5-3-2021

A Novel Task to Assess Outcome Probability Bias for Social Anxiety

Amanda A. Draheim

Georgia State University

A NOVEL TASK TO ASSESS OUTCOME PROBABILITY BIAS FOR SOCIAL ANXIETY

by

AMANDA DRAHEIM

Under the Direction of Page L Anderson, PhD

ABSTRACT

People with social anxiety disorder tend to expect that the likelihood of social embarrassment or negative judgment is much higher when anticipating a social encounter than people without the disorder (Lucock & Salkovskis, 1988). This outcome probability bias is theorized to contribute to the development and maintenance of symptoms (Clark, 2001; Heimberg et al., 2010). The causal role of outcome probability bias in social anxiety disorder is unknown, however, because all known studies have relied exclusively on paper-and-pencil questionnaires to operationalize this construct. The present study aimed to assess the reliability, validity, and factor structure of a novel computer task designed to measure outcome probability bias in social anxiety. Important components of this task included incorporation of social images and assessment of outcome probability bias resulting from both automatic and controlled levels of processing. Results from a pilot study indicated the images in the task were appropriate but modifications were warranted for assessment of outcome probability bias at an automatic level of processing. Results from the main study indicated good to excellent internal consistency among social images ($\alpha = 0.86 – 0.96$). Correlations were consistent with the nomological network. Outcome probability bias task ratings were higher in response to social items compared to nonsocial items and among people with social anxiety symptom scores above a clinical cutoff compared to people with scores below the cutoff (all p values < 0.05). Social images in the task also predicted self-reported symptoms of social anxiety, self-reported safety behaviors during a behavioral avoidance test, and subjective distress during a behavioral avoidance test (all p values < 0.05). The task did not, however, significantly predict performance on a measure of behavioral avoidance. Exploratory factor analyses revealed a tripartite factor structure. These findings offer preliminary support for the reliability, convergent validity, discriminant validity, construct validity, and criterion validity of the outcome probability bias task. This task may potentially be used in future research for multimodal assessment and experimental manipulation of outcome probability bias in social anxiety.

INDEX WORDS: Social anxiety, Outcome probability bias, Mental imagery, Automatic processing, Psychometric properties, Factor structure

A NOVEL TASK TO ASSESS OUTCOME PROBABILITY BIAS FOR SOCIAL ANXIETY

by

AMANDA DRAHEIM

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2020

A NOVEL TASK TO ASSESS OUTCOME PROBABILITY BIAS FOR SOCIAL ANXIETY

by

AMANDA DRAHEIM

Committee Chair: Page Anderson

Committee: Michael Beran

Lindsey Cohen

Erin Tone

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2021

DEDICATION

This dissertation is dedicated to my husband, Christopher Draheim, and my mother, Denise Benbow.

ACKNOWLEDGMENTS

First and foremost, I would like to express my gratitude for my mentor, Dr. Page Anderson, for bringing me into this program and giving me the freedom to pursue my passion. I would also like to thank my co-mentor, Dr. Erin Tone, and my committee members, Dr. Lindsey Cohen and Dr. Michael Beran, for their sound feedback and support of this project.

I am incredibly thankful for the support of my husband, Christopher Draheim, who was instrumental in programming and troubleshooting the outcome probability bias task. He also consulted on some of the statistical analyses. I would also like to acknowledge Jason Tsukahara for his assistance with task programming.

My thanks to my fellow lab mates for their feedback on task development, study procedures, study analyses, and presentations: Tony Molloy, Nancy Su, Donovan Ellis, and Deah Abbott.

Several undergraduate lab members provided invaluable assistance with data collection, including Kate Niven, Shimarith Wallace, Aleysha Delpin, Riven Mendoza, and Taliyah Mosley. Thank you.

Funding support was received from the Brains and Behavior Fellowship and the Health Resources and Services Administration.

I'd also like to thank all of the GSU students who participated in my study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Outcome probability bias, defined as a person's tendency to overestimate the likelihood of a feared outcome (Foa & Kozak, 1986), has long been theorized to contribute to the development and maintenance of social anxiety disorder (Clark, 2001; Clark & Wells, 1995; Heimberg et al., 2010; Rapee & Heimberg, 1997). In line with these theories, widely used cognitive-behavioral treatments for social anxiety disorder such as the Unified Protocol and Exposure Group Therapy include interventions designed to target and reduce outcome probability bias (Barlow et al., 2017; Hofmann, 2004). Over the past thirty years, many studies have demonstrated that outcome probability bias declines significantly from pre- to post-treatment (Calamaras et al., 2015; Foa et al., 1996; Gregory et al., 2015; Hoffart et al., 2009; Lucock & Salkovskis, 1988; McManus et al., 2000; Moscovitch et al., 2012; O'Toole et al., 2015; Poulton & Andrews, 1996; Rapee et al., 2009; Smits et al., 2006; Taylor & Alden, 2008; Voncken & Bögels, 2006). Further, evidence suggests that outcome probability bias is a specific mechanism through which cognitive-behavioral therapy for social anxiety disorder contributes to symptom improvement (Calamaras et al., 2015; McManus et al., 2000; O'Toole et al., 2015; Smits et al., 2006; Taylor & Alden, 2008), though some studies have reported null results (Foa et al., 1996; Moscovitch et al., 2012; Rapee et al., 2009).

A major limitation of this literature, however, is that all studies exclusively relied on paper-and-pencil questionnaires to assess the construct of outcome probability bias. In addition, all analyses were correlational in nature. As such, the causal role of outcome probability bias in treatment for social anxiety disorder is unknown. This represents a notable gap in the literature because accurate understanding of the mechanisms of psychological treatments—simply put, why people get better—is needed to refine theory and enhance treatment outcomes (Kazdin, 2007). Identification of treatment mechanisms may also promote dissemination efforts. For example, by streamlining treatments to only include interventions that target known mechanisms (Cooper et al., 2017).

Further, the use of multimodal assessment would strengthen the scientific support for the mechanistic role of outcome probability bias within social anxiety disorder's treatment (Campbell & Fiske, 1959) because researchers' confidence in a relation between variables is enhanced when the effect is demonstrated using a variety of measures and settings. Computer tasks represent one alternative approach to paper-and-pencil questionnaires that may be used to measure and/or experimentally manipulate a construct of interest. For example, other forms of bias that are characteristic of anxiety disorders such as attention, interpretation, and memory biases have been measured via computer tasks (for a recent review of computer tasks designed to assess interpretation bias see Schoth & Liossi, 2017). Furthermore, computer tasks designed to assess attention and interpretation biases have been adapted to allow for experimental manipulation of these biases, termed cognitive bias modification (Hertel & Mathews, 2011). Cognitive bias modification paradigms have been used to test causal models of anxiety and depression and have demonstrated a small but significant effect on symptoms of anxiety when employed as therapeutic interventions (Hallion & Ruscio, 2011). A computer task designed to assess outcome probability bias in social anxiety could therefore allow for multimodal assessment of the construct. It could also serve as a tool for future experimental manipulation and intervention.

An outcome probability bias computer task may also allow for incorporation of elements that paper-and-pencil questionnaires can't capture. For example, by including visual rather than verbal stimuli researchers can control the variability in mental imagery participants experience when providing responses about their judgments of outcome probability. A computer task can also allow for evaluation of outcome probability bias at different levels of processing, which can inform the mechanisms of this construct. As such, the proposed study aims to develop a computer task to operationalize outcome probability bias in social anxiety and assess its reliability, validity, and factor structure.

1.1 Outcome Probability Bias in Anxiety and Related Disorders

Outcome probability bias has been studied extensively over the past fifty years, and reductions in outcome probability bias have long been considered a primary mechanism of change in cognitive-behavioral treatments for anxiety disorders. Carr (1974) and Beck (1976) first introduced the construct. They proposed that anxiety and related disorders are characterized by exaggerated perceptions of danger, which may be caused by subjective overestimations of the likelihood of a feared outcome (later termed outcome probability bias). Outcome probability bias has been identified in a variety of anxiety and related disorders including generalized anxiety disorder (Butler & Mathews, 1983), social anxiety disorder (Andrews et al., 1994; Foa et al., 1996; Lucock & Salkovskis, 1988; McManus et al., 2000; Poulton & Andrews, 1996; Uren et al., 2004, Voncken et al., 2003), agoraphobia (McNally & Foa, 1987; Poulton & Andrews, 1996), specific phobia (Menzies & Clarke, 1995), panic disorder (Uren et al., 2004) and acute stress disorder (Warda & Bryant, 1998). Across these studies, outcome probability bias was higher among untreated anxious individuals when compared to a non-anxious control group. Further, the content of the bias was specific to the type of disorder (Gilboa-Schechtman et al., 2000; Poulton & Andrews, 1996; Uren et al., 2004). For example, people with social anxiety disorder demonstrated outcome probability bias for social events but not for physical events (Uren et al., 2004). Notably, all of these studies used paper-and-pencil questionnaires to measure outcome probability bias.

Cognitive-behavioral treatments for anxiety and related disorders are thought to reduce outcome probability bias by means of cognitive restructuring and/or exposure. Cognitive restructuring is an intervention that encourages people to assess the validity of their thoughts and assumptions (Beck et al., 2005). An example of a technique designed to reduce outcome probability bias using cognitive restructuring is the pie method (Voncken & Bögels, 2006). When using the pie method, clients are instructed to consider a feared situation and estimate the likelihood of all possible outcomes. The probabilities for all included outcomes must add up to 100 percent. Clients often reduce their initial estimates of the probability of a negative outcome in order to allow for the alternatives. Voncken and Bögels (2006) presented data from a pilot study of 13 people with social anxiety disorder who reported post-treatment reductions in outcome probability bias following nine sessions designed to target this bias using cognitive restructuring techniques.

Exposure to feared stimuli may also reduce outcome probability bias. Exposure therapy is a treatment that arose from classical conditioning research and is based on the idea that anxiety and related disorders are maintained by avoidance (Wolpe, 1958). Classical conditioning experiments that have informed the development of exposure therapy were conducted on animals and typically involved conditioning a fear response. For example, following multiple presentations of a neutral stimulus (a tone) paired with an unconditioned stimulus (shock), animals learn to fear the tone and demonstrate a fear response (freezing) even when the tone is presented in the absence of shock. Fear responding declines during extinction, where the conditioned stimulus (the tone) is presented in the absence of shock repeatedly. Research on the factors that promote successful extinction of fear responding is the basis for exposure therapy in humans. During exposure therapy, clients are encouraged to repeatedly face feared stimuli in a variety of contexts for prolonged periods (Foa & Kozak, 1986). For example, a client with a specific phobia of spiders would be encouraged to repeatedly look at, approach, and eventually hold a spider in multiple settings.

Foa and Kozak (1986) proposed that reductions in outcome probability bias may represent a mechanism through which exposure therapy leads to improvement in symptoms of anxiety. The authors suggested that the reductions in fear that occur over the course of multiple exposures to the feared stimulus (i.e. between-session habituation) could lead to more accurate predictions of the likelihood of a feared outcome. The authors posited that these reductions in outcome probability bias that result from exposure therapy would, in turn, lead to symptom reduction.

Basic science research demonstrates that manipulation of expectations about the likelihood of a feared outcome (achieved by varying the rates of association between the unconditioned and the conditioned stimuli) influences the effectiveness of fear response extinction (Gallistel & Gibbon, 2000; Rescorla & Wagner, 1972). More specifically, greater expectancy violations are associated with decreased fear responding. For example, Rescorla and Wagner (1972) describe a study where 36 rabbits underwent fear conditioning using three different stimuli. During the acquisition phase, one stimulus (randomly assigned as either a flashing light or a chest vibration) was trained to be highly salient when paired with an eye shock at irregular intervals over the course of 224 trials. A second, less salient stimulus (either the flashing light or chest vibration, whichever had not been trained as highly salient), was paired with the eye shock at irregular intervals over 28 trials. As such, the highly salient stimulus was associated with greater expectation for negative outcome (comparable to greater outcome probability bias). All rabbits also underwent 224 trials during which a tone was paired with a shock at irregular intervals (therefore the tone was highly salient across groups). During the extinction phase, the tone was paired with either the highly salient stimulus (in half of the rabbits) or the low salience stimulus (the other half) in the absence of punishment for 32 trials. Arguably, greater expectancy violation would take place when the highly salient stimulus was paired with the tone in the absence of punishment relative to the low salient stimulus. Consistent with this hypothesis, the researchers found that animals' fear responses declined significantly faster during extinction of the highly salient stimulus pair. Reacquisition training, during which the tone was again paired with shock, was used to test the effectiveness of extinction learning. The researchers found that return of the fear response took longer when the tone had been paired with the highly salient stimulus during extinction. These findings suggest that fear extinction was more resistant to reacquisition as a result of greater expectancy violation.

