Journal of
Experimental Psychopathology
JEP Volume 8 (2017), Issue 3, 303-319
ISSN 2043-8087 / DOI:10.5127/jep.059516

Are You Watching Me? Interacting Effects of Fear of Negative Evaluation and Social Context on Cognitive Performance

Erin L. Maresh, Bethany A. Teachman, James A. Coan

Department of Psychology, University of Virginia, Charlottesville, VA, USA

Abstract

Socially anxious individuals exhibit cognitive performance impairments; it is unclear whether this is due to trait differences in abilities or effects of the experimental context. This study sought to determine how social context, individual differences in fear of negative evaluation (FNE), and task difficulty interact to influence working memory performance as indicated by effectiveness (accuracy) and efficiency (reaction times). Participants (N = 61) performed the n-back task at 2-back and 3-back difficulty levels under three conditions: alone ("Anonymous"), in presence of a non-evaluative experimenter ("Presence"), and under explicit performance evaluation by the experimenter ("Threat"). Overall, participants showed improved accuracy during Threat, but only on 2-back trials. FNE was positively associated with longer reaction times during Threat on 3-back trials. FNE did not relate to accuracy, suggesting that threat-related impairments tied to social anxiety may alter efficiency rather than effectiveness. Thus, social anxiety may elicit cognitive performance impairments even in minimally evaluative environments.

Keywords: social anxiety, social-evaluative threat, social presence, working memory, cognitive impairment, social context, fear of negative evaluation

Correspondence to: Erin L. Maresh, Department of Psychology, University of Virginia, 314 Gilmer Hall, PO Box 400400, Charlottesville, VA 22904-4400. Email: emaresh@virginia.edu

Received 14-Sep-2016; received in revised form 14-Mar-2017; accepted 29-Mar-2017

Table of Contents

Introduction

The feeling of being evaluated can produce far-reaching negative consequences on cognitive performance. For example, social-evaluative threat (SET) can lead athletes to "choke" under pressure, can impair one's social abilities, and can heighten errors during cognitive tasks (Baumeister & Showers, 1986; Eysenck & Derakshan, 2011; Kashdan & Roberts, 2004; Mesagno, Harvey, & Janelle, 2012; Schmader, Johns, & Forbes, 2008). Individuals high in social anxiety, a trait characterized by excessive fear of being evaluated by others, may be especially sensitive to the deleterious effects of SET. However, although considerable work has explored the role of SET on *social* performance among socially anxious people, less is known about how social context interacts with social anxiety to affect performance on *non-social*, cognitive tasks. Intriguingly, people high in social anxiety have shown cognitive performance decrements in experimental situations in which threat of social evaluation is not made explicit (Cody et al., 2014; Maresh et al., 2014; O'Toole, Pedersen, Hougaard, & Rosenberg, 2015). While this may indicate that social anxiety corresponds with trait-like differences in cognitive abilities, it is also possible that socially anxious individuals experience cognitive impairments due to interference from social-contextual factors unaccounted for in the lab, such as the presence of an experimenter. This study attempts to disentangle this confound by introducing conditions designed to examine the independent and interacting effects of fear of negative evaluation (FNE), experimenter presence, and effects of overt evaluation on working memory performance. FNE is considered to be a core fear in social anxiety disorder, with individuals high in social anxiety showing considerable concern with being unfavorably evaluated by others (Leary, 1983).

Impairments in cognitive performance have been found following SET manipulations, particularly in anxious individuals (Oei, Everaerd, Elzinga, van Well, & Bermond, 2006; Schoofs, Preuß, & Wolf, 2008). According to attentional control theory, cognitive deficits related to trait anxiety may be due to impairments in executive functioning, leading to difficulty inhibiting task-irrelevant distractors, switching between tasks, and updating working memory (Berggren & Derakshan, 2013; Eysenck & Derakshan, 2011). These deficits appear more prominently during high cognitive load, suggesting task difficulty may be an important moderator of SET-related cognitive impairment (Beilock, Rydell, & McConnell, 2007; Oei et al., 2006).

Attentional control theory suggests that attentional deficits should be present even in the absence of threat, yet some studies have found no difference in cognitive performance based on social anxiety, or, in some cases, enhanced cognitive performance, suggesting that situational factors may be at play (Moriya & Sugiura, 2012; Wenzel & Holt, 2003). Importantly, attentional control theory posits that anxiety is most likely to reduce performance efficiency (such as how quickly someone completes a task) rather than effectiveness (such as a person's accuracy rate), as anxious individuals may use compensatory strategies to overcome impaired performance effectiveness. Thus, an anxious individual may overtly perform as well as a less anxious individual (e.g., show similar accuracy) but will use more resources to do so, which may be evident in longer reaction times. Psychophysiological measures, such as heart rate or skin conductance, can also be used to detect differences in resource expenditure that indicate reduced performance efficiency. Thus, focusing on a single outcome may mask important performance and physiological differences in individuals who fear negative evaluation.

Another limitation seen across social anxiety studies is that they often do not adequately address the potential influence of the experimental social context on cognitive performance. For individuals high in social anxiety, the presence of others may serve as especially salient task-irrelevant distractors, capturing attentional resources even without the overt threat of evaluation. Given that the perception of an audience may be sufficient to activate fears for socially anxious individuals (Clark & Wells, 1995; Rapee & Heimberg, 1997), the mere presence of an experimenter may impair the performance of socially anxious individuals even on tasks that are not explicitly social. Further, if the presence of an experimenter implies (or explicitly includes) real social evaluation, we could expect these impairments to worsen. In either case, it is unclear whether cognitive differences seen in social anxiety are related to trait-like differences in abilities *per se*, or to state differences activated by unaccounted-for SET.

Importantly, while SET manipulations used in previous research provide insight into the effects of social stress on cognitive abilities, these paradigms generally examine cognitive performance either *in anticipation of* or *following* the experience of SET (e.g., performing a task while waiting to give a speech or immediately after giving the speech), as opposed to *during*. Disruptions in task performance in these paradigms may be mediated by different processes than ongoing SET, such as fatigue or reduced motivation. Interestingly, in a study comparing digit span performance in individuals who showed elevated cortisol in response to a social stressor compared to those who did not, cortisol responders were found to exhibit working memory deficits compared to non-responders only when performing the task in front of an audience, but not prior to this social stressor or when recovering from it (Elzinga & Roelofs, 2005). Thus, ongoing SET may be a more potent disruptor of cognitive performance than anticipation of or recovery from SET.

In this study, we examined the interactive effects of FNE and task difficulty on working memory task performance (using the n-back task) across three within-subjects conditions that varied social context: one in which participants completed the n-back task alone ("Anonymous"), one in which they completed the task in the presence of a non-evaluative experimenter ("Presence"), and one in which they completed the task while the experimenter explicitly evaluated their performance ("Threat"). Condition differences were additionally assessed based on self-report data and psychophysiological measures of heart rate and skin conductance to provide additional data on performance efficiency during SET.

