – A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology, 28*(1-2), 76-81.

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states-Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour research and therapy, 33*(3), 335-343.

Lupien, S. J., Maheu, F., Tu, M., Fiocco, A., & Schramek, T. E. (2007). The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. *Brain and Cognition, 65*(3), 209-237. doi: http://dx.doi.org/10.1016/j.bandc.2007.02.007

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21*(1), 8-14. doi: 10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49-100. doi: http://dx.doi.org/10.1006/cogp.1999.0734

Pelegrina, S., Lechuga, M. T., García-Madruga, J. A., Elosúa, M. R., Macizo, P., Carreiras, M., . . . Bajo, M. T. (2015). Normative data on the n-back task for children and young adolescents. *Frontiers in Psychology, 6*(1544). doi: 10.3389/fpsyg.2015.01544

Plessow, F., Fischer, R., Kirschbaum, C., & Goschke, T. (2011). Inflexibly focused under stress: Acute psychosocial stress increases shielding of action goals at the expense of reduced cognitive flexibility with increasing time lag to the stressor. *Journal of Cognitive Neuroscience, 23*(11), 3218-3227. doi: 10.1162/jocn_a_00024

Plessow, F., Schade, S., Kirschbaum, C., & Fischer, R. (2012). Better not to deal with two tasks at the same time when stressed? Acute psychosocial stress reduces task shielding in dual-task performance. *Cognitive, Affective, & Behavioral Neuroscience, 12*(3), 557-570. doi: 10.3758/s13415-012-0098-6

Quintana, D. S., Alvares, G. A., & Heathers, J. A. J. (2016). Guidelines for Reporting Articles on Psychiatry and Heart rate variability (GRAPH): recommendations to advance research communication. *Translational Psychiatry, 6*(5), e803. doi: 10.1038/tp.2016.73

Quintana, D. S., Heathers, J. A. J., & Kemp, A. H. (2012). On the validity of using the Polar RS800 heart rate monitor for heart rate variability research. *European Journal of Applied Physiology, 112*(12), 4179-4180. doi: 10.1007/s00421-012-2453-2

Roberts, K. L., Summerfield, A. Q., & Hall, D. A. (2006). Presentation modality influences behavioral measures of alerting, orienting, and executive control. *J Int Neuropsychol Soc, 12*(4), 485-492.

Robinson, O. J., Bond, R. L., & Roiser, J. P. (2015). The impact of stress on financial decision-making varies as a function of depression and anxiety symptoms. *PeerJ, 3*, e770. doi: 10.7717/peerj.770

Roos, L. E., Knight, E. L., Beauchamp, K. G., Berkman, E. T., Faraday, K., Hyslop, K., & Fisher, P. A. (2017). Acute stress impairs inhibitory control based on individual differences in parasympathetic nervous system activity. *Biological Psychology, 125*(Supplement C), 58-63. doi: https://doi.org/10.1016/j.biopsycho.2017.03.004

Rozenman, M., Vreeland, A., & Piacentini, J. (2016). Thinking anxious, feeling anxious, or both? Cognitive bias moderates the relationship between anxiety disorder status and sympathetic arousal in youth. *Journal of Anxiety Disorders, 45*, 34-42.

Schoofs, D., Pabst, S., Brand, M., & Wolf, O. T. (2013). Working memory is differentially affected by stress in men and women. *Behavioural Brain Research, 241*, 144-153. doi: http://dx.doi.org/10.1016/j.bbr.2012.12.004

Schoofs, D., Preuß, D., & Wolf, O. T. (2008). Psychosocial stress induces working memory impairments in an n-back paradigm. *Psychoneuroendocrinology, 33*(5), 643-653. doi: https://doi.org/10.1016/j.psyneuen.2008.02.004

Schulz, K. P., Fan, J., Magidina, O., Marks, D. J., Hahn, B., & Halperin, J. M. (2007). Does the emotional go/no-go task really measure behavioral inhibition?: Convergence with measures on a non-emotional analog. *Archives of Clinical Neuropsychology, 22*(2), 151-160. doi: https://doi.org/10.1016/j.acn.2006.12.001

Seery, M. D. (2013). The Biopsychosocial Model of Challenge and Threat: Using the Heart to Measure the Mind. *Social and Personality Psychology Compass, 7*(9), 637-653. doi: 10.1111/spc3.12052

Shields, G. S., Sazma, M. A., & Yonelinas, A. P. (2016). The effects of acute stress on core executive functions: A meta-analysis and comparison with cortisol. *Neuroscience & Biobehavioral Reviews, 68*, 651-668. doi: http://dx.doi.org/10.1016/j.neubiorev.2016.06.038

Simmonds, D. J., Pekar, J. J., & Mostofsky, S. H. (2008). Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. *Neuropsychologia, 46*(1), 224-232. doi: 10.1016/j.neuropsychologia.2007.07.015

Spagna, A., Mackie, M.-A., & Fan, J. (2015). Supramodal executive control of attention. *Frontiers in Psychology, 6*(65). doi: 10.3389/fpsyg.2015.00065

Steinman, S. A., & Teachman, B. A. (2011). Cognitive processing and acrophobia: Validating the Heights Interpretation Questionnaire. *J Anxiety Disord, 25*(7), 896-902. doi: https://doi.org/10.1016/j.janxdis.2011.05.001

Tarvainen, M. P., Niskanen, J. P., Lipponen, J. A., Ranta-Aho, P. O., & Karjalainen, P. A. (2014). Kubios HRV--heart rate variability analysis software. *Comput Methods Programs Biomed*, *113*(1), 210-220. doi: 10.1016/j.cmpb.2013.07.024

Villada, C., Hidalgo, V., Almela, M., & Salvador, A. (2016). Individual Differences in the Psychobiological Response to Psychosocial Stress (Trier Social Stress Test): The Relevance of Trait Anxiety and Coping Styles. *Stress and Health, 32*(2), 90-99. doi: 10.1002/smi.2582

Weippert, M., Kumar, M., Kreuzfeld, S., Arndt, D., Rieger, A., & Stoll, R. (2010). Comparison of three mobile devices for measuring R–R intervals and heart rate variability: Polar S810i, Suunto t6 and an ambulatory ECG system. *European Journal of Applied Physiology*, *109*(4), 779-786. doi: 10.1007/s00421-010-1415-9

Wolf, O. T. (2003). HPA axis and memory. *Best Practice & Research Clinical Endocrinology & Metabolism*, *17*(2), 287-299. doi: https://doi.org/10.1016/S1521-690X(02)00101-X

Zanstra, Y. J., & Johnston, D. W. (2011). Cardiovascular reactivity in real life settings: Measurement, mechanisms and meaning. *Biological Psychology, 86*(2), 98-105. doi: 10.1016/j.biopsycho.2010.05.002