Based on these results, Craske et al. (2008) argue that expectancy violations are essential for effective exposure therapy. Violations in expectations may be achieved in humans by means of presentation of a feared stimulus in the absence of punishment (i.e. exposure) or through discussion and logic (i.e. cognitive interventions such as cognitive restructuring). Reductions in outcome probability bias may therefore represent a mechanism of expectancy violation in cognitive restructuring and in exposure therapy.

1.2 Outcome Probability Bias in Social Anxiety Disorder

Expectancy violations may be particularly difficult to achieve during exposure therapy for social anxiety disorder. Social anxiety disorder is characterized by persistent and excessive fear of social interactions in which there is potential for negative evaluation (American Psychiatric Association, 2013). Many social outcomes tend to be ambiguous, and the baseline likelihood that a feared outcome might occur (e.g. someone turns you down when you ask them out on a date) is higher compared to feared outcomes for other anxiety disorders (e.g. having a heart attack, a common fear among people with panic disorder; Foa & Kozak, 1986). As such, outcome probability bias may be more resistant to change in treatment for social anxiety disorder relative to other anxiety disorders (Poulton & Andrews, 1996). If so, further research on ways to effectively measure and reduce outcome probability bias specifically for people with social anxiety disorder would be useful.

Outcome probability bias has been identified as a factor that contributes to the development and maintenance of symptoms in cognitive models of social anxiety disorder (Clark, 2001; Clark & Wells, 1995; Heimberg et al., 2010; Rapee & Heimberg, 1997). Clark and Wells' (1995) model posits that distorted assumptions and beliefs about social performance lead to exaggerated perceptions of social danger – that a socially anxious person believes they will likely behave "in an inept and unacceptable fashion" (p. 69). According to this model, perception of social danger represents a core maintenance factor of social anxiety that contributes to and is reinforced by self-focused attentional shift as well as cognitive, somatic, and behavioral symptoms of social anxiety. Hoffart et al. (2009) explicitly stated that outcome probability bias likely contributes to the development and maintenance of the perception of social danger identified in the Clark and Wells (1995) model. Rapee and Heimberg (1997) also propose that outcome probability bias contributes to fear in social situations, which subsequently exacerbates symptoms of social anxiety and maintains the disorder. Their model emphasizes the role of imagery. Specifically, it postulates that people with social anxiety disorder form mental representations of themselves from a third person perspective (i.e. seeing oneself from the perspective of an audience). Mental imagery, in turn, predicts outcome probability bias. As these cognitive models have been updated to reflect more recent research findings, outcome probability bias has been retained as central to the development and maintenance of social anxiety disorder (Clark, 2001; Heimberg et al., 2010).

1.3 Mental Imagery in Social Anxiety Disorder

Many studies have supported Heimberg et al.'s (2010) emphasis on the role of mental imagery in social anxiety disorder. For example, researchers have found that socially anxious individuals demonstrate enhanced emotional responding when presented with images relative to words (Acosta & Vila, 1990; Holmes & Mathews, 2005; Holmes et al., 2006; Lang et al., 1983; Vrana et al., 1986). Evidence suggests individuals with social anxiety disorder report experiencing higher frequencies of mental images in the context of social encounters relative to non-anxious controls (Hackmann et al., 1998). Further, spontaneous images commonly described by socially anxious individuals typically come from an observer perspective and are often more negative and distorted (Hackmann et al., 1998; Wells et al., 1998). In addition, engagement in negative mental representations using the observer perspective is associated with increased symptoms of social anxiety and impairs performance (Hirsch et al., 2003; Hirsch et al., 2004; Spurr & Stopa, 2003; Stopa & Jenkins, 2007; Vassilopoulos, 2005). Taken together, mental imagery appears to play a central role in anticipatory anxiety experienced by individuals with social anxiety disorder, yet no known studies have evaluated outcome probability bias using visual stimuli.

1.4 Mechanisms of Outcome Probability Bias

An image-based measure of outcome probability bias should be informed by theoretical mechanisms of the construct (Embretson, 1983). Butler and Mathews (1983) published results from the first measure of outcome probability bias and posited that maladaptive judgments about the likelihood of a feared outcome may result from employment of the availability and representativeness heuristics.

The availability heuristic refers to people's tendency to estimate the likelihood of an outcome based on the ease with which the outcome comes to mind (Tversky & Kahneman, 1974). Certain outcomes may come to mind more easily than others due to increased familiarity, salience, or imaginability. For example, after hearing several news reports about shark attacks, people may overestimate the likelihood of being attacked by a shark when considering a trip to the beach. This overestimation is arguably due to the increased ease with which lethal shark attacks come to mind after hearing about them in the news. Tversky and Kahneman (1973) conducted a series of experiments designed to demonstrate how use of the availability heuristic can lead to inaccurate probability estimates. In one such experiment, participants were instructed to estimate whether a given letter of the alphabet, such as the letter r, appears more commonly in the first or the third position of a typical English word. On average, participants reported (inaccurately) that the letter r appears more commonly in the first position. The authors concluded that this common mistake occurs because words that begin with the letter r come to mind more easily than words with r in the third position. Regarding anxiety disorders, Butler and Mathews (1983) posited that feared outcomes may come to mind more easily than alternative outcomes.

The representativeness heuristic also contributes to inaccurate estimates of probability. When relying on the representativeness heuristic, people tend to judge an outcome as more likely to occur when the outcome more closely matches their mental representation of its category. For example, when asked to judge which series of coin toss outcomes was more likely, HTTHTH or HHHTTT, people tended to assume that the first outcome was more probable than the latter despite the fact that both outcomes are equally likely. Tversky and Kahneman (1974) concluded that this inaccurate judgment of probability typically happens because the first outcome best matches common understanding of the randomness and proportion (1/2) of the coin toss process. According to Butler and Mathews (1983), individuals with an anxiety disorder may tend to associate a variety of stimuli with a mental category of fear. For instance, increased heartrate can result from several causes, but an individual with panic disorder may be more likely to assume that rapid heartrate is representative of a heart attack.

Kahneman (2011) argued that the availability and representativeness heuristics occur as a result of automatic processing, as described by the dual-process information-processing model (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977), which distinguishes between two different types of information-processing: automatic and controlled. Automatic processing can be defined as speeded processing that occurs with minimal effort and without conscious awareness. For example, a skilled reader can decode the letters in this sentence and access the meaning of these words quickly and with little effort. Controlled processing involves relatively slower and more intentional processing. For example, quickly counting all the commas and periods on this page would require conscious and sustained attention and effort. Research conducted with individuals with brain damage demonstrates the dissociation between automatic and controlled processing (Birnboim, 2003).

Studies evaluating the ease-of-retrieval effect support Kahneman's (2011) assertion that the availability heuristic manifests in automatic processing more so than controlled processing. Schwartz et al. (1991) reported an ease-of-retrieval effect when they found that people who were instructed to recall 6 instances of engaging in assertive behavior rated themselves as more assertive than those who were required to recall 12 examples. The authors attributed this difference to less difficulty in retrieving examples of assertiveness when the number of examples was low because easier retrieval facilitates the availability heuristic and biases judgments (Tversky & Kahneman, 1974). This effect was also demonstrated for judgments of probability. In a similar study, Raghubir and Menon (1998) found that people judged their risk of contracting AIDS as higher when it was easier for them to retrieve examples of AIDS-related behaviors from memory. Menon and Raghubir (2003) conducted a series of experiments designed to assess the automaticity of this ease-of-retrieval effect. In the first study, 133 undergraduate students saw an advertisement highlighting ten product features for a type of personal computer. Some students were instructed to recall two of the features (easy retrieval) and others were instructed to recall eight (difficult retrieval); participants in the easy retrieval condition rated the personal computer more favorably than participants in the difficult retrieval condition. This result replicated findings from Schwartz et al. (1991) and Raghubir and Menon (1998), demonstrating an ease-of-retrieval effect. In the second study, participants were informed that the task of recalling eight product features was difficult, which brought attention to the fact that ease-of-retrieval influences judgment. This discredited the informational value of the ease-of-retrieval cue by promoting more intentional thinking (i.e. controlled processing) and reversed the ease-of-retrieval effect. The authors concluded reversal of the ease-of-retrieval effect under conditions allowing for controlled processing supports the automaticity of this effect. Results from the third study revealed that the timing of the presentation of cue discreditation, before or after the recall task, influenced the ease-of-retrieval effect. When the difficulty of the task was communicated after the recall task was completed and students had already experienced difficulty in retrieval, evaluative judgments of the favorability of the product were immune to discounting information. This supports the uncontrollability of the influence of ease-of-retrieval on judgment, supporting automaticity. Finally, in the fourth study, discrediting information about the difficulty of the task failed to influence the impact of ease-of-retrieval on favorability judgments when participants were under high cognitive load, and thus were unable to engage in the controlled processing necessary to override responding that results from known bias. This finding further supports the conclusion that the ease-of-retrieval effect is dependent upon conditions where automatic processing predominates, and controlled processing is not possible. Taken together, the authors argued that these studies support the claim that the availability heuristic influences judgements by means of automatic processing.

Based on the impact ease-of-retrieval has on judgments of probability and likeability, simply imagining an event can make it seem more likely to occur. For example, Carroll (1978) conducted two experiments that demonstrated the effect of imagining on the availability heuristic in a sample of undergraduate students. In the first study, participants who imagined that Jimmy Carter would win the presidential election later judged that Carter was more likely to win than participants who were instructed to imagine Gerald Ford winning. This effect was present regardless of students' prior beliefs about the election. In the second study, participants imagined either a good or bad future college football season. Those who imagined a good season predicted that their team had a greater chance of achieving a major bowl bid than those who imagined a poor season.

The tendency to engage in negative mental imagery in anticipation of social encounters (Hackmann et al., 1998; Wells et al., 1998) may activate the availability heuristic and contribute to outcome probability bias among people with social anxiety disorder. In light of the findings discussed above (Carroll, 1978; Menon & Raghubir, 2003), outcome probability bias may therefore manifest at automatic levels of processing. To date, outcome probability bias has only been assessed using paper-and-pencil questionnaires, reflective of controlled processing. The present study is the first to assesses outcome probability bias at both automatic and controlled levels of processing.

1.5 Constructs Similar to Outcome Probability Bias

In addition to identifying the theoretical mechanisms of a construct of interest when developing a new measure, it is also important to identify conceptually similar constructs in the literature (Embretson, 1983). It is necessary to use theory to demonstrate how the construct of interest is distinguishable from other constructs in order to justify the need for a new measure. Identification of similar constructs is also useful for informing evaluation of convergent validity: the new measure should correlate with measures of theoretically related constructs (Campbell & Fiske, 1959; DeVellis, 2011). Two such constructs related to outcome probability bias include outcome cost bias and interpretation bias.

1.5.1 Outcome Cost Bias

Cognitive models of social anxiety disorder posit that a related, yet distinct construct, outcome cost bias, also contributes to the development and maintenance of symptoms of social anxiety disorder (Clark, 2001; Clark & Wells, 1995; Heimberg et al., 2010; Rapee & Heimberg, 1997). Outcome cost bias is defined as a person’s tendency to overestimate the negative consequences of a feared outcome (Foa & Kozak, 1986).

Several studies have compared the differential impact of outcome probability and cost biases on symptoms of social anxiety. Foa et al. (1996) reported that outcome cost bias was a stronger predictor of reduction in symptoms of social anxiety than was outcome probability bias. McManus et al. (2000) found that both outcome probability bias and outcome cost bias were equally predictive of symptom reduction. They also reported that reductions in outcome cost bias did not explain a significant degree of additional variance in symptom reduction after controlling for reductions in outcome probability bias. Notably, these studies used hierarchical linear modeling to explore these differences and were limited by small sample sizes. In addition, temporality was not established in these analyses (assessments only took place at pre- and post-treatment). Temporality is necessary to establish causal mediation (Kazdin, 2007). In contrast, two studies have used more rigorous methodological designs (e.g. longitudinal) and statistical methods (e.g. path analysis) to explore and compare the mechanistic role of outcome probability and cost biases in treatment for social anxiety disorder (Calamaras et al., 2015; Smits et al., 2006).