We hypothesized first that individuals higher in FNE would show deficits in n-back performance during the Threat condition compared to the Anonymous and Presence conditions, given that SET has been shown to interact with social anxiety to impair cognitive and working memory performance (Norton & Hope, 2001; Wenzel & Holt, 2003). We predicted that this impairment would be most evident as reduced efficiency (slower RTs, greater heart rate and skin conductance) rather than reduced effectiveness (lower accuracy). Second, we hypothesized that individuals higher in FNE would additionally show reduced efficiency during the Presence condition compared to the Anonymous condition, due to evidence that mere presence of an experimenter may activate socially anxious fears and that anxiety may confer susceptibility to task-irrelevant distractors (Clark & Wells, 1995; Eysenck & Derakshan, 2011; Rapee & Heimberg, 1997). We additionally hypothesized that performance interference would be more prominent during harder trials, as task disruptions may only be evident when difficulty exceeds a certain threshold (Beilock et al., 2007; Oei et al., 2006).

Method

Participants

Participants (N = 61; 47 women) were recruited through the University's psychology participant pool based on their responses to the 20-item Social Phobia Scale (SPS; Mattick & Clarke, 1998), administered as part of the department-wide pre-selection survey. Participants were recruited using a stratified sampling method to ensure roughly equal numbers of participants with low, moderate, and high levels of social anxiety, with low levels defined as scores less than 1 SD below the mean of the total SPS scores in the participant pool, moderate levels as those between -1 and +1 SD, and high levels as scores greater than 1 SD above the mean. The participant pool (N = 728) showed slightly higher levels of social anxiety on the SPS (*M* = 18.69, *SD* = 14.17) compared to published norms on a sample of 482 undergraduate students (*M* = 14.1, *SD* = 10.2; Mattick & Clarke, 1998). Additional information on this sample (e.g., age, race, ethnicity) is provided in Table 1.

Table 1: Sample characteristics.

	Total
N (females)	61 (47)
Age *M* (SD)	18.52 (1.04)
Race N (%)	
White	32 (52.5%)
Asian	17 (27.9%)
Black	6 (9.8%)
Other/Mixed	6 (9.8%)
Ethnicity N (%)	
Hispanic	2 (3.3%)
Non-Hispanic	59 (96.7%)
BFNE-S *M* (SD)	22.69 (7.82)
DASS-D *M* (SD)	6.98 (7.20)
STICSA *M* (SD)	37.34 (10.31)

BFNE-S=Brief Fear of Negative Evaluation Scale, straightforward-worded items; DASS-D=Depression Anxiety Stress Scales, Depression subscale; STICSA=Trait version of the State-Trait Inventory for Cognitive and Somatic Anxiety.

Social context manipulation

The social context paradigm being tested in this study consisted of performing the n-back task under three conditions – "Anonymous," "Presence," and "Threat" – that varied in their degree of social and evaluative characteristics. In all cases, the experimenter administering the conditions was one of three undergraduate research assistants (two males and one female). Each participant underwent all three conditions with order counterbalanced across participants, and measures of state affect were collected after each condition. After performing a brief practice version of the n-back, participants were informed that they would be performing "several rounds" of the n-back task but were not given descriptions of the condition manipulations until debriefing.

Anonymous condition. In the Anonymous condition, participants were given the following information: "Your scores on this task will be recorded for data analysis but will be completely confidential. I will have no knowledge of your performance." Specific to this condition, the experimenter then stated that he or she needed to go prepare some paperwork and then left the room. This manipulation was intended to reduce, to the degree possible in a psychology experiment, feelings of social presence and evaluation.

Presence condition. In the Presence condition, the experimenter provided the same information as in the Anonymous condition but did not state that he or she would be leaving the room to complete paperwork. Instead, the experimenter sat adjacent to the participant at an angle at which the participant's monitor was not visible. The experimenter then became engaged in reading a document to appear occupied. This was intended to reduce feelings that the participant's performance was being evaluated while still maintaining social presence.

Threat condition. In the Threat condition, the experimenter informed the participant that he or she would be watching their performance and recording each time a mistake was made for the purpose of comparing the participant's performance to that of his or her peers. The experimenter then sat adjacent to the participant as he or she performed the task, similar to the Presence condition but with the monitor turned slightly to be visible to the experimenter, and marked each response made by the participant on a sheet of paper. This condition was designed to contain both social presence and an evaluative component.

Materials

Assessment of working memory: Visual n-back task. In visual n-back tasks, participants view a sequence of letters or numbers on a computer screen with the goal of identifying when the stimulus presented matches the one seen *n* steps earlier (Gevins & Cutillo, 1993). This task is thought to measure working memory abilities related to recognition (Redick & Lindsey, 2013) and can be modified for different levels of difficulty by changing the number of steps back one must remember. In this study, the *n*-back task was administered using a modified version of the task available from PsyToolkit (psytoolkit.org; Stoet, Windows, & Macintosh, 2010). Participants were presented with sequences of letters of the alphabet, each randomly displayed from one of fifteen letters, chosen on the basis of their visual distinctiveness. Participants were instructed to press the letter "M" for target letters and the letter "N" for all other letters (thus, each trial required a response in order to be correct); 33% of the stimuli were target letters and 67% were non-target letters. Each letter was presented for 2000 ms followed by an interstimulus interval of 500 ms during which the participant could still make a response. After responding (or failing to respond), two bars above and below the letter flashed either green or red to reflect that the participant's answer was correct or incorrect, respectively.

To examine the potential moderating effect of task difficulty, participants completed eight alternating blocks of 2-back and 3-back trials within each condition (cf. Schoofs et al., 2008). Each block contained 20 trials (if 2-back) or 21 trials (if 3-back). Because the first two trials for 2-back and three trials for 3-back cannot be target trials, they were removed from analysis, leaving 18 analyzable trials within each block. This yielded a total of 72 trials of 2-back and 72 trials of 3-back within each condition, for a total time of ~6 minutes for each condition. Performance on the n-back task was measured both by reaction times (RTs) and accuracy for each trial. Internal consistency across all trials was high for both RTs (Cronbach’s alpha = .99) and accuracy (Cronbach’s alpha = .93).

State responses to social context manipulation. Following each of the three social context conditions, participants completed four items on the computer using a 0 to 100 slider to assess the greatest degree to which they experienced certain states during the task they just completed. These questions were "To what extent did you feel like you were alone during this task?" (with higher scores indicating feeling less alone), "To what extent did you feel like you were being monitored?", "How anxious did you feel?", and "To what extent did you feel like you were being negatively evaluated?" These measures served as manipulation checks for the social context conditions.

