# Domain-specific working memory loads selectively increase negative interpertations of surprised facial expressions

Nicholas R. Harp[1] & Maital Neta[1]

[1] University of Nebraska-Lincoln

## Author Note

Nicholas R. Harp, Department of Psychology, Center for Brain, Biology, and Behavior, University of Nebraska-Lincoln Maital Neta, Department of Psychology, Center for Brain, Biology, and Behavior, University of Nebraska-Lincoln

Correspondence concerning this article should be addressed to Nicholas R. Harp, Postal address. E-mail: nharp@huskers.unl.edu

## Abstract

Individual differences in interpretations of emotional ambiguity are a useful tool for measuring affective biases.

While trait-like, these biases are also susceptible to experimental manipulations. In the present study, we capitalize on this malleability to expand on previous research suggesting that subjective interpretations are stable independently of cognitive load.

We tested the effects of working memory loads containing either neutral or emotional content on concurrent interpretations of surprised facial expressions.

Here we show that interpretations of surprise are more negative during maintenance of working memory loads with emotional content compared to those with neutral content.

Two or three sentences explaining what the **main result** reveals in direct comparison to what was thought to be the case previously, or how the main result adds to previous knowledge.

One or two sentences to put the results into a more **general context**.

Two or three sentences to provide a **broader perspective**, readily comprehensible to a scientist in any discipline.

*Keywords:* ambiguity, working memory, bias

Word count: X

# Domain-specific working memory loads selectively increase negative interpertations of surprised facial expressions

## Introduction

Completion of even simple tasks in daily life require regular recruitment of working memory, and this often happens without consideration of how these working memory demands may affect concurrent processes. Many types of information can be held in working memory: strings of numbers, neutral and emotional images, stimulus characteristics like brightness, spatial locations, and more. There is debate in the literature around the neurophysiological substrates of working memory and whether this construct is a unified system or comprised of several domain-specific subsystems (see Baddeley, 1986; Shah & Mikaye, 1996). Recent work suggests though that working memory does indeed contain domain-specific subsystems (e.g., affective working memory; see Mikels and Reuter-Lorenz (2019))–or at least that interactions between working memory and visual or affective tasks are variable depending on working memory content. Similar domain-specific effects have been observed in a number of other lines of research, including those outside of the working memory domain. For example, the Stroop task (Stroop, 1935) has been modified to include emotional stimuli (Whalen, Bush, Shin, & Rauch, 2006) which has pronounced effects when the emotional words are population specific (e.g., trauma words in a PTSD sample). Further, neuroimaging work supports the notion that separate systems handle attentional biasing for domain-specific (emotional vs. non-emotional) task relevancy (Egner, Etkin, Gale, & Hirsch, 2008). The present work aims to clarify the interaction of working memory demands on a concurrent and ecologically valid emotionoal task (i.e., interpretations of emotional facial expressions), specifically when that content varies in both amount and affective quality.

There is a large body of research concerning the effects of distractor and task irrelevant stimuli on performance in a variety of tasks [cite, cite]. Notably, domain-specific interference may further exacerbate these effects compared to domain-general stimuli (Gruber, 2001).

Within an affect-specific context, the literature reports conflicting results for the role of affective stimuli on working memory processing (see (**???**) and (**???**)), with some suggesting ehnaced recall for emotional stimuli and others reporting impairment in working memory processes. The relationship between these systems is bidirectional though, and often times working memory affects affective processing. One such way is through dampening emotional responses, and converging work suggests that cognitive demands can mitigate negative emotional effects (**???**). One explanation for mitigated emotional responding during concurrent working memory demands is load theory. This theory posits that cognitive resources are finite, thus limiting the availability for simultaneous processing under large task demands (Lavie, Hirst, Fockert, & Viding, 2004). One limitation to this previous work is that the affective processes under study involved clearly interpretable stimuli/feelings. A more malleable affective process, such as interpreting ambiguously valenced emotional facial expressions, may offer better understanding of how concurrent working memory demands affect emotional processes.