Using archival data from a randomized controlled trial ($n = 53$), Smits et al. (2006) demonstrated that reductions in outcome probability bias predicted symptom reduction; reductions in outcome cost bias, however, were the result of symptom reduction. The authors used a longitudinal design and cross-lagged panel analysis to test this mediation. Of note, treatment outcome was defined as reductions in fear levels following exposure sessions. Reductions in subjective reports of fear during exposure are not considered to be adequate predictors of symptom reduction (Craske et al., 2008), thus limiting the value of these findings.

Calamaras et al. (2015) also employed a cross-lagged panel design to assess the mediational effect of outcome probability and cost biases on social anxiety symptom improvement. They included a rival mediator in their analyses: the working alliance, defined as the quality of the relationship between the client and the therapist (Horvath et al., 2011). Self-report measures with known psychometric properties were used for all variables of interest. Analyses were conducted using combined archival data from a randomized controlled trial assessing the efficacy of virtual reality exposure therapy for social anxiety disorder (Anderson et al., 2013) as well as an uncontrolled trial examining the predictive relation between amygdala activity and symptom reduction (total $n = 86$). The authors found that early decline in outcome probability bias, but not outcome cost bias, was a significant mediator of symptom reduction while controlling for the working alliance.

Taken together, these findings support the mechanistic role of outcome probability bias, but not outcome cost bias, in cognitive-behavioral treatment for social anxiety disorder. They also support the dissociation of these two constructs. Though a computer measure of outcome cost bias could further facilitate evaluation of cognitive models of social anxiety disorder, this is outside the scope of the current study.

1.5.2 Interpretation Bias

It appears that outcome probability bias and interpretation bias may have been conflated by some researchers. Interpretation bias refers to the tendency to interpret ambiguous social stimuli as threatening. For example, lack of an immediate response to a text message sent to a romantic partner (ambiguous) can be interpreted in a benign manner (they are currently busy at work and will get back to me later) or a threatening manner (they are avoiding me because they want to break up). Negative interpretation bias has been identified as a potential mechanism for the development and maintenance of anxiety disorders, including social anxiety disorder (see Blanchette & Richards, 2010; Mathews & MacLeod, 2005).

Butler and Mathews' (1983) first assessment of cognitive processes in anxiety evaluated hypotheses related to interpretation bias, outcome cost bias, and outcome probability bias. Three different measures were used, one for each construct. For the interpretation bias measure, participants filled out a paper-and-pencil questionnaire that presented them with a variety of ambiguous scenarios. Participants wrote down their initial interpretation of the situation, then ranked three possible disambiguating explanations from most to least likely. One of the possible disambiguating explanations was threatening. More frequent selections of threatening explanations reflected interpretation bias. The measure of outcome cost involved a list of threatening items. Participants rated the items based on the prompt "how bad would it be for you?" Higher ratings reflected greater outcome cost bias. To measure outcome probability bias, participants rated the likelihood that positive and negative events would happen to themselves and someone else from "not at all likely" to "extremely likely." Participants' ratings of greater likelihood that negative events would happen to them rather than someone else reflected outcome probability bias. Current measures of outcome probability bias are very similar to this original measure, though they only include negative events (e.g. Outcome Probability Questionnaire; Uren et al., 2004). They only allow for measurement of these biases in response to a limited number of social situations and can only reflect bias present at a controlled level of processing. In addition, mental imagery participants experience in response to items on these questionnaires is likely to be highly variable, and some participants may struggle to generate mental imagery in response to a paper-and-pencil questionnaire format.

The simultaneous introduction of these constructs could imply that outcome cost and probability biases are subtypes of interpretation bias. Supporting this view, Mathews and MacLeod (1994) stated outcome probability bias reflects "patterns of judgment [that] reflect the operation in anxious and depressed individuals of an interpretative bias that favors emotionally negative interpretations of ambiguous information" (p. 31). They subsequently described different measures of interpretation bias, suggesting that outcome probability bias and interpretation bias are the same construct.

Outcome probability bias is arguably distinct from interpretation bias because it manifests *in anticipation of* a negative outcome and may result from more automatic processing relative to interpretation bias, which is commonly assessed *during* an ambiguous situation. In support of this view, Beck and Clark (1997) presented an information processing model of anxiety composed of three stages: initial registration, immediate preparation, and secondary elaboration. Initial registration is considered to be driven by automatic processing, immediate preparation by a combination of automatic and controlled processing, and secondary elaboration primarily by controlled processing. The second stage, immediate preparation, includes “a constriction or narrowing of cognitive processing that leads to certain biases and inaccuracies” such that “there is an overestimation of the probability and severity of the threatening situation” (p. 53). The subsequent secondary elaboration stage involves a “secondary appraisal process” (p. 53) where people may interpret the level of threat and evaluate their ability to cope. The authors’ description of this model indicates that outcome probability bias manifests in the immediate preparation stage and interpretation bias manifests in the secondary elaboration stage. Given Beck and Clark’s (1997) assertion that these stages differ in the degree to which automatic processing predominates, we might expect outcome probability bias to manifest at both automatic and controlled levels of processing and interpretation bias to only manifest at controlled levels of processing.

Automatic processes are often assessed by instructing participants to respond rapidly, before conscious processing can influence performance. Manipulation of stimulus presentation rates may therefore be used to distinguish between automatic and controlled processes. Calvo and Castillo (1997) conducted a series of experiments designed to evaluate the extent to which interpretation bias is a result of automatic processing among a sample of undergraduate students with high trait anxiety. Participants were instructed to read ambiguous sentences that were designed to prime a threatening interpretation (e.g., “At night, the old woman was crossing the motorway on foot when a lorry approached her at high speed"). Sentences were presented at a rapid pace designed to minimize exposure time to the words and curtail controlled processing. Participants subsequently read disambiguating sentences that contained a single target word (e.g., *death*), which resolved the scenario in either a threatening or non-threatening manner. Participants were instructed to read the target word out loud as quickly as possible without making any errors. The target word was presented for either 500ms or 1,250ms duration. Presence of interpretation bias (operationalized as faster naming of the threatening words and slower naming of non-threatening words) at the shorter or longer durations reflected automatic or controlled processing, respectively. The authors found that interpretation bias was present with the longer, but not the shorter, stimulus duration. Consistent with Beck and Clark's (1997) model, Calvo and Castillo (1997) concluded that interpretation bias results only from controlled processing. Several other studies have also found that interpretation bias is not present at short (500ms) stimulus durations (Calvo, 2000; Calvo & Castillo, 1998; Calvo & Castillo, 2001; Calvo et al., 1997; Calvo et al., 1999; Richards & French, 1992). These findings support the conclusion that interpretation bias manifests in the secondary elaboration stage and results from controlled, but not automatic processing.

Given that outcome probability bias may manifest in the immediate preparation stage and result from automatic *and* controlled levels of processing (Beck & Clark, 1997), we may expect outcome probability bias to manifest following rapid and slow presentations of stimuli. This would support the dissociation of interpretation bias from outcome probability bias, but is not possible to test with current measures of outcome probability bias.

1.6 Important Considerations for an Outcome Probability Bias Computer Task

Two important factors to consider in the development of this task included the incorporation of mental imagery of social situations that elicit an observer perspective as well as assessment of outcome probability bias at both automatic and controlled levels of processing. Given that emotional responding is enhanced in socially anxious individuals when presented with images relative to words (Acosta & Vila, 1990; Holmes & Mathews, 2005; Holmes et al., 2006; Lang et al., 1983; Vrana et al., 1986) and that socially anxious individuals tend to imagine social situations using third person imagery (Hackmann et al., 1998; Wells et al., 1998), incorporation of social images that elicit an observer perspective may enhance the ecological validity of the present task. In addition, Beck and Clark’s (1997) information processing model for anxiety suggests that outcome probability bias may manifest at automatic and controlled levels of processing. The present task was designed to assess bias that may result from automatic and/or controlled processing by varying the duration of stimulus presentation rates.

1.7 Present Study

A pilot study was conducted to assess the feasibility of images included in the task and to evaluate whether the automatic block (involving rapid presentation of images) successfully evaluated automatic responding with a minimum of controlled processing. Second, the main study was conducted to evaluate the task’s reliability, validity, and factor structure in a larger sample. In both studies, a secondary aim was to evaluate the role of the availability heuristic in outcome probability bias.

1.7.1 Pilot Study

The pilot study aimed to evaluate the following: 1) internal consistency, 2) the effect of the outcome probability bias task on state anxiety, 3) adequate distinction between social and nonsocial images, 4) differences in outcome probability bias task ratings among those with symptoms of social anxiety above and below a clinical cutoff score, and 5) whether social images from the outcome probability bias task were related to paper-and-pencil measures of outcome probability bias and symptoms of social anxiety. Participants also provided feedback about the images and the task design (e.g. if it was easy to imagine themselves in the social situations presented by the images).

Internal consistency is a measure of reliability and refers to the degree of correlation among items in a measure (Cronbach, 1970). Items that measure the same construct should demonstrate strong intercorrelations. In the pilot study, individual images that reduced internal consistency or demonstrated weak correlation with the mean rating for that stimulus type would be removed. In addition, I hypothesized that: 1) state anxiety would increase following completion of the outcome probability bias task for all participants, 2) outcome probability bias task ratings for social images would be higher relative to nonsocial images among all participants, 3) ratings for social images would be higher for those reporting symptoms of social anxiety above the clinical cutoff compared to those with symptoms below the cutoff, and 4) social images would be related to paper-and-pencil measures of outcome probability bias and symptoms of social anxiety, but nonsocial images would not.

The stimulus presentation interval during the automatic block (involving rapid presentation of images) was evaluated by accuracy and confidence ratings for comprehension questions that followed the automatic trials. Accuracy ratings above the level of chance would likely indicate controlled processing was not sufficiently minimized. Participants also provided anecdotal feedback about the automatic trials.

Regarding the role of the availability heuristic in outcome probability bias, I hypothesized that participants who were familiar with the images would respond to automatic trials with higher outcome probability bias task ratings relative to those who were unfamiliar with the images. Half of the participants were familiar with the images because they completed the automatic trials after rating the same images at slower presentation rates. The other half had never been exposed to the images before completing the automatic trials and were therefore unfamiliar with the images. Because familiarity promotes engagement of the availability heuristic at the automatic level of processing (Kahneman, 2011), higher ratings in response to automatic trials among the half of participants who were familiar with the images would support the effect of the availability heuristic on outcome probability bias, as Butler and Mathews (1983) posited.

1.7.2 Main Study

The aim of the main study was to evaluate the reliability, convergent validity, discriminant validity, construct validity, criterion validity, and factor structure of the outcome probability bias task. Reliability was assessed with internal consistency. Convergent and discriminant validity were evaluated by examining the strength of associations between conceptually similar and distinct constructs, respectively (Campbell & Fiske, 1959; DeVellis, 2011). Stronger associations were expected between the outcome probability bias task and conceptually similar measures (paper-and-pencil measures of outcome probability and cost biases) relative to conceptually distinct measures (paper-and-pencil measures of symptoms of depression and stress).

Construct validity refers to a measure's operationalization: does it truly measure what it is intended to measure? (Cronbach & Meehl, 1955; DeVellis, 2011; Embretson, 1983). In the present study, people scoring above the social anxiety symptom clinical cutoff score were expected to report a greater likelihood of being embarrassed, humiliated, or negatively judged than people scoring below the clinical cut-off. I expected this effect would be present in response to social images but not for nonsocial control images. Evaluation of theorized mechanisms of outcome probability bias may also support construct validity. Butler and Mathews (1983) posited that outcome probability bias results from the availability heuristic, which manifests at the automatic level of processing and is enhanced by familiarity (Tversky & Kahneman, 1974). As such, higher outcome probability bias task ratings in response to rapidly presented familiar images relative to rapidly presented novel images would support the theorized effect of the availability heuristic on outcome probability bias.

Criterion validity refers to the extent to which a measure is associated with an outcome (DeVellis, 2011). In the present study, I expected that higher scores on the outcome probability bias task would predict higher self-report and behavioral symptoms of social anxiety and self-reported safety behaviors. If these relations remained significant while controlling for interpretation bias, stress and depression, this would provide additional support for the discriminant validity of the task. This result would also support the dissociation of outcome probability bias and interpretation bias.