Questionnaires[1]. The Social Phobia Scale (SPS; Mattick & Clarke, 1998) was administered as part of the university participant pool pre-test to use as the initial social anxiety screener for recruitment, as described in the Participants and Recruitment section. The SPS is a 20-item questionnaire that assesses fears of being scrutinized or observed (e.g., *"I get nervous that people are staring at me as I walk down the street," "I worry I might do something to attract the attention of others").* Participants are asked to indicate the degree to which each statement is characteristic of them using a 5-point Likert scale ranging from 0 = "Not at all" to 4 = "Extremely." This scale showed excellent reliability in this sample (Cronbach's alpha = .94).

The Brief Fear of Negative Evaluation Scale-Straightforward Items (BFNE-S; Leary, 1983) was administered at the time of the lab study session. Although we recruited participants based on SPS, we chose to examine fear of negative evaluation during the lab session specifically, rather than trait social anxiety more generally, because of our hypothesis that FNE is a specific cognitive mechanism related to social anxiety that may impair performance under SET. However, it is worth noting that the BFNE-S and SPS were moderately to strongly correlated in our sample ($r = .62$, $p<.0001$). The original BFNE is a 12-item scale that measures the degree to which a person experiences apprehension at the prospect of being evaluated negatively (e.g., *"I am afraid that others will find fault with me," "I am usually worried about what kind of impression I make"*) using a 5-point Likert scale ranging from 0 = "Not at all characteristic of me" to 4 = "Extremely characteristic of me." Because the reverse-scored items on the original BFNE may exhibit poor reliability and validity (e.g., Rodebaugh et al., 2011), only the eight straightforward-worded items were used (BFNE-S). These eight items showed excellent reliability in this sample (Cronbach's alpha = .92).

The depression subscale from the short-form version of the Depression Anxiety Stress Scales (DASS-21; Lovibond & Lovibond, 1995) and the State-Trait Inventory of Cognitive and Somatic Anxiety (STICSA; Ree, French, MacLeod, & Locke, 2008) were administered during the lab study. Because social anxiety is characterized by both decreased positive affect, as seen in depression, and increased negative affect, as seen in general anxiety (Brown, Chorpita, & Barlow, 1998), these scales were used to examine the unique contribution of social anxiety to working memory performance after controlling for depression and general anxiety symptoms. The DASS-D and STICSA both showed good reliability in this sample (Cronbach's alphas=.87 and .90, respectively).

Psychophysiological recording. Heart rate and skin conductance data were collected using a Biopac data acquisition unit (MP150; Biopac Systems Inc., US). Heart rate was collected using a two-lead electrocardiogram (ECG) with lead placement on both wrists and was sampled at a rate of 1,000 Hz. Skin conductance was collected with the placement of two disposable electrodes on the medial phalanges on the first two digits on the left hand and was sampled at a rate of 250,000 Hz. The data were cleaned, filtered, and analyzed using Biopac's AcqKnowledge 4.1 software. Specifically, heart rate data was band pass filtered between 0.05 and 1.0 Hz, while skin conductance data was low pass filtered at 1 Hz. Smoothing at a factor of 500 was applied to skin conductance data. Heart rate and skin conductance data were then averaged across seven discrete 60-s epochs for each condition. Due to data collection errors, heart rate and skin conductance data were not recorded for three participants, leaving $n = 58$ participants with psychophysiological data.

Procedure

Upon arrival to the laboratory, participants were informed that this study examines how cognitive performance affects physiological activity, that they would be completing several rounds of a memory task while heart rate and skin conductance are recorded, and that they would then complete some questionnaires assessing personality characteristics. Information regarding the different social-contextual conditions was not provided until the end of the experiment.

Following informed consent, participants were attached to psychophysiological equipment, and a two-minute baseline recording of heart rate and skin conductance was conducted. After completing a brief practice version of both the 2-back and 3-back task to ensure comprehension, participants performed the n-back task again under the

---

[1] Several additional measures were collected as part of a larger study. Given the specific hypotheses for this study, only BFNE-S was used as a trait moderator for analyses. A full list of measures is available from the first author.

three social context conditions (Anonymous, Presence, and Threat), with condition order counterbalanced across participants. Psychophysiological data were continuously recorded during each condition. State responses to the social context manipulation were collected immediately following each condition. After the n-back task, participants completed a battery of questionnaires assessing several facets of social anxiety, as well as general anxiety and depressive symptoms. Questionnaires were administered last in the study to prevent explicit priming of social anxiety constructs prior to the social context manipulation. Finally, participants were debriefed and received course credit for their participation.

Plan for analyses

Data scoring and reduction. The SPS (used for recruitment) and BFNE-S (used for analyses) were each scored by summing across all items, following scoring procedures described in the original publications. BFNE-S scores were mean-centered prior to subsequent analyses. For n-back data analysis, trials were excluded if the participant did not give a response (<0.1% of trials) or if a trial RT was less than 200 ms (<1% of trials), as RTs faster than 200 ms may indicate preemptive guessing (Luce, 1986). All resulting trials were used for analyses of accuracy data. For analyses of RTs, only trials in which participants made a correct response were used (89% of trials).

State measures. To confirm that the social context manipulation had the intended effects, a series of linear mixed models were conducted with each state measure predicted by a fixed effect of condition (three-level factor: Anonymous, Presence, and Threat), with a random intercept for each participant. Pairwise t-tests with Satterthwaite approximations to degrees of freedom were then performed to determine significant differences between conditions.

Psychophysiological measures. To further characterize the social context manipulation, linear mixed models were conducted with heart rate and skin conductance each predicted by a fixed effect of condition (three-level factor: Anonymous, Presence, and Threat), with a random intercept for each participant. Pairwise t-tests with Satterthwaite approximations to degrees of freedom were then performed to determine significant differences between conditions.

Interactive effects of social context and fear of negative evaluation on working memory performance. To account for the clustered nature of the data in this study, a mixed effects regression modeling approach was used. This analytic method has several well-documented advantages over more traditional repeated-measures ANOVA approaches, including improved flexibility in modeling continuous-categorical variable interactions, more robust, unbiased handling of missing data, and the ability to model variations in responses across time (Baayen, Davidson, & Bates, 2008; Nich & Carroll, 1997). We used the maximal random effects structure justified by the study design (i.e., that random intercepts and slopes should be specified where they are expected to exist in the data) as recommended by Barr et al. (Barr, Levy, Scheepers, & Tily, 2013) due to inflated Type I error rates present in under-specified models. This approach was additionally justified through the use of likelihood ratio tests to compare the maximal model with models in which one random effect term was dropped at a time. The maximal model (described below) showed the greatest model fit. Each model was fitted using the “lme4” package in R (Bates, Mächler, Bolker, & Walker, 2015; R Core Team, 2013) and p-values were determined using the package "lmerTest" (Kuznetsova, Brockhoff, & Christensen, 2014).