Facial expressions are rich with affective information, and correctly interpreting these social cues is critical for successfully navigating the social world. Often, facial expressions serve as a clear social signal, but this is not always the case. While a smile likely expresses a positive affective state, other cues are not so clear. For instance, a surprised expression could signal either a positive (e.g., winning the lottery) or negative (e.g., seeing a snake in the woods) affective state in the expresser. When contextual information is limited, individuals differ in their tendency to interpret surprised facial expressions as positive or negative. Importantly, this affective bias extends beyond facial expressions, as individuals often show a similar bias to surprised faces as they do for ambiguous scenes (Neta, Kelley, & Whalen, 2013) or even words (Harp, Petro, Brown & Neta, in prep). This bias towards positive or negative interpretations is known as one's valence bias and myriad factors contribute to one's bias.

Both bottom-up (e.g., perceptual input) and top-down (e.g., stereotypes) processes are recruited for interpretation of facial expressions, and a growing body of work suggests that the initial interpretation of emotionally ambiguous stimuli is negative and driven by bottom-up processes. For instance, reaction times are faster for negative interpretations of ambiguous stimuli (just faces??) (Neta & Tong, 2016). Additionally, presentation of surprised facial expressions as low spatial frequency images, which is processed more readily than high spatial frequency images, biased interpretations towards negativity (**???**). Consequently, under this framework, arriving at a positive interpretation requires additional, top-down regulatory processes and there is evidence to support this as well. For example, forcing participants to slow their responding during interpretations of ambiguous images shifts individuals' biases towards positivity (Neta, Tong, & Henley, 2018). Neuroimaging evidence supports this initial negativity hypothesis as well, more positive individuals show higher levels of BOLD activation in brain regions recruited during emotion regulation (**???**). In short, slowing response time allows individuals to better regulate the potentially negative information in surprised expressions and to see it in a more positive light.

Given that a regulatory mechanism likely contributes to positive interpretations of surprised facial expressions, domain-specific interference may cause more negative interpretations of ambiguity compared to a more domain-general interference. Mattek and colleagues (2016) recently showed that different levels of cognitive load (i.e., holding either a single or seven digit number in working memory) does not affect subjective interpretations of surprised facial expressions, but that high cognitive loads do mitigate mouse trajectories. While the authors interpret this as a distinction between trait-like biases and dynamic cognitive-motor processes, there may be more domain-specific processes (e.g., emotional components) that span across these two measures of valence bias. Given the task irrelevance of the numeric distractors in Mattek and colleagues' (2016) work, it follows that the resources required for interpreting ambiguity as positive (Neta, Norris, & Whalen, 2009) may not have been recruited for working memory maintenance, and thus no change in subjective ratings was observed.

In the present study, we aim to test the effects of low and high working memory loads in both emotional and neutral domains. We expect that trials in which participants are maintaining an emotional working memory load will be more negative than neutral trials. Further, we predict that higher working memory laod trials, specifically in the emotional domain, will result in even more exaggerated negative interpretations.

Recent work suggests that ambiguity resolution in this context requires more cognitive resources/processing compared to clearly valenced faces (Mattek et al., 2016; Neta & Tong, 2016). The valence bias is trait-like (Neta et al., 2009) and generalizes to non-face stimuli (Neta et al., 2013); however, it is also malleable and may differ depenending on experimental manipulations, including stress inductions or instructions to slow responding (Brown, Raio, & Neta, 2017; Neta & Tong, 2016). Importantly, the valence bias relates to behavior outside of the laboratory; specifically, it is known to relate to depressive symptomology (**???**), at least in children. Chronic negativity biases are common in numerous psychopathologies, including depression and anxiety (**???**).

# Methods

## Participants

Fifty-eight subjects were recruited from the undergraduate research pool at the University of Nebraska-Lincoln. The data from eight subjects were excluded due to technical difficulties resulting from an error in one of the experiment scripts. This left 50 individuals in the final sample for analysis. The mean age of the remaining sample was 18.82 (1.19), a majority of participants were female (82.00%), and all were white/caucasian without hispanic/Latinx ethnicity. All subjects provided written informed consent in accordance with the Declaration of Helsinki and all procedures were approved by the University of Nebraska-Lincoln Institutional Review Board (Approval #20141014670EP). Each participant received course credit for completing the study.