Factor structure of the outcome probability task was assessed using exploratory factor analysis. Given that this was an exploratory evaluation of a novel measure, the only hypothesis I identified was that social and nonsocial images might represent two distinct factors.

2 METHOD

2.1 Outcome Probability Bias Task

The outcome probability task contained three types of stimuli: social images that elicited an observer perspective (social 3rd person), social images that elicited a first-person perspective (social 1st person), and images that did not have a social component (control). Details about the stimuli, including the images themselves, are included in Appendix A. Some images (primarily control images) came from the Set of Fear Inducing Pictures (SFIP) database (Michałowski et al., 2017) or the Nencki Affective Picture System database (NAPS; Marchewka et al., 2014). The SFIP and NAPS are composed of several high-quality images with established fear, valence, and arousal ratings. The SFIP database includes 40 social images shown to elicit higher levels of fear among undergraduate students with high levels of social anxiety. All control images in the present outcome probability bias task came from the SFIP and the NAPS. Some social images also came from the SFIP (12 images) or the NAPS (three images). The remainder of the social images were stock photos downloaded from Canva (https://www.canva.com/).

Social images depicted commonly feared and avoided social situations identified in questionnaires used to assess socially anxious populations with known psychometric properties, including the Outcome Probability Bias Questionnaire (Uren et al., 2004) and the Liebowitz Social Anxiety Scale - Self Report Version (Liebowitz, 1987). For social 3rd person images, one person in the photo was photoshopped so that their gender, age, and race could not easily be perceived. Participants were instructed to imagine themselves as the photoshopped person to elicit a 3rd person perspective.

All images were presented in a randomized order. The task consisted of two blocks (automatic and controlled) that differed in terms of stimulus presentation rates. Blocks were counterbalanced to assess for possible order effects: participants with an odd identification number completed the automatic block first and participants with an even identification number completed the controlled block first. E-prime was used to administer the task.

2.1.1 Controlled Block

In the controlled block, participants received the following instructions:

> For this part of the task, you are going to see a series of images showing situations you might encounter in your daily life. In some of these images, you will see a blurry outline of a person that represents you (3rd person perspective). In others, imagine you are seeing what is shown in the picture in the real world (1st person perspective).

Participants then saw three example images: a social 3rd image, a social 1st image, and a control image. Images were accompanied by instructions to "imagine you are the blurry figure" for the 3rd person image and to "imagine you are seeing what is shown in the picture in the real world" for the social 1st and control images. The following instruction screens stated:

> You may take as much time as you like to imagine you will soon experience what is shown in the picture. Then, when you are ready, press the spacebar. For each picture, you will use the slide bar to rate the likelihood that you are about to be embarrassed, humiliated, or negatively judged from 0 – 100% likely. Move the mouse left and right to adjust the slider. Left click the mouse to submit your response. A rating of 100% (all the way to the right side of the scale) indicates that you feel you will DEFINITELY be embarrassed, humiliated, or negatively judged. A rating of 0% (all the way to the left side of the scale) indicates that you feel you DEFINITELY WILL NOT be embarrassed, humiliated, or negatively judged. You may give any rating between 0-100%.

Participants completed six practice trials, two of each stimulus type, before completing the task. During the task, participants viewed each image for as long as they liked. After viewing the image, participants pressed the space bar to indicate they successfully imagined they were about to encounter the pictured situation. On the next screen, participants rated the likelihood of being embarrassed, humiliated, or negatively judged from 0-100% using a visual analogue scale. The slider always started in the middle of the visual analogue scale, at 50%. The controlled block sequence is presented in Figure 2.1.

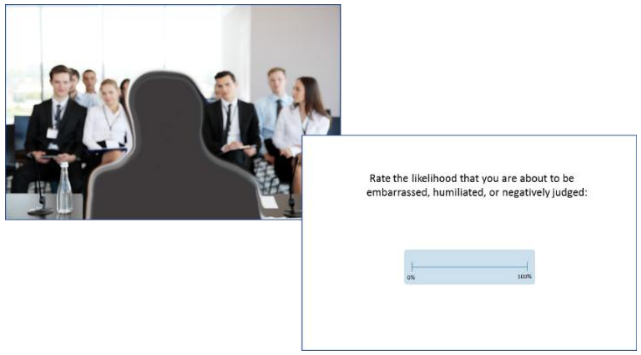


Figure 2.1 Sequence for the Controlled Block of the Outcome Probability Bias Task Using a Social 3rd Image.

2.1.2 Automatic Block

In the automatic block, images were presented rapidly (67ms each). Given the parameters of the equipment used for task administration, this was the shortest possible presentation rate that has been shown to be long enough to influence responding (Ionescu, 2016). Participants received the following instructions:

> For this part of the task, you are going to be rating some images of situations you might encounter in your daily life. First, you will see a cross in the middle of the screen. Focus your attention on the cross. Then there will be several brief flashes of images. Do your best to pay attention to the images, even if they go by too quickly for you to see them. You will use a slide bar to rate the likelihood that you are about to be embarrassed, humiliated, or negatively judged from 0 – 100% likely (if you were about to encounter what you see in the images). Move the mouse left and right to adjust the slider. Left click the mouse to submit your response. Give your rating as quickly as possible. A rating of 100% (all the way to the right side of the scale) indicates that you feel you will

DEFINITELY be embarrassed, humiliated, or negatively judged. A rating of 0% (all the way to the left side of the scale) indicates that you feel you DEFINITELY WILL NOT be embarrassed, humiliated, or negatively judged. You may give any rating between 0-100%. After you rate the images you will then be asked to identify a description of an image that was just shown, and how confident you are in your choice

Participants completed two practice trials, one with control images and the other with social images. During the task, participants were presented with a fixation cross for 2,500 milliseconds. Then, 3 (main study) or 5 (pilot study) images from the same stimulus type were presented for 67ms each. After the last image, participants rated the likelihood of being embarrassed, humiliated, or negatively judged from 0-100% likely on a visual analogue scale. The slider always started in the middle of the visual analogue scale, at 50%. Participants subsequently completed a multiple-choice comprehension question asking them to identify the correct description of an image that was just shown from three options. Next, participants rated the degree of confidence they had in their choice when responding to the comprehension question on a four-point Likert-type scale. Possible responses included “not at all confident,” “somewhat confident,” “mostly confident,” and “completely confident.” The automatic block sequence is presented in Figure 2.2.

Figure 2.2 Sequence for the Automatic Block of the Outcome Probability Bias Task in the Main Study.

2.2 Participants

Participants were recruited through the Georgia State University undergraduate psychology research participant pool and received class credit for their participation. To be eligible for the pilot study, participants were adults who were fluent in English. For the main study, participants were also required to be able to wear a virtual reality headset and could not have a history of seizures.

2.3 Measures

2.3.1 ***Virtual Reality Behavioral Avoidance Test (BAT)***

The BAT is based on a commonly used standardized speech assessment protocol (Beidel et al., 1989) and involves the delivery of an impromptu speech to a small group of confederates who are instructed to maintain a neutral facial expression. Participants are instructed to speak for ten minutes but are given the option to discontinue the speech at any point. Speech topics (e.g. healthcare, abortion) are challenging or controversial in order to elicit high anxiety. The BAT is commonly used as a behavioral measure of social anxiety, particularly in social anxiety treatment studies (Coles & Heimberg 2000; Heimberg et al., 1998). Speech length has been shown to be a reliable measure of "escape" or behavioral avoidance (Beidel et al., 1989).

In the present study, this measure was used to test the criterion validity of the outcome probability bias task. A virtual audience was used rather than confederates. Research shows that virtual reality exposure therapy significantly reduces public speaking fears from pre- to post-treatment among individuals with social anxiety disorder (Anderson et al., 2005), and treatment using virtual reality is just as effective as group exposure therapy (Anderson et al., 2013). Application of virtual reality for a BAT has been shown to be feasible for assessment of spider phobia (Mühlberger et al., 2008) but has not been assessed for public speaking anxiety.

A Shinecon virtual reality headset was used to administer the virtual reality BAT. An iPhone was placed in the headset, and the VirtualSpeech app was used for the virtual environment. Participants gave their speech in the “meeting room,” which is composed of 11 avatars arranged around a conference table. All the avatars were dressed professionally and maintained neutral facial expressions. They sometimes nodded, drank coffee, and took notes. Two avatars were standing, one on each side of the conference table. One of the standing avatars never made eye contact and was seen using his phone throughout the speech. Avatars did not speak. A laptop was present on the table in front of the viewer and acted as a stopwatch that began as soon as the participant entered the virtual environment.

2.3.2 Depression, Anxiety, Stress Scale (DASS; Lovibond & Lovibond, 1995)

The DASS is a 21-item measure of symptoms of depression, anxiety, and stress. This measure was used to assess discriminant validity of the outcome probability bias task. Participants rate how frequently they experienced symptoms over the past week from 0 (did not apply to me at all-never) to 3 (applied to me very much, or most of the time – almost always). Scores range from 0 – 63; higher scores reflect higher levels of symptomology. Relative to many other commonly used measures of anxiety and depression, the DASS has been shown to distinguish between symptoms of anxiety and depression with greater reliability (Lovibond & Lovibond, 1995). The DASS has demonstrated good psychometric properties in nonclinical samples, with internal consistency and concurrent validity scores ranging from acceptable to excellent (Antony et al., 1998; Henry & Crawford, 2005; Sinclair et al., 2012). Internal consistency in the present study was good for the depression ($\alpha = 0.87$), anxiety ($\alpha = 0.81$), and stress subscales ($\alpha = 0.80$).

Liebowitz Social Anxiety Scale– Self Report Version (LSAS; Liebowitz, 1987). The LSAS is a 48-item paper-and-pencil self-report instrument that instructs participants to rate the level of fear and avoidance they experience across a variety of social interactions and performance situations. This measure was used to assess the construct and criterion validity of the outcome probability bias task. Participants are instructed to rate their fear and avoidance using a Likert-type scale ranging from 0 (none/never) to 3 (severe/usually). Scores range from 0 – 144. Higher scores indicate greater anxiety and avoidance. The LSAS has demonstrated high internal consistency and strong convergent and discriminant validity among a clinical sample of socially anxious individuals ($\alpha = .95$) as well as among individuals without an Axis I disorder ($\alpha = .94$). A total score of 60 or higher has been shown to be associated with clinical levels of symptoms of social anxiety disorder, generalized subtype (Mennin et al., 2002). Internal consistency in the present study was excellent ($\alpha = 0.96$).

2.3.3 Outcome Cost Questionnaire (OCQ; Uren et al., 2004)

The OCQ is a 12-item paper-and-pencil questionnaire of people's beliefs about the negative consequences of a socially threatening event. This measure was used to assess convergent validity of the outcome probability bias task in the present study. The OCQ asks participants to rate "how bad or distressing the following outcomes would be for you if they were to occur" on a scale of 0 ("not at all distressing") to 8 ("extremely distressing"). Scores range from 0 – 96; higher scores reflect higher outcome cost bias. Internal consistency for the OCQ was excellent in the original scale development paper ($\alpha = 0.92$–0.94; Uren et al. 2004) and in the present study ($\alpha = 0.93$).

2.3.4 Outcome Probability Questionnaire (OPQ; Uren et al., 2004)

The OPQ is a 12-item paper-and-pencil questionnaire that assesses an individual's estimate of the probability that negative socially threatening events will occur. This measure was used to assess convergent validity of the outcome probability bias task in the present study. Items are scored on a nine-point Likert-type scale, with higher scores indicating higher outcome probability bias. Scores range from 0 – 96. Internal consistency for the measure has been found to range from good to excellent (α = .89 - .90; Uren et al., 2004). Internal consistency in the present study was excellent (α = 0.91).

2.3.5 Subtle Avoidance Frequency Examination (SAFE; Cuming et al., 2009)

The SAFE is a 19-item measure of safety behaviors among people with social anxiety. Safety behaviors are a subtle form of avoidance that people with anxiety often engage in during exposure to feared stimuli, such as avoiding eye-contact with audience members while giving a speech (Cuming et al., 2009). This measure was used to assess criterion validity of the outcome probability bias task in the present study. The SAFE instructs participants to rate how frequently they engaged in a number of safety behaviors during their speech (e.g. "blank out or switch off mentally") on a five-point Likert-Type scale from "never" to "always." Scores range from 19 – 95 with higher scores reflecting engagement in more safety behaviors. The SAFE has demonstrated excellent psychometric properties in clinical and non-clinical samples (Cuming et al., 2009) and good internal consistency in the present study (α = 0.84).