Separate linear mixed models were conducted for each performance outcome variable (RT and accuracy). For both the RT and accuracy models, fixed effects included n-back difficulty (two-level factor: 2-back and 3-back), condition (three-level factor: Anonymous, Presence, and Threat), scores on the BFNE-S (continuous; mean-centered), all two-way interaction terms, and the three-way interaction term. When predicting RT, random effects included random by-participant slopes for condition, n-back level, and their interaction and a random intercept for participant; random intercepts for block and trial nested within block were additionally included to account for general variability in blocks and trials within blocks across participants. In mixed models, failures of convergence can occur when there is too little data given the number of parameters in the model (Bates, Kliegl, Vasishth, & Baayen, 2015). Accordingly, when predicting accuracy, the random effects structure had to be reduced due to the reduction in data when looking at a binomial outcome variable, such that only a by-participant random slope for condition was fitted. Planned comparisons using Satterthwaite approximations to degrees of freedom were then conducted to probe pairwise effects.

Secondary analyses. To examine other potential important moderators of performance, we ran several additional analyses assessing the effects of gender, depression, trait anxiety, and condition order. Because there were only 14 men in our sample, there was insufficient power to use gender as a factor. Instead, we ran the same models predicting RTs and accuracy described above on just the sample of women (N = 47). To assess whether effects of FNE on performance were unique to FNE, depression and trait anxiety scores were each included as a covariate in separate models predicting RTs and accuracy. To assess effects of condition order on performance, separate mixed models were conducted for RTs and accuracy with condition, condition order, and their interaction as fixed effects. Random effects included a random by-participant slope for condition order. Condition itself could not be included as a random slope with condition order due to lack of convergence.

Sample size considerations. Straightforward guidelines to conduct power analyses for mixed effects models are not currently available. We selected a sample size of N = 61 based on several sources. The "30-30 rule" established by Kreft and DeLeeuw (1998) recommends at least 30 upper-level units (e.g., participants) each with at least 30 lower-level units (e.g., trials) to obtain sufficient power. However, it has been suggested that this heuristic may not be universally applicable (Mathieu, Aguinis, Culpepper, & Chen, 2012). A simulation study of cross-level interaction effects identified >90% power when using 60 upper-level units with 18 lower-level units (Mathieu et al., 2012). Thus, we recruited 61 participants to increase our ability to detect reliable effects. Further, we had well over 18 trials within each participant (72 trials at each n-back level within each condition for a total of 432 trials per participant).

Results

Sample characteristics

Participant characteristics are reported in Table 1. Participants reported levels of FNE on the BFNE-S in line with norms previously determined on a university student sample (*M* = 22.69, *SD* = 7.82; see Rodebaugh et al., 2011). Levels of FNE did not differ by gender ($t(59) = 0.14$, $p = 0.889$).

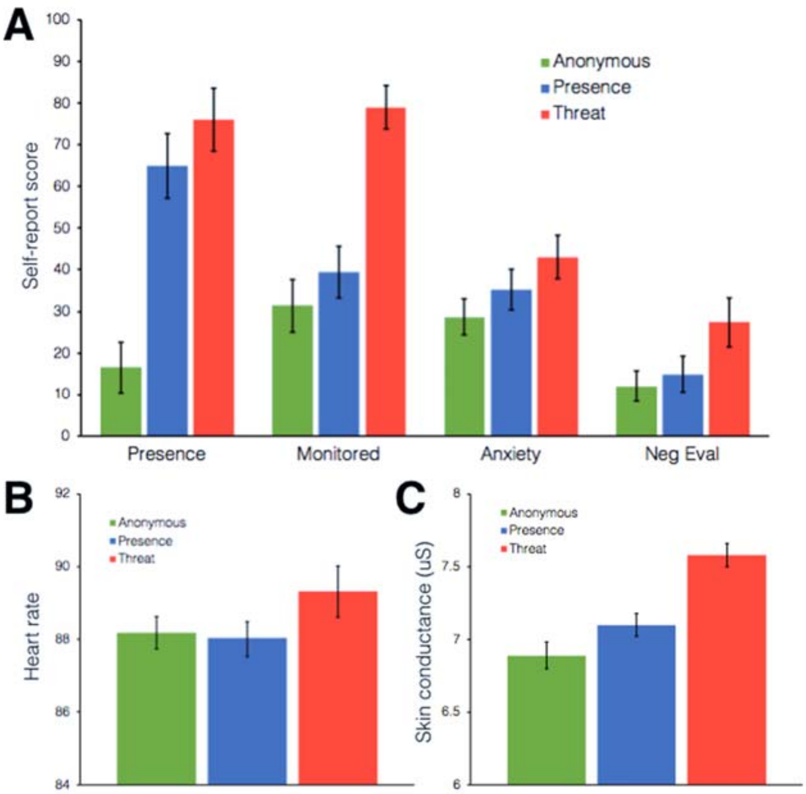

Figure 1: A) Means of state measures (feelings of social presence, feelings of being monitored, feelings of anxiety, and feelings of being negatively evaluated) by social context condition. Each measure was self-reported by the

participant on a 0-100 scale at the end of the condition. B) Mean heart rate averaged across each condition. C) Mean skin conductance averaged across each condition. All error bars represent 95% confidence intervals.

Manipulation check

State measures. To evaluate whether the social context manipulation had the intended effect on participants, differences among state measures (feeling not alone, monitored, anxious, and negatively evaluated) were assessed. As hypothesized, participants reported monotonic increases in all state measures from the Anonymous to Presence to Threat conditions (Figure 1A, all ps<0.05), except for the perceived negative evaluation item. In other words, participants felt less alone, more monitored, and more anxious when another person was in the room than when they were alone; this effect was further heightened if the other person was overtly evaluating them. Levels of perceived negative evaluation were greater only in the Threat condition compared to the Presence and Anonymous conditions (ps<0.05). Moderation of condition effects on state measures by FNE was also examined. A significant interaction emerged between FNE and condition only in regards to level of perceived negative evaluation, such that higher FNE associated with greater feelings of negative evaluation in the Threat condition compared to the Presence and Anonymous conditions (ps<0.05).

Psychophysiological data. In addition to self-report, we collected measures of peripheral psychophysiology to assess whether the social context manipulation altered participants' mean heart rate and skin conductance. Both heart rate and skin conductance significantly increased in the Threat compared to Presence and Anonymous conditions (Figure 1B & Figure 1C, all ps<0.05); the Presence and Anonymous conditions did not differ on heart rate or skin conductance. Additionally, FNE showed a main effect on heart rate, such that FNE associated with higher heart rate activity across conditions ($F(1, 56) = 6.60$, $p = .013$); FNE did not associate with higher skin conductance ($F(1, 56) = 3.64$, $p = .061$). No significant interactions between FNE and condition emerged.