## Material

**Stimuli.** The stimuli included faces from the NimStim (Tottenham et al., 2009) and Karolinska Directed Emotional Faces (Lundqvist, Flykt, & Öhman, 1998) stimuli sets, as in previous work (**???**; Brown et al., 2017). The faces consisted of 34 unique identities including 11 angry, 12 happy, and 24 surprised expressions organized pseudorandomly. The scene stimuli were selected from the International Affective Picture System (Lang, Bradley, & Cuthbert, 2008). A total of 288 scenes (72 positive, 72 negative, and 144 neutral) were selected for the image matrices. The positive and negative images did not differ on arousal ($Z = -0.23$, $p = 0.82$). The scenes were organized into low (two images) and high (six images) cognitive load of either neutral or emotional (equal number of positive and negative) images (Figure 1).

## Procedure

After arriving at the lab, participants provided informed consent prior to completing the task. Participants were randomly assigned to complete one of the task versions, which included 144[1] trials split between working memory probe and face rating trials. The task was completed using MouseTracker software (Freeman & Ambady, 2010) and participants responded with a mouse to indicate the appropriate response for the face ratings (i.e., “POSITIVE” or “NEGATIVE”) and the memory probe (i.e., “YES” or “NO”). The trials were self-initiated; that is, the participant clicked a “start” button at the bottom of the screen at the beginning of each trial at their own pace. After initiating the trial, a fixation cross appeared (1000 ms), then participants viewed an image matrix, which the participants were instructed to remember for the duration of the trial. The image matrix was presented for 4000 ms and the image was either a low or high load matrix consisting of either

[1] Some versions of the task only included 142 trials due to a programming error.

emotional (equal positive and negative) or neutral images. After the image matrix a happy, angry, or surprised face appeared for 1000 ms and the participants rated the face by clicking on either the positive or negative response option. After the face rating, a single image probe appeared (5000 ms), and participants indicated whether or not the image probe was present in the previous image matrix.

## Data analysis

We used R (Version 3.6.0; **???**) and the R-packages * }dplyr* [@ }R-dplyr], *BayesFactor* (Version 0.9.12.4.2; **???**), *broom* (Version 0.5.2; **???**), *circlize* (Version 0.4.6; **???**), *coda* (Version 0.19.2; **???**), *cstab* (Version 0.2.2; **???**), *diptest* (Version 0.75.7; **???**), *dotCall64* (Version 1.0.0; **???**; **???**), *fastcluster* (Version 1.1.25; **???**), *fields* (Version 9.8.3; **???**), *forcats* (Version 0.4.0; **???**), *foreach* (Version 1.4.7; **???**), *ggplot2* (Version 3.1.1; **???**), *jpeg* (Version 0.1.8; **???**), *lattice* (Version 0.20.38; **???**), *magrittr* (Version 1.5; **???**), *maps* (Version 3.3.0; **???**), *Matrix* (Version 1.2.17; **???**), *mousetrap* (Version 3.1.2; **???**), *openxlsx* (Version 4.1.0; **???**), *papaja* (Version 0.1.0.9842; **???**), *plyr* (Version 1.8.4; @ }R-dplyr; **???**), *pracma* (Version 2.2.5; **???**), *processx* (Version 3.3.1; **???**), *psych* (Version 1.8.12; **???**), *purrr* (Version 0.3.2; **???**), *RColorBrewer* (Version 1.1.2; **???**), *Rcpp* (Version 1.0.1; **???**; **???**), *readbulk* (Version 1.1.2; **???**), *readr* (Version 1.3.1; **???**), *readxl* (Version 1.3.1; **???**), *Rmisc* (Version 1.5; **???**), *scales* (Version 1.0.0; **???**), *spam* (Version 2.2.2; **???**; **???**; **???**), *stringr* (Version 1.4.0; **???**), *tibble* (Version 2.1.3; **???**), *tidyr* (Version 0.8.3.9000; **???**), *tidyverse* (Version 1.2.1; **???**), and *yarrr* (Version 0.1.5; **???**) for all our analyses. Data preprocessing was completed in R using the mousetrap package (**???**). First, percent negative ratings were calculated for happy, angry, and surprised faces across all trial types, as well as a percent correct score for the memory probe trials. After, trials were screened for RT outliers. Any trials that were greater than three standard deviations from the mean were removed from the analyses. Additionally, we removed the preceding face rating trial for any incorrect memory probe trials, as these trials can be considered a manipulation failure.