2.3.6 Simulator Sickness Questionnaire (SSQ; Kennedy et al., 1993)

The SSQ is a 16-item measure of cybersickness. Cybersickness refers to common symptoms that may result from engaging with a virtual reality environment such as dizziness or blurred vision. This measure was used to assess the feasibility of the virtual reality BAT in the present study. Participants rate the severity of symptoms on a four-point Likert-type scale from none to severe. Scores range from 16 – 64 with higher scores reflecting more symptoms of cybersickness. Internal consistency in the present study was good (α = 0.84).

2.3.7 State-Trait Anxiety Inventory-State Form (STAI; Spielberger et al., 1970)

The STAI is a 20-item measure that assesses in-the-moment symptoms of general anxiety. This measure was used to assess the feasibility of the virtual reality BAT in the present study. Participants are instructed to rate their current anxiety using a four-point Likert-type scale, (1= not at all, 2= somewhat, 3= moderately so, and 4= very much so). Scores range from 20 – 80 with higher scores reflecting higher state anxiety. The measure has demonstrated good psychometric properties, with internal consistency coefficients ranging from 0.86 to 0.95 (Spielberger & Gorsuch, 1983). Internal consistency in the present study was excellent (α = 0.90 - 0.92).

2.3.8 Subjective Units of Distress Scale (SUDS; Wolpe, 1973)

The SUDS is a self-report face-valid scale that is used to assess current distress or anxiety. A SUDS of 0 represents the absence of any distress or anxiety, 5 represents moderate distress or anxiety, and 10 represents the most distress or anxiety imaginable. SUDS ratings are commonly used during exposure therapy as well as during assessment (e. g., during a behavioral avoidance test; Kim et al., 2008; Wolpe & Lazarus, 1966). In the present study, participants stated their SUDS at two-minute intervals throughout the BAT. SUDS were used to assess criterion validity of the outcome probability bias task.

2.3.9 Temple Presence Inventory (TPI; Lombard et al., 2009)

Participants' experience of fear in response to the virtual environment has been shown to be related to presence, or the ability to feel connected to and engaged with virtual stimuli (Lee, 2004; Schubert et al., 2001; Witmer & Singer, 1998). The TPI is a 42-item multidimensional measure that assesses spatial presence, social presence, engagement, social richness, social realism, and perceptual realism. This measure was used to assess the feasibility of the virtual reality BAT in the present study. It has demonstrated good reliability, validity, and applicability across a wide range of media (Lombard et al., 2009). Participants rate the degree to which an item applied to their virtual experience from 1 (not at all) to 7 (very much). A subset of 20 items relevant to the present study were used. As such, scores ranged from 20 – 140. Internal consistency in the present study was excellent (α = 0.92).

2.3.10 Word Sentence Association Paradigm (WSAP; Beard & Amir, 2008)

The WSAP is a computer task designed to measure interpretation bias in social anxiety. This measure was used to assess discriminant validity of the outcome probability bias task. Participants first see a fixation cross for 500ms, followed by a target word for 500ms. An ambiguous sentence appears (e.g., "people laugh after something you said") and participants are instructed to press the spacebar as soon as they have read the sentence. Participants then see the prompt, "are they related?" Participants press "1" on the keypad if the word and sentence appear to be related, and "3" if they do not appear to be related. The task consists of 110 trials, 76 of which include social themes and 34 of which are non-social filler sentences (e.g., "a loud noise is heard at night"). The target words that quickly flash on the screen prior to the ambiguous sentence can be interpreted as threatening (e.g., "embarrassing") or benign (e.g., "funny"). Interpretation bias is calculated by self-report (the number of threat trials that received a response of "yes" to the relatedness prompt) or reaction times (mean reaction time to endorse relatedness of a threatening target word minus mean reaction time to reject relatedness of a threatening target word). Based on previous research, the WSAP has been shown to correlate with symptoms of social anxiety and validly discriminated between socially anxious individuals and non-anxious controls (Beard & Amir, 2009). A review of 41 studies that used the WSAP indicated the task has good reliability and validity (Gonsalves et al., 2019).

3 PILOT STUDY

3.1 Participant Characteristics

Participants' (*n* = 60) mean age was 18.95 (range 18 - 22). The majority of the sample identified as female (*n* = 46, 77%). Thirteen identified as male (22%) and one identified as gender-fluid. Regarding sexual orientation, 45 participants identified as heterosexual (75%), eight identified as bisexual (13%), four as questioning (7%), two as gay (3%) and one as lesbian (2%). Twenty-four participants identified as African American/Black (40%), 14 as White (23%), 11 as Asian American/Asian (18%), seven as Hispanic/Latinx (12%), three as multi-racial (5%), and one self-identified as Pacific Islander and White. Most reported the United States as their country of origin (*n* = 42, 70%). The remaining participants' countries of origin included Canada, France, Pakistan, Nigeria, Antigua, Bangladesh, India, Puerto Rico, the Bahamas, Vietnam, and South Korea. Most were single (*n* = 47, 78%), 12 reported being in a serious dating or committed relationship (20%), and one was in a civil union, domestic partnership, or equivalent. Many of the participants reported being Christian (*n* = 24, 40%), eight were agnostic (13%), eight stated no preference (13%), six were Muslim (10%), six stated "other," including spiritual, general belief in God, Christian Buddhist, and Regla de Ocha, four Catholic (7%), two Atheist (3%), one Jewish, and one Hindu.

3.2 Pilot Study Procedure

This project was approved by the Georgia State University Institutional Review Board (H20007). Following completion of the informed consent, participants completed the pre-test

STAI, the outcome probability bias task, and the post-test STAI. Subsequently, participants completed a battery of questionnaires, including the LSAS, OPQ, OCQ, and the DASS. Participants also filled out a demographics questionnaire and a questionnaire designed to elicit participants' feedback about the outcome probability task (included in Appendix B.9). This questionnaire included questions about the images (e.g. if participants were able to tell the race or gender of the blurry figures; which images, if any, induced the most anxiety), any strategies they may have used while completing the outcome probability bias task, and any general feedback they may have had.

3.3 Pilot Study Results

Power analyses using g*power indicated a minimum sample size of 24 was needed to detect a small effect for pilot study hypotheses. A conservative estimate was chosen given the outcome probability bias task is a new measure. Due to the iterative nature of the pilot process, data were collected from 60 participants. Changes made during completion of the pilot included adding confidence ratings to the automatic block and editing task instructions for clarity.

Data from the outcome probability bias task were missing for two participants in the controlled block and eight participants in the automatic block due to programming error. Significant skew was detected in post-STAI scores and mean outcome probability bias task ratings. Notably, the *F*-statistic is considered to be robust with regard to violations of normality and homogeneity when group sizes are equal (Donaldson, 1968). In addition, linear regression analyses do not assume normal distribution of predictor variables (Field, 2013). As such, no corrections were made to present analyses due to skew.

3.3.1 Descriptive Statistics

Means and standard deviations are reported in Table 3.1 (overall) and Table 3.2 (separated by social anxiety symptom scores above and below the clinical cutoff of 60; Mennin et al., 2002).

Table 3.1 Means and Standard Deviations for Measures Used in the Pilot Study

Measure	M (SD)
Pre-STAI	35.22 (9.45)
Post-STAI	39.57 (11.45)
OPQ	56.10 (23.01)
OCQ	64.22 (20.71)
LSAS	67.28 (28.40)
Social 1st (Controlled)	47.78 (16.64)
Social 3rd (Controlled)	49.27 (19.37)
Control (Controlled)	9.59 (10.67)
Social 1st (Automatic)	59.90 (29.30)
Social 3rd (Automatic)	50.74 (26.42)
Control (Automatic)	8.58 (12.84)

Note. STAI = State-Trait Anxiety Inventory, OPQ = Outcome Probability Questionnaire, OCQ = Outcome Cost Questionnaire, LSAS = Liebowitz Social Anxiety Scale, Social 1st, Social 3rd, and Control measures all refer to mean outcome probability bias task ratings from 0 – 100%.

Table 3.2 Means and Standard Deviations for Measures Used in the Pilot Study Separated by Social Anxiety Symptom Clinical Cutoff

Measure	Below LSAS Cutoff M (SD) *n* = 27	Above LSAS Cutoff M (SD) *n* = 31
Pre-STAI	31.04 (7.31)	38.88 (9.66)
Post-STAI	34.93 (9.52)	43.63 (11.58)
OPQ	39.00 (18.04)	71.06 (15.13)
OCQ	51.32 (19.92)	75.50 (13.76)
LSAS	45.14 (20.05)	86.66 (18.84)
Social 1st (Controlled)	38.30 (17.72)	56.03 (10.15)
Social 3rd (Controlled)	38.15 (20.03)	58.96 (12.52)
Control (Controlled)	8.33 (9.00)	10.69 (11.99)
Social 1st (Automatic)	49.49 (29.65)	69.23 (26.01)
Social 3rd (Automatic)	38.68 (22.33)	61.25 (25.47)
Control (Automatic)	7.27 (10.21)	9.71 (14.84)

Note. STAI = State-Trait Anxiety Inventory, OPQ = Outcome Probability Questionnaire, OCQ = Outcome Cost Questionnaire, LSAS = Liebowitz Social Anxiety Scale, Social 1st, Social 3rd, and Control measures all refer to mean outcome probability bias task ratings from 0 – 100%.

The distribution of means and standard deviations was consistent with expectations that outcome probability bias task ratings would be higher for social images compared to control images and for people who reported symptoms of social anxiety above the clinical cutoff relative to those who reported symptoms below the clinical cutoff.

Zero-order correlations are reported in Table 3.3. Zero-order correlations were also consistent with expectations. Social images from the outcome probability bias task were positively and significantly correlated with paper-and-pencil measures of symptoms of social anxiety (LSAS), outcome probability bias (OPQ), and outcome cost bias (OCQ). Correlations were strong for the controlled block and moderate for the automatic block.

Table 3.3 Zero-order Correlations Among Blocks of the Outcome Probability Task and Paper-and-pencil Measures of State Anxiety, Outcome Probability Bias, Outcome Cost Bias, and Symptoms of Social Anxiety Included in the Pilot Study

	Pre-STAI	Post-STAI	OPQ	OCQ	LSAS	Social 1st	Social 3rd	Control	Social 1st (Automatic)	Social 3rd (Automatic)
Pre-STAI	1									
Post-STAI	.74*	1								
OPQ	.54*	.56*	1							
OCQ	.45*	.35*	.70*	1						
LSAS	.56*	.52*	.72*	.68*	1					
Social 1st (Controlled)	.48*	.50*	.62*	.50*	.60*	1				
Social 3rd (Controlled)	.51*	.50*	.61*	.51*	.68*	.90*	1			
Control (Controlled)	.06	.34*	.20	.14	.24	.27*	.19	1		
Social 1st (Automatic)	.14	.11	.44*	.33*	.39*	.53*	.56*	-.03	1	
Social 3rd (Automatic)	.17	.30*	.45*	.33*	.44*	.66*	.65*	.14	.81*	1
Control (Automatic)	.01	.17	.24	.14	.14	.15	.15	.40*	-.01	.22

Note. STAI = State-Trait Anxiety Inventory, OPQ = Outcome Probability Questionnaire, OCQ = Outcome Cost Questionnaire, LSAS = Liebowitz Social Anxiety Scale, Social 1st, Social 3rd, and Control measures all refer to mean outcome probability bias task ratings from 0 – 100%.

3.3.2 Feasibility of Stimuli

Feasibility of task stimuli was evaluated by assessing internal consistency among the images, analysis of hypotheses 1 – 4, and anecdotal feedback from participants.

3.3.2.1 Internal Consistency

Internal consistency was evaluated using Cronbach's alpha. Internal consistency was excellent for all stimulus types in the controlled block ($\alpha = 0.94$, 0.96, and 0.91 for social 1st, social 3rd, and control images, respectively). Internal consistency was good for social 1st images ($\alpha = 0.89$), acceptable for social 3rd images ($\alpha = 0.79$), and poor for control images ($\alpha = 0.58$) in the automatic block. Alpha-if-item-deleted scores indicated that removal of the control trial depicting different objects (e.g. coffee cup, padlock, etc.) would improve internal consistency to the questionable range ($\alpha = 0.67$). All images demonstrated a correlation greater than $r = 0.30$ with the mean score for a given stimulus type in both blocks, with one exception ($r = 0.29$), and were therefore retained.