Table 2: Summary table for model predicting reaction times in milliseconds.

Fixed effect	**Beta**	**SE**	**t**	**p**
Intercept (Anonymous)	758.10	19.50	38.87	<.001***
Intercept (Presence)	755.91	16.71	45.24	<.001***
Intercept (Threat)	773.04	17.74	43.58	<.001***
Presence vs. Anonymous	-2.19	15.48	-0.14	0.888
Threat vs. Anonymous	14.94	15.15	0.99	0.328
Threat vs. Presence	17.13	14.33	1.2	0.237
N-back (Anonymous)	54.70	16.86	3.24	0.003**
N-back (Presence)	55.29	18.56	2.98	0.005**
N-back (Threat)	77.90	17.26	4.51	<.001***
FNE (Anonymous)	-0.57	2.41	-0.24	0.813
FNE (Presence)	1.70	2.04	0.84	0.407
FNE (Threat)	0.92	2.18	0.42	0.674
Condition (Presence vs. Anonymous) x N-back	0.56	7.95	0.07	0.944
Condition (Threat vs. Anonymous) x N-back	23.03	7.91	2.91	0.004**
Condition (Threat vs. Presence) x N-back	22.47	7.90	2.85	0.004**
Condition (Presence vs. Anonymous) x FNE	2.30	2.00	1.15	0.254
Condition (Threat vs. Anonymous) x FNE	1.50	1.95	0.77	0.445
Condition (Threat vs. Presence) x FNE	-0.80	1.84	-0.43	0.666
N-back x FNE (Anonymous)	-1.96	1.67	-1.17	0.245
N-back x FNE (Presence)	-3.33	1.94	-1.71	0.092
N-back x FNE (Threat)	1.35	1.73	0.78	0.438
Condition (Presence vs. Anonymous) x N-back x FNE	-1.28	1.02	-1.26	0.209
Condition (Threat vs. Anonymous) x N-back x FNE	3.34	1.02	3.29	0.001**
Condition (Threat vs. Presence) x N-back x FNE	4.62	1.02	4.55	<.001***

Treatment coding was used for Condition and simple coding was used for n-back level. Condition in parentheses represents baseline condition. Beta weights and standard errors are unstandardized values.

**p ≤ .01; *** p ≤ .001

Interactive effects of social context, n-back difficulty, and trait fear of negative evaluation on working memory performance.

To evaluate the effect of SET on working memory performance, the interactive effects of social context, FNE, and n-back difficulty level on both RT and accuracy were assessed using mixed effects models. Table 2 and Table 3 show the results for the models predicting RT and accuracy.

Reaction times

A significant main effect of n-back level on RTs emerged ($F(1, 23) = 14.90$, $p<0.001$), such that participants were slower to respond to 3-back trials compared to 2-back trials (2-back $M = 730.30$, $SD = 268.55$ ms; 3-back $M = 793.17$, $SD = 298.50$ ms). This was subsumed by a significant three-way interaction between condition, n-back level, and FNE, indicating that the interactive effects of condition and n-back level on RT were moderated by FNE ($F(2, 58) = 5.15$, $p = 0.009$), as expected. Visual inspection of the interaction (Figure 2) indicated that during 3-back trials, higher (compared to lower) FNE was associated with shorter RTs in the Anonymous condition, no difference in the Presence condition, and longer RTs in the Threat condition. During 2-back trials, higher (compared to lower) FNE was associated with longer RTs in the Presence condition, with no FNE difference in the Anonymous or Threat conditions.

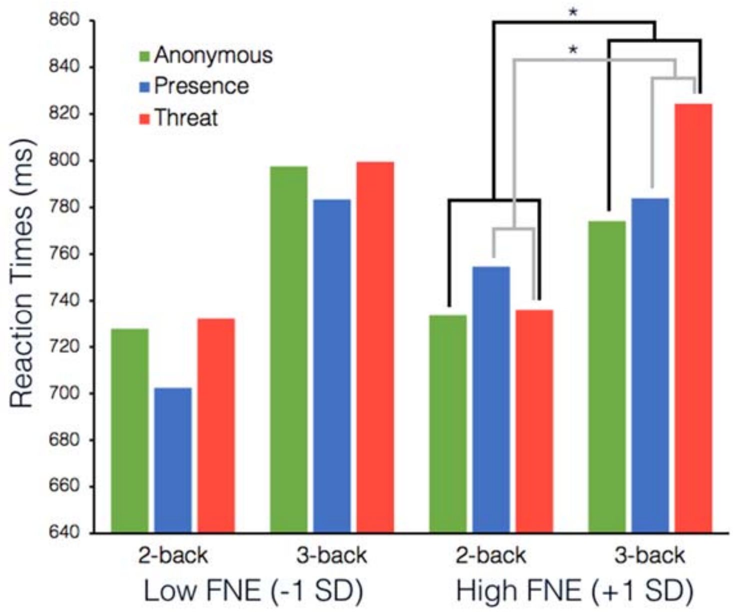

Figure 2: Reaction times by condition and n-back level at different values (+1 standard deviation and -1 standard deviation) of fear of negative evaluation (FNE). Asterisks indicate significant (p<.05) 2-way interactions within high FNE.

To further decompose this interaction, all possible two-way interactions were tested at different levels of the third variable. First, two-way interactions between condition and FNE were tested at each n-back level; no significant interactions emerged. Second, two-way interactions between condition and n-back level were tested at high (+1 SD) and low (-1 SD) levels of FNE using the simple slopes method recommended by Aiken and West (1991). For participants with high levels of FNE, a significant n-back level by condition interaction emerged such that the increase in RTs from the Anonymous to Threat condition was greater during 3-back compared to 2-back trials ($t(58) = 3.35$, $p = 0.001$; Figure 2). Additionally, for those high in FNE, the slope between the Presence and Threat condition differed based on n-back level ($t(58) = 3.30$, $p = 0.002$), such that RTs increased from the Presence to Threat condition during 3-back trials but *decreased* from the Presence to Threat condition during 2-back trials. No significant interactions emerged in participants with low levels of FNE. Finally, two-way interactions between n-back level and FNE were tested within each condition. No significant interactions emerged in any condition. In sum, at high levels of FNE, but not at low levels, participants showed greater impairment on harder trials during Threat compared to Presence and Anonymous.