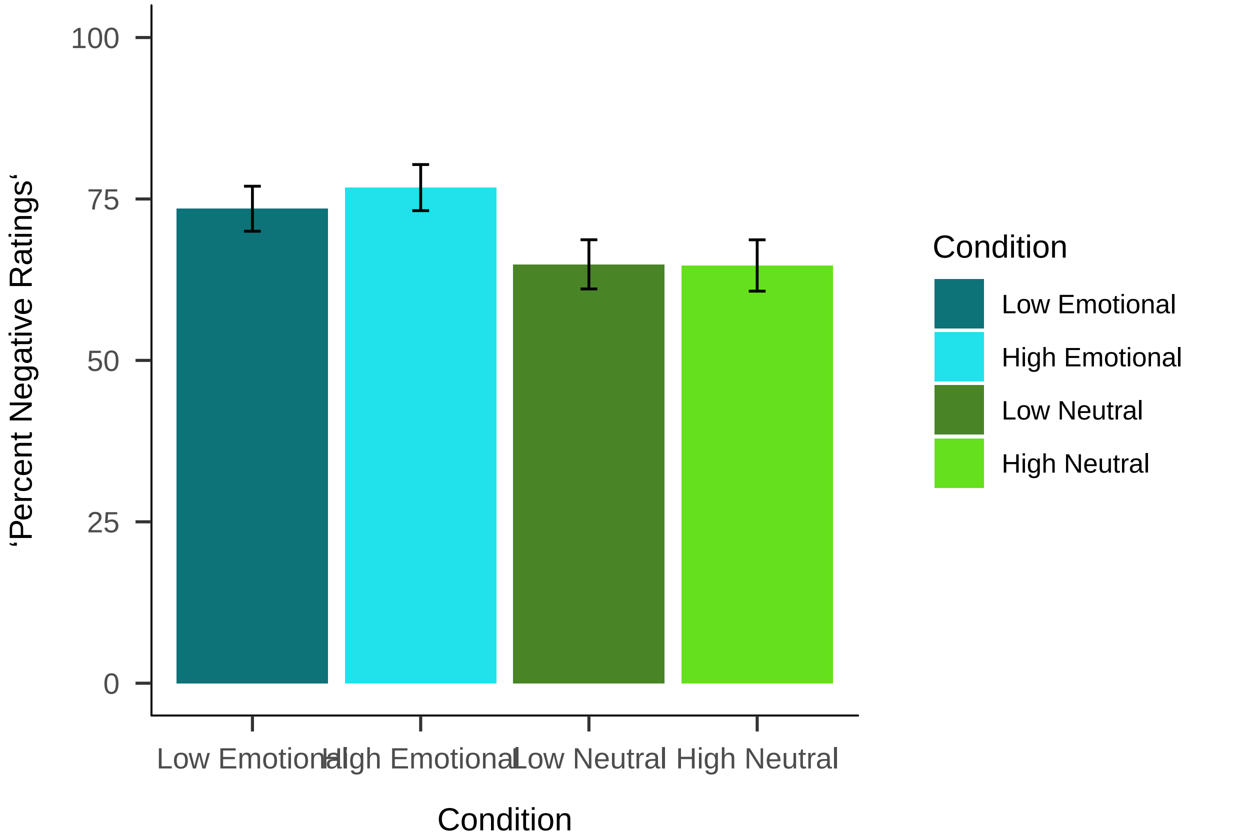
Prior to completing the analyses, all data were assessed for normality using Shapiro-Wilks tests. We tested for differences in valence bias among the different working memory load conditions. Friedman's test was used to assess overall differences and pairwise comparisons were completed using Wilcoxon signed rank tests using Bonferroni correction. Next, we tested for differences among maximum deviations in each working memory load condition using a load (low, high) X domain (emotional, neutral) repeated-measures ANOVA.