3.3.2.2 Quantitative Analyses

Hypothesis 1: A 2x2 repeated measures ANOVA with a within-subjects factor of time (pre- and post-STAI) and a between-subjects factor of symptoms of social anxiety (above or below the clinical cutoff on the LSAS) was used to evaluate the hypothesis that state anxiety would increase following completion of the outcome probability bias task for all participants. The analysis revealed significant main effects of time (F (1, 58) = 18.26, $p < 0.01$, *partial* $\eta^2 = 0.24$) and symptoms (F (1, 58) = 12.92, $p < 0.01$, *partial* $\eta^2 = 0.18$). State anxiety significantly increased from pre- to post-outcome probability bias task. State anxiety scores were significantly higher for those with social anxiety symptom scores above the clinical cutoff relative to those with scores below the clinical cutoff. The interaction was not significant ($p < 0.05$).

Hypothesis 2: A series of one-way ANOVA analyses were conducted to evaluate the hypothesis that social images would elicit significantly higher outcome probability bias task ratings relative to control images. In both analyses, the assumption of sphericity was violated so Greenhouse-Geisser corrected values are reported. *A priori* repeated contrasts were used to evaluate differences in outcome probability bias task ratings among the three stimulus types. These contrasts compared social 1st images to social 3rd images and social 3rd images to control images. Social 1st and control images were not compared using this method.

The model was significant for the controlled block, indicating a significant difference in mean outcome probability bias task ratings among the three stimulus types: F (2, 72.58) = 225.25, p < 0.01, *partial* η^2 = 0.80. Repeated contrasts indicated that ratings in response to social 1st and social 3rd images did not differ significantly (p < 0.05). Outcome probability bias task ratings in response to social 3rd images were significantly higher relative to ratings in response to control images (F (1, 57) = 222.55, p < 0.01, *partial* $\eta 2$ = 0.80). Results are depicted in Figure 3.1.

The model was also significant for the automatic block, reflecting a significant difference in mean outcome probability bias task ratings among the three stimulus types: F (2, 82.36) = 127.21, p < 0.01, *partial* η^2 = 0.69. Repeated contrasts revealed responses to social 1st images were significantly greater than responses to social 3rd images (F (1,57) = 16.12, p < 0.01, *partial* η^2 = 0.22. Responses to social 3rd images were significantly greater than responses to control images (F (1,57) = 144.31, p < 0.01, *partial* η^2 = 0.72 (see Figure 3.2).

Results from both of these analyses were consistent with the hypothesis that outcome probability bias task ratings would be higher in response to social images relative to control images.

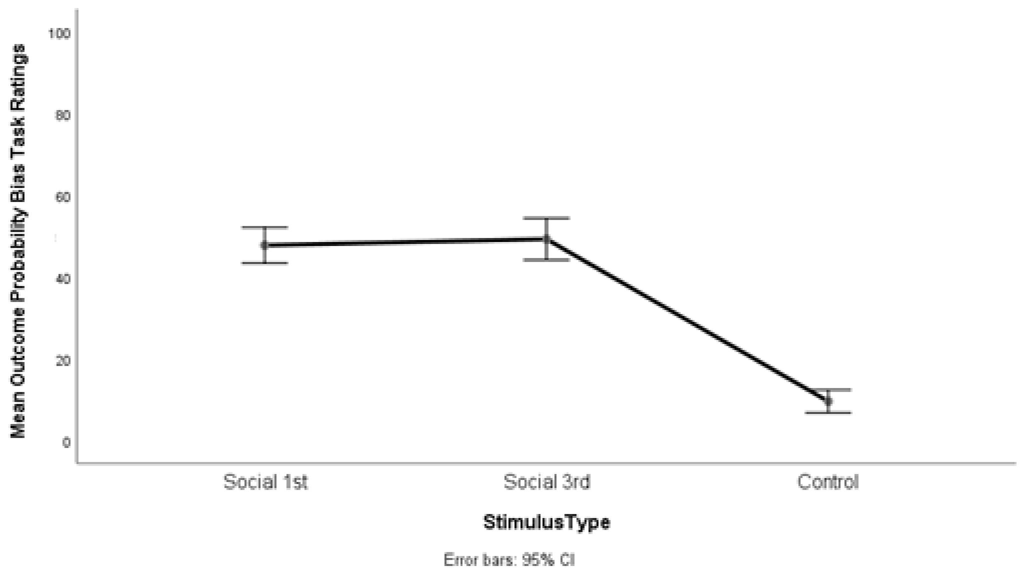


Figure 3.1 Mean Outcome Probability Bias Task Ratings for Social 1st, Social 3rd, and Control Images in the Controlled Block During the Pilot Study

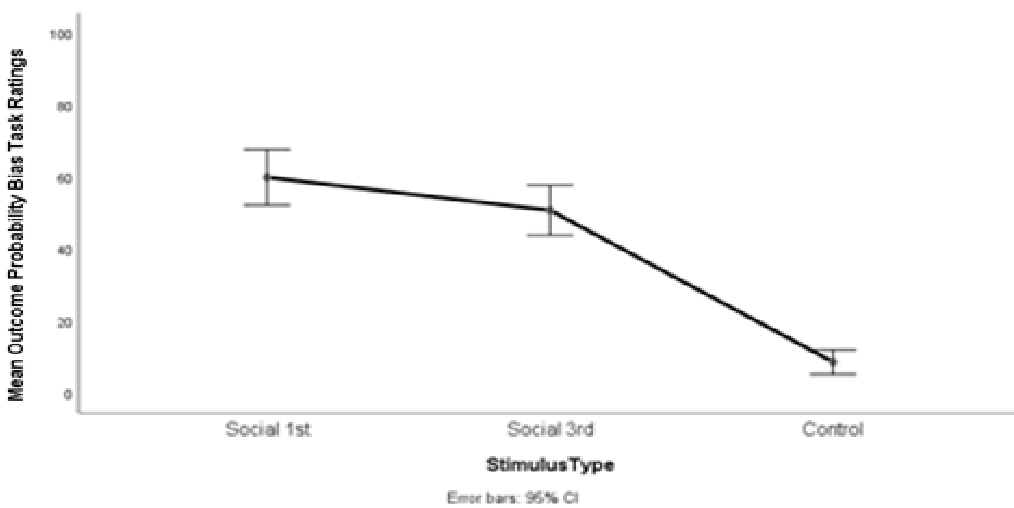


Figure 3.2 Mean Outcome Probability Bias Task Ratings for Social 1st, Social 3rd, and Control Images in the Automatic Block During the Pilot Study

Hypothesis 3: A one-way ANOVA analysis with a between-subjects factor of symptoms of social anxiety (above or below the clinical cutoff on the LSAS) and outcome probability bias task stimuli entered as the dependent variables was used to test the hypothesis that ratings for social images would be higher for those reporting symptoms of social anxiety above the clinical cutoff (60; Mennin et al., 2012). Consistent with the hypothesis, outcome probability bias task ratings in response to all social images were significantly higher among those reporting higher levels of social anxiety relative to those reporting lower levels of social anxiety (Social 1st (controlled): F (1, 57) = 22.56, $p < 0.01$; Social 3rd (controlled): F (1, 57) = 23.12, $p < 0.01$; Social 1st (automatic): F (1, 57) = 7.52, $p < 0.01$; Social 3rd (automatic): F (1, 57) = 12.69, $p < 0.01$). Ratings did not significantly differ by group in response to control images (all p values > 0.05).

Hypothesis 4: A series of regression analyses were used to test the hypothesis that social images would significantly predict responses to paper-and-pencil measures of outcome probability bias and symptoms of social anxiety, but control images would not. Paper-and-pencil measures of outcome probability bias (OPQ) and symptoms of social anxiety (LSAS) were entered as dependent variables and social and control images from the controlled and automatic blocks of the outcome probability bias task served as the predictor variables. Social images were collapsed across stimulus type (social 1st and social 3rd) due to the high degree of correlation between them, which could contribute to problems with multicollinearity (Field, 2013). All regression models were significant (all p values < 0.01). The controlled and automatic blocks of the outcome probability bias task accounted for 40% and 25% of the variability in the OPQ and 45% and 20% of the variability in the LSAS, respectively. Across analyses, social images were identified as significant predictors and control images were nonsignificant (see Table 3.4).

Table 3.4 Results from Regression Analyses of the Outcome Probability Bias Task with Paper-and-pencil Measures of Outcome Probability Bias and Symptoms of Social Anxiety in the Pilot Study

DV	Predictor	Model	β	P	95% CI
OPQ	Controlled Block	F (2, 57) = 18.36, p < 0.01			
	Social Images		0.79	<0.01	0.52 – 1.07
	Control Images		0.06	0.60	-0.34 – 0.58
	Automatic Block	F (2,57) = 9.28, p < 0.01			
	Social Images		0.38	<0.01	0.18 – 0.58
	Control Images		0.34	0.11	-0.08 – 0.75
LSAS	Controlled Block	F (2, 57) = 14.44, p < 0.01			
	Social Images		0.98	<0.01	0.30 – 1.65
	Control Images		0.32	0.26	-0.25 – 0.89
	Automatic Block	F (2, 57) = 6.64, p < 0.01			
	Social Images		0.42	<0.01	0.19 – 0.72
	Control Images		0.21	0.44	-0.33 – 0.76

3.3.2.3 Feedback Survey

Fifty-four participants (90%) indicated they had little or no difficulty imagining themselves in the social situations depicted in the outcome probability bias task images. Thirty-six participants (60%) gave some indication that they could determine the race, age, or gender of the blurry figures. Generally, participants indicated they could guess the gender in some of the images. For example, one participant stated, “sometimes I could tell the gender of the blurred figures based on the shape of the silhouette.” Another participant indicated, “just gender, occasionally,” though others indicated they could also sometimes tell the age or race of the figure. Many participants indicated they could not tell, with one participant stating, “no, because I was only imagining it being me.”

A total of 41 participants (68%) indicated that some of the images made them feel anxious. Feedback about which images induced the most anxiety supported the face validity of the task.

Among participants who indicated the images made them feel anxious, many participants stated the images with public speaking (n = 20) and interviews (n = 11) were particularly anxiety provoking. Images with disapproving facial expressions (mad, judgmental, bored), direct eye contact, and close-up images of faces were also frequently endorsed (n = 15). Other descriptions of anxiety provoking images were typically consistent with items on the paper-and-pencil measure of symptoms of social anxiety (LSAS) including talking to someone in authority or parties. One response also emphasized the race and age of the people in the image: "meeting with powerful white men who were older than me."

When asked "what were you afraid might happen in response to the images that made you feel anxious?" participants indicated that their "mind would go blank," that they would be negatively judged or mocked, they would be embarrassed, they would fail in some way (e.g. not achieve their goals), be punished, or they would be ostracized. Some also indicated concern about their emotional experience as a result (feeling anxiety, feeling jittery, or feeling regret). Images that tended to be least likely to provoke anxiety included those of plants, buildings, and objects. Some also indicated that images of parties or showing people having fun and smiling did not induce anxiety.

Most participants indicated they didn't use any strategies while completing the outcome probability bias task. Among those who did endorse a strategy, one participant reported they would "zone out and focus on the image until it comes to life." Another indicated they would "try to really think about what was about to happen."

When asked for general feedback, many reported enjoying the task, stating they were engaged throughout or that the images were relatable. For example, one participant stated, "the images were well placed and not hard to picture yourself in." Another said, "[the images] were all easy to focus on and understand the situation pictured." Others described how the images elicited anxiety, including some who stated they realized the extent of their social anxiety as a result of participating in the task. For example, one participant noted that the images "made me physically feel very anxious" and another shared that "they just showed me that I might actually have anxiety." Others noted that the images didn't have much of an emotional impact on them: "overall none of the images caused me stress or made me feel uncomfortable." Finally, some participants expressed confusion surrounding inclusion of the control items, stating they "didn't understand why inanimate objects make you feel embarrassed or judged."