Accuracy

When predicting accuracy, a significant main effect of n-back level emerged that was subsumed by a significant two-way interaction between condition and n-back level ($X^2(2) = 20.04$, $p<0.0001$). As seen in Figure 3, comparisons of slopes indicated that during 2-back trials, accuracy was greater in the Threat condition compared to the Presence condition ($z = 2.96$, $p = 0.003$) and Anonymous condition ($z = 4.46$, $p<0.0001$). Accuracy was not greater during 2-back trials in the Presence compared to Anonymous condition ($z = 1.67$, $p = .095$). During 3-back trials, accuracy did not differ based on condition. FNE did not moderate any effects of condition or n-back level on accuracy, nor was the three-way interaction between FNE, condition, and n-back significant ($X^2(2) = 1.57$, $p = .455$).

Table 3: Summary table for model predicting accuracy in percent.

Fixed effect	Beta	SE	t	p
Intercept (Anonymous)	87.90	0.90	97.36	<.001***
Intercept (Presence)	88.38	0.87	101.91	<.001***
Intercept (Threat)	89.66	0.78	114.77	<.001***
Presence vs. Anonymous	0.48	0.70	0.69	0.493
Threat vs. Anonymous	1.77	0.69	2.55	0.014*
Threat vs. Presence	1.28	0.65	1.97	0.054
N-back (Anonymous)	-6.86	0.66	-10.37	<.001***
N-back (Presence)	-8.12	0.66	-12.28	<.001***
N-back (Threat)	-9.63	0.66	-14.62	<.001***
FNE (Anonymous)	0.08	0.12	0.71	0.479
FNE (Presence)	-0.02	0.11	-0.18	0.855
FNE (Threat)	-0.02	0.10	-0.24	0.814
Condition (Presence vs. Anonymous) x N-back	-1.26	0.94	-1.35	0.176
Condition (Threat vs. Anonymous) x N-back	-2.78	0.93	-2.98	0.003**
Condition (Threat vs. Presence) x N-back	-1.52	0.93	-1.62	0.105
Condition (Presence vs. Anonymous) x FNE	-0.10	0.09	-1.14	0.257
Condition (Threat vs. Anonymous) x FNE	-0.11	0.09	-1.20	0.237
Condition (Threat vs. Presence) x FNE	0.00	0.08	-0.04	0.968
N-back x FNE (Anonymous)	-0.01	0.09	-0.10	0.921
N-back x FNE (Presence)	0.08	0.09	0.93	0.352
N-back x FNE (Threat)	-0.01	0.08	-0.06	0.952
Condition (Presence vs. Anonymous) x N-back x FNE	0.09	0.12	0.73	0.466
Condition (Threat vs. Anonymous) x N-back x FNE	0.00	0.12	0.03	0.977
Condition (Threat vs. Presence) x N-back x FNE	-0.08	0.12	-0.70	0.483

Treatment coding was used for Condition and simple coding was used for n-back level. Condition in parentheses represents baseline condition. Beta weights and standard errors are unstandardized values.

*p ≤ .05; **p ≤ .01; *** p ≤ .001

Secondary moderator analyses.

Gender, depression, trait anxiety, and condition order were additionally examined as moderators of performance to assess possible alternative explanations for our findings.

Gender. We examined whether our findings changed when looking at the subset of women only (n = 47). The pattern of significance did not differ for RTs; for accuracy, condition became a significant predictor, but there were still no interactions with FNE.

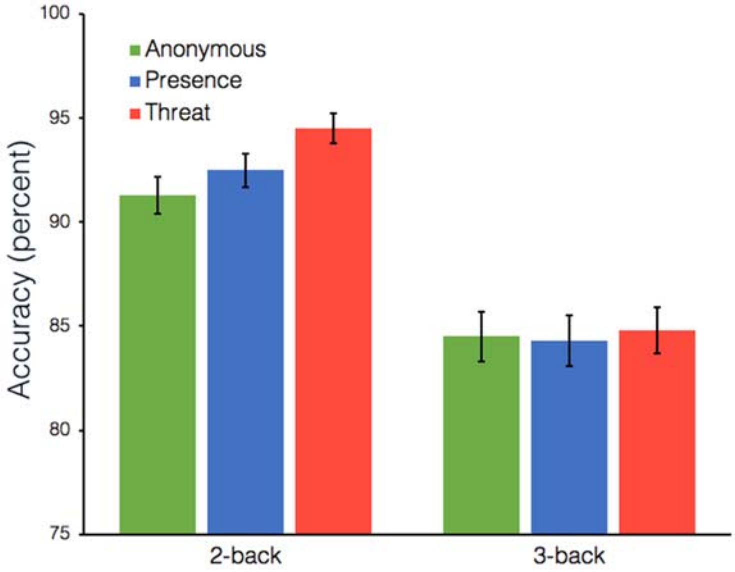

Figure 3: Two-way interaction between social context condition and n-back difficulty predicting accuracy. Error bars represent 95% confidence intervals.

Depression and trait anxiety. When including depression or trait anxiety as a covariate, the pattern of significance did not change for RTs. For accuracy, including either depression or trait anxiety scores resulted in a significant condition effect that continued to be subsumed by a two-way interaction between condition and n-back.

Order effects. Because of the within-subjects design of this study, practice and/or fatigue may have affected performance on later conditions. We examined whether condition order interacted with condition and FNE to influence RTs or accuracy. Condition order showed a main effect on RTs and accuracy such that RTs were faster and accuracy was higher on each subsequent condition ($ps<0.05$), suggesting improvement rather than fatigue. Levels of FNE and condition did not interact with order effects ($ps>0.05$).

Discussion

The primary goal of this study was to examine how FNE interacts with different social and evaluative contexts to influence cognitive performance on a working memory task. Specifically, we tested whether the mere presence of an experimenter affected working memory performance differentially from performing the task either alone or under overt evaluative threat and, further, whether trait FNE influenced effects of condition on performance. Because task difficulty was expected to moderate these effects, two difficulty levels of the n-back task, 2-back and 3-back, were included in the study design. The results indicated that social context and FNE do interact to influence working memory performance, and these interactive effects depend considerably on task difficulty. As predicted by attentional control theory (Eysenck et al., 2007), higher FNE corresponded with longer reaction times during Threat on 3-back trials but did not correspond with differences in accuracy.

Effects of Ongoing Social-Evaluative Threat

Participants higher in FNE displayed slower RTs on the n-back task when being overtly evaluated by an experimenter compared to when completing the n-back alone or in the presence of a non-evaluative experimenter, suggesting that FNE reduces performance efficiency. This pattern only emerged during 3-back trials, the more difficult n-back level, as would be expected given evidence that SET interference might become apparent only when task difficulty exceeds a certain threshold (Beilock et al., 2007; Oei et al., 2006). While confirming previous studies of the effects of SET, this finding is particularly notable given that we examined the effects of an *ongoing* SET context, rather than the anticipation or after-effects of SET, as has often been done in the past. In this way, the Threat condition more closely mirrors many real-life testing, academic, and work situations. It is worth noting that, while the Threat condition in this study did indeed increase feelings of anxiety and negative evaluation, this condition is rather mild in terms of SET for many reasons: the experimenter did not provide negative feedback, the experimenter was a similarly-aged peer (vs.

an authority figure), task difficulty did not increase as performance continued, and there was no tangible negative outcome to performing poorly. This suggests that even mildly evaluative situations, which are largely unavoidable in day-to-day life, are sufficient to induce performance impairments in individuals with elevated FNE.