# Results

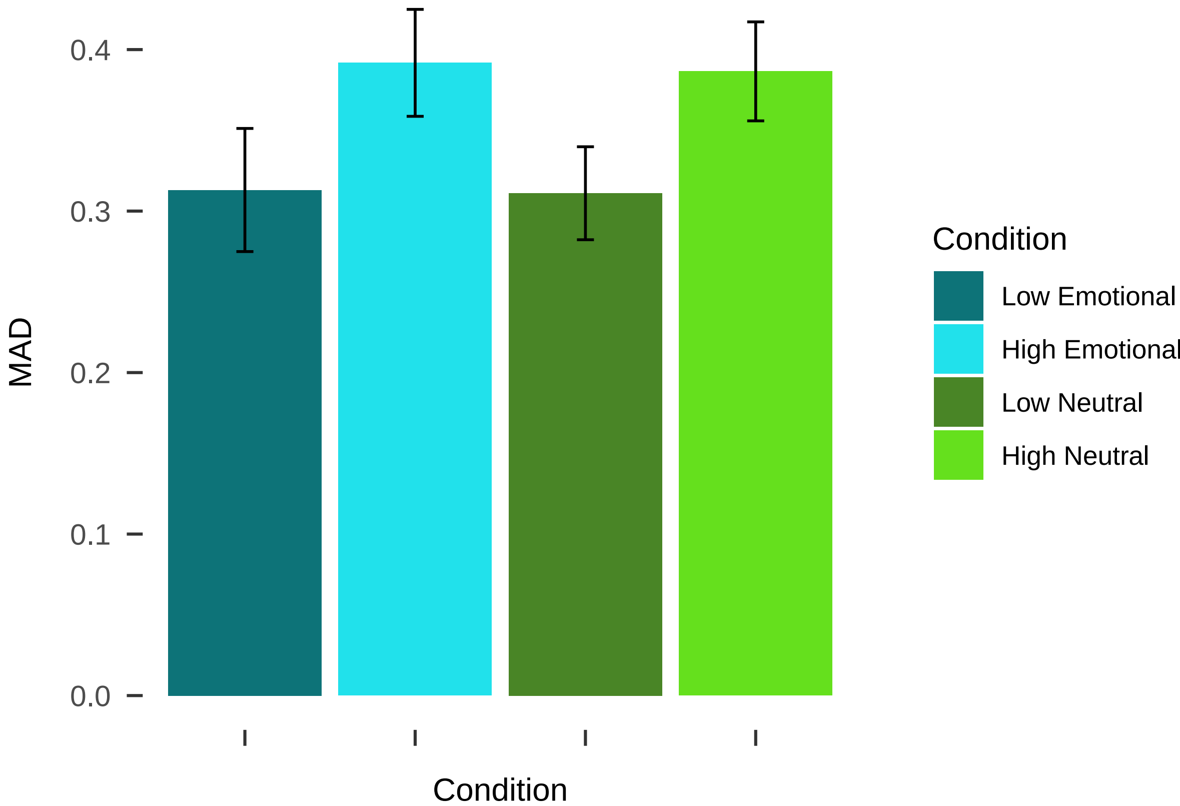
## Subjective ratings

Distributions of ratings were first tested for normality using Shapiro-Wilk's test. The results of all four tests were highly significant (p's $< .001$), so non-parametric tests were used for data analysis. Friedman's test results showed significantly different rank-order distributions across the conditions $\chi^2(3.00) = 27.79$, $p < .001$. Follow up Wilcoxon signed rank tests revealed that surprise is rated as more negative when holding emotional content in working memory compared to neutral content, and this was true for both low and high loads. Low emotional load ratings were significantly more negative than low, $Z = 3.27$, $p = .001$, neutral and high, $Z = 3.67$, $p < .001$, neutral loads. The same was true for high emotional load ratings and low, $Z = 4.55$, $p < .001$, and high, $Z = 3.81$, $p < .001$, neutral loads. However, there was no effect of load. That is, the comparisons between low and high load ratings for both emotional, $Z = -1.35$, $p = .176$, and neutral, $Z = -0.06$, $p = .954$, load ratings were not significantly different.[2]

[2] These results are qualitatively the same when analyzing these data with a repeated measures ANOVA.