3.3.3 Feasibility of Stimulus Duration in Automatic Block

Assessment of the feasibility of the stimulus duration in the automatic block involved evaluating the degree to which responses may have been influenced by automatic or controlled processing. This was assessed by examination of accuracy and confidence ratings and anecdotal feedback.

3.3.3.1 Accuracy and Confidence Ratings

In the automatic task, participants completed comprehension questions about the images they just saw. They also rated how confident they were in their response to the comprehension question. Participants' responses to comprehension items were highly accurate, with a mean accuracy rate of 8.11 out of nine questions (n = 50). High accuracy (above the level of chance) indicates there may have been opportunity for controlled processing to take place when participants rated the images in the automatic task. The average confidence rating was 2.27 out of 4 (n = 18; confidence ratings were added towards the end of data collection in the pilot study). Lower ratings reflected less confidence.

3.3.3.2 Anecdotal Feedback

When asked if the participants were able to tell what the images were showing during the automatic block, 23 (38%) indicated they could discern nearly all of the images; the remaining 37 participants (62%) indicated they could tell what was shown some of the time. Some participants indicated they could remember at least one of the images (often the first or the last image) and used this to influence their responses to the comprehension questions: “[my ratings were] based off the freshest picture in my memory.” Another participant said “if I didn't know, I just focused on finding one picture specifically and answering the question based off of that singular picture.” Because a substantial proportion of participants stated they could tell what the images were showing for nearly all images, controlled processing could have impacted responding during the automatic block, particularly if participants took time to consider their response before submitting their outcome probability rating.

3.3.4 Role of the Availability Heuristic in Outcome Probability Bias

I hypothesized that outcome probability bias task ratings would be higher for participants who completed the controlled block first due to increased familiarity with the images. Participants who completed the controlled block first viewed and rated the same images when they completed the automatic block (and were therefore familiar with the images). Participants who completed the automatic block first were seeing the images for the first time (and were therefore not familiar with the images). Familiarity exacerbates the availability heuristic at an automatic level of processing (Menon & Raghubir, 2003; Kahneman, 2011). As such, higher outcome probability bias task ratings during the automatic block among participants who were familiar with the images (participants who completed the controlled block first) relative to those who were unfamiliar with the images (participants who completed the automatic block first) could support Butler and Mathews' (1983) hypothesis that outcome probability bias results from the availability heuristic.

A one-way ANOVA was conducted to evaluate the effect of block order on mean outcome probability bias task ratings in the automatic block. Consistent with the hypothesis, block order had a significant effect on mean outcome probability bias task ratings in response to social images (F (1, 57) = 4.71, p = 0.03), such that mean ratings were significantly higher if participants completed the controlled block first (and were familiar with the images; M = 62.64) compared to participants who completed the automatic block first (and were unfamiliar with the images; M = 48.00).

3.4 Pilot Discussion

The pattern of results indicated that the images included in the task were appropriate. Internal consistency ranged from acceptable to high for all social images (α = 0.79 - 0.96), hypotheses 1 – 4 were supported in quantitative analyses, and anecdotal feedback indicated that the task images were evocative and relatable. The controlled block was retained for the main study with no alterations, but, as explained below, changes were made to the automatic block.

To further limit the degree to which controlled processing could influence responding during the automatic block, the number of images included in the automatic trials was reduced from five to three for the main study. Accuracy ratings were high (well above the level of chance) in response to comprehension questions following automatic trials during the pilot, indicating there was potential for participants to engage in controlled processing. Inclusion of five images (all depicting the same social situation) of 67ms each (335ms total) likely allowed sufficient time for controlled processing to take place. Reducing the number of images within each trial for the main study shortened the total amount of time participants were exposed to the stimuli, thus limiting the potential for controlled processing to occur. In addition to reducing the number of images, the automatic block task instructions were amended to emphasize that participants should give their outcome probability ratings as quickly as possible to further minimize opportunity for controlled processing.

I also removed social 3rd images from the automatic block. Outcome probability bias task ratings were significantly higher for social 1st images relative to social 3rd images, potentially because social 3rd images are not conducive to automatic processing. Controlled processing is required to insert oneself into a blurry figure. Given the short duration of the stimulus presentation (67ms), participants may not have noticed the presence of the blurry figure in the image.

Social 3rd images were replaced with novel images that were not included in the controlled block to clarify the nature of the order effect detected in the automatic block ratings. Participants who completed the controlled block first had higher outcome probability bias task ratings in response to automatic trials. This order effect could be due to greater familiarity with the images because participants who completed the controlled block first would have previously seen the same images once before and were therefore more familiar with the images than participants who completed the automatic block first. Research indicates that familiarity can promote engagement with the availability heuristic at an automatic level of processing (Kahneman, 2011). Given that outcome probability bias has been hypothesized to result from the availability heuristic (Butler & Mathews, 1983), familiarity should also promote engagement of outcome probability bias. As such, higher outcome probability bias task ratings among participants who completed the controlled block first is consistent with Butler and Mathews' (1983) claim. An alternative interpretation for order effects is that outcome probability bias task ratings were higher among those who completed the controlled block first because of demand bias. Participants were aware the study was about social anxiety and may have had sufficient exposure to the task to infer that social items should be rated higher and responded accordingly. In order to evaluate whether this order effect was due to familiarity or demand bias, new social 1st images were added to the automatic block in the main study. These new images depicted the same social situations as the familiar images but were not included in the controlled block and would therefore be unfamiliar to all participants. Presence of an order effect for familiar, but not novel social images would rule out the interpretation that demand bias contributed to higher ratings among those who completed the controlled block first.

4 MAIN STUDY

4.1 Participant Characteristics

Participants' (*n* = 148) mean age was 20.61 (range 18-47). The majority of participants identified as women (*n* = 101, 68%). Forty-six identified as men (31%) and one identified as other and did not specify. Regarding sexual orientation, 118 participants identified as heterosexual (80%), thirteen identified as bisexual (9%), seven as questioning (5%), seven as gay (5%) and two as pansexual (1%). Seventy participants identified as African American/Black (47%), 25 as White (17%), 23 as Asian American/Asian (16%), 15 as Hispanic/Latinx (10%), 10 as multi-racial (7%), two as Native Hawaiian or Pacific Islander (1%) and two as American Indian or Alaska Native (1%). The majority reported the United States is their country of origin (*n* = 91, 62%). Most were single (*n* = 96, 65%), 45 reported being in a serious dating or committed relationship (30%), four were married (3%), one was separated, one was divorced, and one was in a civil union, domestic partnership, or equivalent. Only six of the participants were international students (4%). Many of the participants were freshman (*n* = 65, 44%). None reported any physical disabilities.

4.2 Main Study Procedure

This project was approved by the Georgia State University Institutional Review Board (H20200). This study was pre-registered through the Open Science Framework (https://osf.io/fr8zv). Results from the pilot study were used to influence the design of the outcome probability bias task. No changes were made to the controlled block. Changes to the automatic block were intended to reduce opportunities for controlled processing and included a reduction in the number of images presented in each trial (from five to three), removal of a control trial that negatively impacted internal consistency, removal of trials using social 3rd images, and inclusion of new trials containing novel images that were not also included in the controlled block. New images added to the automatic block were consistent in content with familiar images, which were also present in the controlled block (i.e. public speaking, interviews, threatening faces). In addition, comprehension question items were altered so correct responses were specific to the second image presented in the automatic trials, and the instructions were modified to specify that participants should submit their ratings on the visual analogue scale as quickly as possible.

The procedure for the main study differed from the procedure used for the pilot study. After signing the informed consent, participants completed the outcome probability bias task followed by a virtual reality behavioral avoidance test. Participants then completed a measure of state anxiety (STAI) and rated their performance while standing to maintain physiological arousal. Subsequently, participants completed a battery of questionnaires on the computer including measures of cybersickness (SSQ), presence (TPI), safety-behaviors (SAFE), outcome probability bias (OPQ), outcome cost bias (OCQ), symptoms of social anxiety (LSAS), symptoms of depression, anxiety, and stress (DASS), exploratory measures of stereotype confirmation concerns and mindfulness, a demographics questionnaire, and an exploratory measure of post-event processing. Finally, participants completed the WSAP interpretation bias computer task.

4.3 Main Study Results

Power analyses using g*power indicated a minimum sample size of 67 was needed to detect a small effect for hypotheses related to convergent, discriminant, construct, and criterion validity. A minimum sample size of 200 is recommended to conduct an exploratory factor analysis (Fabrigar, Wegener, MacCallum, & Strahan, 1999). Data collected from the main study (*n* = 148) were combined with data from the pilot study (*n* = 55; 5 participants completed both studies) when possible to ensure adequate power (outcome probability bias task controlled block, OPQ, OCQ, and LSAS), including for the exploratory factor analysis of the controlled block from the outcome probability bias task. Research shows that the sample size needed to achieve a point of stability for correlations of 0.5 and power level of 0.80 is *n* = 143 (Schönbrodt and Perugini, 2013), which further increased confidence that the study was adequately powered.

4.3.1 Missing Data Strategy

Missing data can lead to significant loss of information that could potentially influence the sensitivity of statistical tests (Gold & Bentler, 2000). Little’s (1988) MCAR test indicated that missing data were likely missing completely at random for all self-report measures included in the present analyses. Expectation maximization was used to replace missing values with predicted values. Expectation maximization has been shown to be advantageous in a Monte-Carlo examination of different approaches to estimating missing data (Gold & Bentler, 2000).

For the interpretation bias task (WSAP), reaction times shorter than 50ms or longer than 2,000ms were removed, consistent with the original task development paper (Beard & Amir, 2009). No other outliers were identified in the dataset. Assumptions of normality were evaluated using P-P plots, histograms, and scatterplots of residuals (Field, 2013). All assumptions were met with two exceptions. Positive skew and platykurtosis were detected for outcome probability bias task control images, which was not surprising because most participants tended to rate little to no likelihood of negative judgment, embarrassment, or humiliation in response to nonsocial images. In addition, platykurtosis and negative skew were detected for BAT speech length.

4.3.2 Descriptive Statistics

Overall means and standard deviations for all primary measures are reported in Table 4.1. Means and standard deviations for all measures separated by social anxiety symptom scores above and below a clinical cutoff score (60; Mennin et al., 2012) are reported in Table 4.2. Outcome probability bias task ratings were higher for social images relative to control images and among participants with social anxiety symptom scores above a clinical cutoff relative to those with scores below the clinical cutoff. The mean social anxiety symptom score on the LSAS was just below the clinical cutoff for all participants. Overall mean depression, anxiety, and stress scores on the DASS fell in the mild, moderate, and mild ranges, respectively. Symptoms of general anxiety (DASS) fell in the severe range for participants with social anxiety symptom scores above the clinical cutoff and the mild range for participants with scores below the clinical cutoff. Zero-order correlations among measures included in validity analyses are reported in Table 4.3. Correlations were consistent with expectations with two exceptions. The outcome probability bias task was not significantly related to behavioral avoidance (BAT) and only social 1st images from the controlled block significantly correlated with interpretation bias (WSAP).

Table 4.1 Means and Standard Deviations for Measures in the Main Study

Measure	M (SD)
Social 1st (Controlled)	43.61 (18.59)
Social 3rd (Controlled)	45.18 (20.98)
Control (Controlled)	7.72 (9.46)
Social (Automatic)	53.30 (26.34)
Control (Automatic)	9.53 (18.74)
OPQ	47.78 (23.06)
OCQ	56.57 (24.56)
LSAS	58.00 (27.59)
SAFE	46.62 (10.55)
DASS Stress	15.00 (8.59)
DASS Anxiety	11.26 (8.77)
DASS Depression	10.87 (8.94)
STAI	46.19 (14.23)
BAT Speech Length (s)	349.86 (233.10)
Peak SUDS	6.61 (2.63)
WSAP Threat Endorse	23.58 (6.75)
WSAP Benign Endorse	25.71 (5.85)
WSAP Threat Bias (RT)	44.75 (139.74)
WSAP Benign Bias (RT)	-77.28 (148.92)

Note. Social 1st, Social 3rd, Social, and Control measures all refer to mean outcome probability bias task ratings from 0 – 100%. OPQ = Outcome Probability Questionnaire, OCQ = Outcome Cost Questionnaire, LSAS = Liebowitz Social Anxiety Scale, SAFE = Subtle Avoidance Frequency Examination, DASS = Depression Anxiety Stress Scale, STAI = State-Trait Anxiety Inventory, BAT = Behavioral Avoidance Test, SUDS = Subjective Units of Distress, WSAP = Word Sentence Association Paradigm, RT = reaction time.