Of note, although SET yielded slower RTs for socially anxious individuals in 3-back trials compared to 2-back trials, this did not correspond with improved accuracy; indeed, social anxiety, whether by itself or in interactions with condition and n-back level, did not correspond with any differences in the number of correct trials. Attentional control theory states that anxiety may reduce *efficiency* more so than *effectiveness* (Eysenck & Derakshan, 2011)-in other words, although overt performance (in terms of number of correct trials) was not impaired in anxious individuals during SET, this may be because they were expending more resources to perform on par with their less anxious peers. In accordance with this, individuals high in FNE did display elevated heart rate activity (compared to low FNE individuals), which is consistent with heightened resource utilization. FNE, and social anxiety more generally, may incur "hidden costs"–subtle performance and physiological effects that may not be readily apparent to others but that are likely to be taxing nonetheless, especially in the aggregate.

Effects of Social Presence

The impact that the presence of others can have on general performance has long been a topic of study by social psychologists, who for decades have observed that social presence can either facilitate or inhibit behaviors, and that personality can be a strong moderator of these effects (see Uziel, 2007, for a review). Social facilitation and inhibition are thought to be related to the potential for evaluation, and, correspondingly, we observed facilitation of correct responses to easier 2-back trials in the Threat condition (relative to both the Presence and Anonymous conditions), irrespective of level of FNE.

Notably, hypothesized performance differences on reaction times and accuracy were not observed when comparing the Presence condition to the Anonymous condition, even when examining moderation by FNE, although it is worth mentioning that participants did report greater anxiety and feelings of being monitored during Presence compared to Anonymous. The null findings for Presence vs. Anonymous differences could indicate a true absence of performance effects related to mere social presence. Alternatively, it is also possible that performance interference due to social presence is only evident in individuals with more severe social anxiety. We recruited participants using a continuous measure of social anxiety to sample a wide range of behavior, in line with the Research Domain Criteria (RDoC; http://www.nimh.nih.gov/research-priorities/rdoc/index.shtml) framework, rather than recruiting based on meeting diagnostic criteria for social anxiety disorder. Although approximately 10% of our participants exceeded the average BFNE-S score seen in samples diagnosed with social anxiety disorder (Rodebaugh et al., 2011), future studies may wish to expand the number of individuals with clinical levels of social anxiety. Regardless, our present findings suggest that, at least in an undiagnosed sample, the presence of a non-evaluative person may not interfere with cognitive performance.

Limitations and Future Directions

There are some additional limitations to our study design. First, we elected to use a within-subjects design to examine how context changes performance within an individual rather than between individuals. This design may have led to practice and/or habituation effects that dampened any interference from the social context conditions. However, while participants showed improved n-back performance over time, this improvement did not differ based on level of social anxiety or condition order. Second, although participants were not explicitly evaluated in the Anonymous and Presence conditions, they may have felt evaluation from the application of psychophysiological measures that signify monitoring is occurring (in addition to the implicit evaluation likely to be felt in any psychology study). While this effect was presumably constant across all conditions, minimizing impact on the observed condition differences, it is likely that we were unable to capture performance in a purely non-evaluative state. Third, there is some question about what aspect of working memory the n-back task measures, as it often shows weak or modest correlations with other working memory measures, including complex span tasks (Jaeggi et al., 2010; Kane et al., 2007). It has been hypothesized that n-back tasks tap into recognition memory, whereas complex span tasks tap into serial recall (Jaeggi et al., 2010; Kane et al., 2007). Thus, future studies are warranted to explore whether SET interferes with performance efficiency across other facets of working memory. Finally, this study does not allow us to make claims about *why* performance interference occurs for socially anxious individuals during SET. For example, socially anxious people may be taking extra time to be careful, undergoing interference from negative self-talk, or experiencing distraction from scanning the environment for threat cues. Future research should focus on identifying mediators of this task disruption.

Conclusions

This study showed that individuals high in FNE experienced performance interference on a working memory task during ongoing, overt social-evaluative threat compared to mere social presence or being alone, but only during high task difficulty. This interference was evident in task efficiency (RTs) but not task effectiveness (accuracy), suggesting the negative effects of evaluation on individuals high in FNE may most influence efficiency. Participants did not differ in task performance during social presence compared to being alone but did report greater feelings of anxiety and being monitored during social presence, irrespective of FNE. In sum, the detrimental effects of FNE on performance may extend beyond the social domain to impair performance on non-social, cognitive tasks as well, even when the environment is minimally evaluative as occurs in many settings. This suggests potentially ubiquitous effects on performance that educators, clinicians, researchers, and socially anxious individuals may need to consider before making any inferences about differences in ability for persons with social anxiety.

Acknowledgements

This work was supported by an NIMH grant (R01MH080725) awarded to J. A. Coan and an NIMH grant (R34MH106770), as well as a Templeton Science of Prospection Award, to B. A. Teachman.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testin and interpreting interactions. Multiple regression: Testing and interpreting interactions*.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D. M., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv Preprint arXiv*:1506.04967, 1–27. http://doi.org/arXiv:1506.04967

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Baumeister, R. F., & Showers, C. J. (1986). A review of paradoxical performance effects: Choking under pressure in sports and mental tests. European *Journal of Social Psychology*, 16(4), 361–383. https://doi.org/10.1002/ejsp.2420160405

Beilock, S. L., Rydell, R. J., & McConnell, A. R. (2007). Stereotype threat and working memory: mechanisms, alleviation, and spillover. *Journal of Experimental Psychology. General*, 136(2), 256–76. https://doi.org/10.1037/0096-3445.136.2.256

Berggren, N., & Derakshan, N. (2013). Attentional control deficits in trait anxiety: Why you see them and why you don't. *Biological Psychology*. https://doi.org/10.1016/j.biopsycho.2012.03.007

Brown, T. a, Chorpita, B. F., & Barlow, D. H. (1998). Structural relationships among dimensions of the DSM-IV anxiety and mood disorders and dimensions of negative affect, positive affect, and autonomic arousal. *Journal of Abnormal Psychology*, 107(2), 179–192. https://doi.org/10.1037/0021-843X.107.2.179

Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. *Social Phobia: Diagnosis, Assessment, and Treatment*, 41, 68.