Next, we assessed differences in maximum absolute deviation (MD) across the working memory trial conditions. While one of the conditions, low emotional MD, was not normally distributed (p = .024), all other conditions were normally distributed and repeated-measures ANOVA was used to analyze the MDs across conditions. There was a significant effect of load, F(1.00,196.00) = 5.51, p = .020, such that MDs under high load were larger than trials with low load. There was no significant effect of domain on MDs, F(1.00 196.00) = 0.01, p = .912, nor an interaction of load by domain, F(1.00 196.00) = 0.00, p = .960.




## Discussion

The effect of high vs. low load is still not apparent in these data, just like Mattek et al. 2016. An alternative explanation is that the high load manipulation is not sufficiently difficult to recruit the targeted cognitive resources; however, future work will be needed to better test this alternative.

Previous work has shown that more positive interpretations of surprised faces are related to slower RTs. Our working hypothesis suggests that this delayed reaction is a result of deliberation and slower, top-down cognitive processing. It is interesting to note that, at least in these data, there is no such difference observed between the neutral and emotional WM trials, *even though* the emotional WM trials are overall more negative. Future work should tease apart why this may be. For instance, . . .

Future work should consider whether the representations of these emotional images in

AWM (Reuter-Lorenz), or

## References

Brown, C. C., Raio, C. M., & Neta, M. (2017). Cortisol responses enhance negative valence perception for ambiguous facial expressions. *Scientific Reports*, *7*(1), 15107. doi:10.1038/s41598-017-14846-3

Egner, T., Etkin, A., Gale, S., & Hirsch, J. (2008). Dissociable neural systems resolve conflict from emotional versus nonemotional distracters. *Cerebral Cortex (New York, N.Y.: 1991)*, *18*(6), 1475–1484. doi:10.1093/cercor/bhm179

Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods*, *42*(1), 226–241. doi:10.3758/BRM.42.1.226

Gruber, O. (2001). Effects of domain-specific interference on brain activation associated with verbal working memory task performance. *Cerebral Cortex (New York, N.Y.: 1991)*, *11*(11), 1047–1055. doi:10.1093/cercor/11.11.1047

Lang, P., Bradley, M. M., & Cuthbert, B. N. (2008). International affective picture system (IAPS): Affective ratings of pictures and instruction manual., Technical Report A–8. University of Florida, Gainesville, FL.

Lavie, N., Hirst, A., Fockert, J. W. de, & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339–354. doi:10.1037/0096-3445.133.3.339

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The karolinska directed emotional faces—KDEF (CD ROM)., Stockholm: Karolinska Institute, Departmentof Clinical Neuroscience, PsychologySection.

Mattek, A. M., Whalen, P. J., Berkowitz, J. L., & Freeman, J. B. (2016). Differential effects of cognitive load on subjective versus motor responses to ambiguously valenced facial expressions. *Emotion*, *16*(6), 929–936. doi:10.1037/emo0000148

Mikels, J. A., & Reuter-Lorenz, P. A. (2019). Affective working memory: An integrative psychological construct. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *14*(4), 543–559. doi:10.1177/1745691619837597

Neta, M., Kelley, W. M., & Whalen, P. J. (2013). Neural responses to ambiguity involve domain-general and domain-specific emotion processing systems. *Journal of Cognitive Neuroscience*, *25*(4), 547–557. doi:10.1162/jocn_a_00363

Neta, M., Norris, C. J., & Whalen, P. J. (2009). Corrugator muscle responses are associated with individual differences in positivity-negativity bias. *Emotion (Washington, D.C.)*, *9*(5), 640–648. doi:10.1037/a0016819

Neta, M., & Tong, T. T. (2016). Don't like what you see? Give it time: Longer reaction times associated with increased positive affect. *Emotion (Washington, D.C.)*, *16*(5), 730–739. doi:10.1037/emo0000181

Neta, M., Tong, T. T., & Henley, D. J. (2018). It's a matter of time (perspectives): Shifting valence responses to emotional ambiguity. *Motivation and Emotion*, *42*, 258–266. doi:10.1007/s11031-018-9665-7

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. doi:10.1037/h0054651

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., … Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. doi:10.1016/j.psychres.2008.05.006

Whalen, P. J., Bush, G., Shin, L. M., & Rauch, S. L. (2006). The emotional counting stroop: A task for assessing emotional interference during brain imaging. *Nature Protocols*, *1*(1), 293–296. doi:10.1038/nprot.2006.45