Cody, M. W., Clerkin, E. M., Stevens, E. S., Gasser, M. L., Pasciuti, M. L., & Teachman, B. A. (2014). Social anxiety disorder and global/local performance on a visuospatial processing task. *Journal of Experimental Psychopathology*. https://doi.org/10.5127/jep.035013

Elzinga, B. M., & Roelofs, K. (2005). Cortisol-induced impairments of working memory require acute sympathetic activation. *Behavioral Neuroscience*, 119(1), 98–103. https://doi.org/10.1037/0735-7044.119.1.98

Eysenck, M. W., & Derakshan, N. (2011). New perspectives in attentional control theory. *Personality and Individual Differences*, 50(7), 955–960. https://doi.org/10.1016/j.paid.2010.08.019

Gevins, A., & Cutillo, B. (1993). Spatiotemporal dynamics of component processes in human working memory. *Electroencephalography and Clinical Neurophysiology*, 87(3), 128–143. https://doi.org/10.1016/0013-4694(93)90119-G

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, 18(4), 394-412. https://doi.org/10.1080/09658211003702171

Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attention control, and the n-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 615-622. https://doi.org/10.1037/0278-7393.33.3.615

Kashdan, T. B., & Roberts, J. E. (2004). Social Anxiety's Impact on Affect, Curiosity, and Social Self-Efficacy During a High Self-Focus Social Threat Situation. *Cognitive Therapy and Research*, 28(1), 119–141. https://doi.org/10.1023/B:COTR.0000016934.20981.68

Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781849209366

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2014). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). *R Package Version*. http://doi.org/http://CRAN.R-project.org/package = lmerTest

Leary, M. R. (1983). A Brief Version of the Fear of Negative Evaluation Scale. *Personality and Social Psychology Bulletin*, 9(3), 371–375. https://doi.org/10.1177/0146167283093007

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales. Psychology Foundation of Australia* (Vol. 56). https://doi.org/10.1016/0005-7967(94)00075-U

Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization* - R. Duncan Luce - Google Books. New York: Oxford University Press. Retrieved from https://books.google.com/books?hl = en&lr = &id = WSmpNN5WCw0C&oi = fnd&pg = PA1&ots = XqLSR9a_eN&sig = R0LFVPKxicc1GztismUerEckfSQ#v = onepage&q&f = false

Maresh, E. L., Allen, J. P., Coan, J. A., Watson, D., Clark, L., Gray, = EK, … Ernst, M. (2014). Increased default mode network activity in socially anxious individuals during reward processing. *Biology of Mood & Anxiety Disorders*, 4(1), 7. https://doi.org/10.1186/2045-5380-4-7

Mathieu, J. E., Aguinis, H., Culpepper, S. A., & Chen, G. (2012). Understanding and estimating the power to detect cross-level interaction effects in multilevel modeling. *Journal of Applied Psychology*, 97(5), 951–966. https://doi.org/10.1037/a0028380

Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy*, 36(4), 455–470. https://doi.org/10.1016/S0005-7967(97)10031-6

Mesagno, C., Harvey, J. T., & Janelle, C. M. (2012). Choking under pressure: The role of fear of negative evaluation. *Psychology of Sport & Exercise*, 13, 60–68. https://doi.org/10.1016/j.psychsport.2011.07.007

Moriya, J., & Sugiura, Y. (2012). High visual working memory capacity in trait social anxiety. *PLoS ONE*, 7(4). https://doi.org/10.1371/journal.pone.0034244

Nich, C., & Carroll, K. (1997). Now you see it, now you don't: a comparison of traditional versus random-effects regression models in the analysis of longitudinal follow-up data from a clinical trial. *J Consult Clin Psychol,* 65(2), 252–261. https://doi.org/10.1037/0022-006X.65.2.252

Norton, P. J., & Hope, D. A. (2001). Kernels of truth or distorted perceptions: Self and observer ratings of social anxiety and performance. *Behavior Therapy*, 32(4), 765–786. https://doi.org/10.1016/S0005-7894(01)80020-4

O'Toole, M. S., Pedersen, A. D., Hougaard, E., & Rosenberg, N. K. (2015). Neuropsychological test performance in social anxiety disorder. Nordic *Journal of Psychiatry*, 69(6), 444–452. https://doi.org/10.3109/08039488.2014.997288

Oei, N. Y. L., Everaerd, W. T. a M., Elzinga, B. M., van Well, S., & Bermond, B. (2006). Psychosocial stress impairs working memory at high loads: an association with cortisol levels and memory retrieval. *Stress (Amsterdam, Netherlands)*, 9(3), 133–141. https://doi.org/10.1080/10253890600965773

R Core Team. (2013). R Core Team. *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria*., ISBN 3-900051-07-0, URL http://www.R-project.org/. Retrieved from http://www.mendeley.com/research/r-language-environment-statistical-computing-96/\npapers2://publication/uuid/A1207DAB-22D3-4A04-82FB-D4DD5AD57C28

Rapee, R. M., & Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy*, 35(8), 741–756. https://doi.org/10.1016/S0005-7967(97)00022-3

Redick, T. S., & Lindsey, D. R. (2013). Complex span and n-back measures of working memory: a meta-analysis. *Psychonomic Bulletin & Review*, 20(6), 1102-1113. https://doi.org/10.3758/s13423-013-0453-9

Ree, M. J., French, D., MacLeod, C., & Locke, V. (2008). Distinguishing cognitive and somatic dimensions of state and trait anxiety: development and validation of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA). *Behavioural and Cognitive Psychotherapy*, 36(3), 313–332. https://doi.org/10.1017/S1352465808004232

Rodebaugh, T. L., Heimberg, R. G., Brown, P. J., Fernandez, K. C., Blanco, C., Schneier, F. R., & Liebowitz, M. R. (2011). More reasons to be straightforward: Findings and norms for two scales relevant to social anxiety. *Journal of Anxiety Disorders*, 25(5), 623–630. https://doi.org/10.1016/j.janxdis.2011.02.002

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. Psychological Review, 115(2), 336–56. https://doi.org/10.1037/0033-295X.115.2.336

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

Stoet, G., Windows, M., & Macintosh, A. (2010). PsyToolkit: a software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42(4), 1096–104. https://doi.org/10.3758/BRM.42.4.1096

Uziel, L. (2007). Individual differences in the social facilitation effect: A review and meta-analysis. *Journal of Research in Personality*, 41(3), 579–601. https://doi.org/10.1016/j.jrp.2006.06.008

Wenzel, A., & Holt, C. S. (2003). Social-evaluative threat and cognitive performance in socially anxious and nonanxious individuals. *Personality and Individual Differences*. https://doi.org/10.1016/S0191-8869(02)00044-2