Table 4.2 Means and Standard Deviations Separated by Self-Report Social Anxiety Symptom Clinical Cutoff for Measures in the Main Study

Measure	Below LSAS Cutoff M (SD) *n* = 86	Above LSAS Cutoff M (SD) *n* = 62
Social 1st (Controlled)	34.63 (18.41)	52.01 (14.47)
Social 3rd (Controlled)	35.33 (19.84)	54.85 (17.57)
Control (Controlled)	6.07 (8.74)	8.13 (8.68)
Social (Automatic)	45.35 (27.18)	64.32 (20.73)
Control (Automatic)	5.94 (13.79)	14.52 (23.19)
OPQ	36.49 (20.52)	55.40 (20.01)
OCQ	45.36 (25.44)	65.92 (20.03)
LSAS	36.59 (15.48)	80.10 (15.28)
SAFE	44.38 (10.48)	49.72 (9.91)
DASS Stress	12.23 (7.72)	18.84 (8.30)
DASS Anxiety	8.12 (6.17)	15.61 (9.96)
DASS Depression	8.00 (7.19)	14.84 (9.64)
STAI	40.85 (11.61)	53.60 (14.29)
BAT Speech Length (s)	368.56 (228.09)	323.94 (239.31)
Peak SUDS	5.64 (2.5)	7.95 (2.14)
WSAP Threat Endorse	23.05 (6.63)	24.33 (6.89)
WSAP Benign Endorse	26.03 (5.25)	25.26 (6.62)
WSAP Threat Bias (RT)	38.31 (145.51)	54.13 (131.55)
WSAP Benign Bias (RT)	-78.78 (146.54)	-75.14 (153.48)

Note. Social 1st, Social 3rd, Social, and Control measures all refer to mean outcome probability bias task ratings from 0 – 100%. OPQ = Outcome Probability Questionnaire, OCQ = Outcome Cost Questionnaire, LSAS = Liebowitz Social Anxiety Scale, SAFE = Subtle Avoidance Frequency Examination, DASS = Depression Anxiety Stress Scale, STAI = State-Trait Anxiety Inventory, BAT = Behavioral Avoidance Test, SUDS = Subjective Units of Distress, WSAP = Word Sentence Association Paradigm, RT = reaction time.

4.3.3 Reliability: Internal Consistency

Internal consistency for the controlled block was excellent for social 1st (α = 0.95) and social 3rd images (α = 0.96), and good for control images (α = 0.89). For the automatic block, internal consistency was good for the social images (α = 0.86) and the control images (α = 0.82).

4.3.4 Convergent and Discriminant Validity

Table 4.3 Zero-Order Correlations Among Blocks of the Outcome Probability Task and Paper-and-pencil Measures of Outcome Probability Bias, Outcome Cost Bias, Social Anxiety Symptoms, Safety Behaviors, Depression, Stress, an Interpretation Bias Task, and Behavioral Avoidance for the Main Study

	Social 1st (Ctrl)	Social 3rd (Ctrl)	Control (Ctrl)	Social (Auto)	Control (Auto)	OPQ	OCQ	LSAS	SAFE	DASS-S	DASS-D	WSAP
Social 1st (Ctrl)	1											
Social 3rd (Ctrl)	.89*	1										
Control (Ctrl)	.34*	.26*	1									
Social (Auto)	.63*	.60*	.20*	1								
Control (Auto)	.29*	.31*	.37*	.20*	1							
OPQ	.46*	.48*	.13	.33*	.00	1						
OCQ	.42*	.45*	.05	.29*	-.03	.66*	1					
LSAS	.58*	.62*	.20*	.41*	.24*	.59*	.56*	1				
SAFE	.41*	.43*	.11	.28*	.17*	.51*	.42*	.40*	1			
DASS Stress	.31*	.36*	-.02	.18*	.16	.50*	.43*	.51*	.40*	1		
DASS Depression	.28*	.33*	.12	.21*	.07	.51*	.43*	.42*	.30*	.67*	1	
WSAP Threat Bias (RT)	.21*	.15	.15	.12	.11	.07	-.07	.17*	.01	.04	.09	1
BAT Speech Length	-.09	-.13	-.04	.00	-.07	-.16	-.08	-.21*	-.28*	-.15	-.05	-.17*

Note. Statistically significant correlations ($p < 0.05$) are designated with an asterisk. Social 1st, Social 3rd, Social, and Control measures all refer to mean outcome probability bias task ratings from 0 – 100%. Ctrl = Controlled block, Auto = Automatic block, OPQ = Outcome Probability Questionnaire, OCQ = Outcome Cost Questionnaire, LSAS = Liebowitz Social Anxiety Scale, SAFE = Subtle Avoidance Frequency Examination, DASS = Depression Anxiety Stress Scale, WSAP = Word Sentence Association Paradigm, RT = reaction time, BAT = Behavioral Avoidance Test.

The pattern of correlations was consistent with hypotheses of convergent and discriminant validity. Social items on the outcome probability bias task demonstrated moderate positive correlations with paper-and-pencil measures of outcome probability and cost biases for the controlled block (r = 0.42 – 0.48). Correlations were weak to moderate for the automatic block (r = 0.29 – 0.33). Weaker correlations were detected with paper-and-pencil measures of depression and stress for the controlled block (r = 0.28 - 0.36) and the automatic block (r = 0.18 – 0.21).

4.3.5 Construct Validity

Construct validity was evaluated by exploring differences in outcome probability bias task ratings based on stimulus type, social anxiety symptoms above and below the clinical cutoff (60; Mennin et al., 2012), and the order in which participants completed the automatic and controlled blocks. To demonstrate construct validity, I hypothesized that both blocks of the outcome probability bias task would show significantly higher mean ratings for social images relative to control images. I also hypothesized that mean outcome probability bias task ratings would be higher among people with social anxiety symptoms above the clinical cutoff relative to those with scores below the clinical cutoff. For the automatic block, I hypothesized outcome probability bias task ratings would be higher among people who completed the automatic block after completing the controlled block, particularly for familiar images.

4.3.5.1 Controlled Block

A 3 (image type: social 1st, social 3rd, control) x 2 (social anxiety symptom score above or below the clinical cutoff) x 2 (block order: automatic first, controlled first) mixed-design ANOVA was conducted to assess construct validity of the controlled block. The assumption of sphericity was violated so Greenhouse-Geisser corrected values are reported for within-subjects variables (image type). Consistent with hypotheses and replicating results from the pilot study, main effects of image type (F (1.45, 209.04) = 578.91, $p < 0.01$, partial $\eta^2 = 0.80$) and social anxiety symptoms (F (1, 144) = 39.91, $p < 0.01$, partial $\eta^2 = 0.22$) were significant. Regarding the main effect of image type, Bonferroni corrected pairwise comparisons revealed significantly higher outcome probability bias task ratings in response to both social 1st and social 3rd images when compared to control images (M differences = 36.24 and 38.00, respectively, both p values < 0.01). There was no significant difference between social 1st and social 3rd images (M difference = 1.76, $p > 0.05$). Regarding the main effect of social anxiety symptoms, outcome probability bias task ratings were significantly higher among those with symptoms above the clinical cutoff relative to those below the clinical cutoff. There was also a significant main effect of block order (F (1, 144) = 16.54, $p < 0.01$, partial $\eta^2 = 0.10$) such that outcome probability bias task ratings were higher among those who completed the automatic block first.

These main effects were qualified by significant interactions between image type and block order (F (2, 144) = 7.00, $p < 0.01$, partial $\eta^2 = .05$; see Figure 4.1) and between image type and social anxiety symptom score (F (2, 144) = 27.05, $p < 0.01$, partial $\eta^2 = .16$; see Figure 4.2). There was not a significant interaction between social anxiety symptom score and block order ($p > 0.05$) and the three-way interaction was nonsignificant ($p > 0.05$). These interaction effects reflected that higher outcome probability bias task ratings among participants who completed the automatic block first and with symptoms of social anxiety above the clinical cutoff were specific to social images (image type x block order: F (1, 144) = 8.61, $p < 0.01$, partial $\eta^2 = .06$; image type x social anxiety symptom score: F (1, 144) = 30.74, $p < 0.01$, partial $\eta^2 = .18$). Evidence that outcome probability bias ratings were higher among people with symptoms of social anxiety above the clinical cutoff for social images, but not control images was consistent with our hypotheses and supports the construct validity of the outcome probability bias task.

An order effect was unexpected for the controlled block. The fact that the interaction between block order and social anxiety symptom score was nonsignificant indicates the number of people with higher scores was not greater in one block order compared to the other. As such, people who completed the automatic block first may have given higher ratings due to priming or demand characteristics.

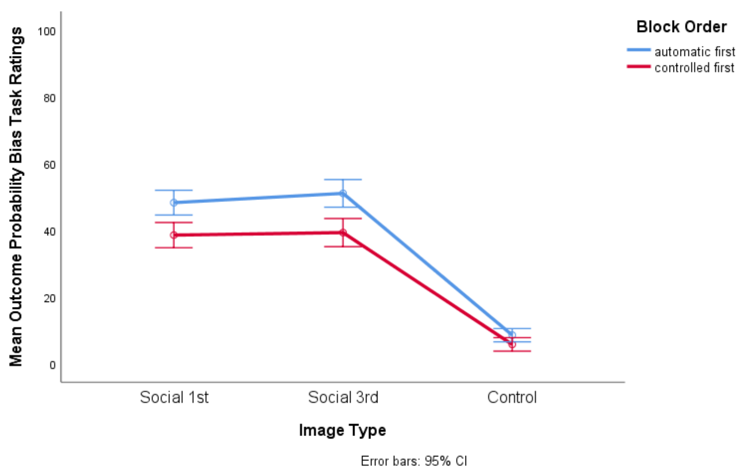


Figure 4.1 Main Study Outcome Probability Bias Task Ratings by Image Type and Block Order in the Controlled Block Collapsing Across Social Anxiety Symptom Score

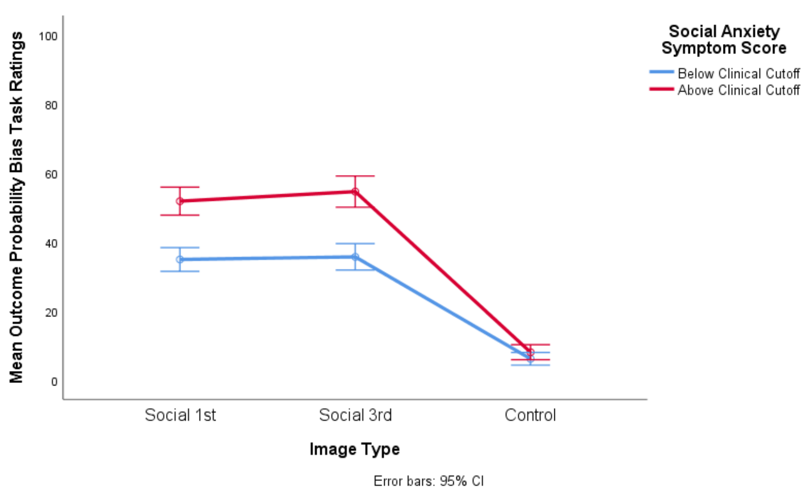


Figure 4.2 Main Study Outcome Probability Bias Task Ratings by Image Type and Social Anxiety Symptom Score in the Controlled Block Collapsing Across Block Order

4.3.5.2 Automatic Block

Assessment of participants' assigned block order was particularly relevant for evaluation of construct validity of the automatic block because it allowed for evaluation of Butler and Mathews' (1983) hypothesis that outcome probability bias results from the availability heuristic. I expected to replicate findings from the pilot study: mean outcome probability bias task ratings were higher among participants who completed the controlled block first. Novel images that were not present in the controlled block were added to the automatic block for the main study. I hypothesized that higher outcome probability bias task ratings in response to familiar images relative to novel images would support the interpretation that the order effect was due to the availability heuristic, which can result from familiarity, rather than demand bias.

A 3 (image type: novel, familiar, control) x 2 (block order: automatic first, controlled first) x 2 (social anxiety symptom score above or below the clinical cutoff) mixed-design ANOVA was used to test the hypothesis that outcome probability bias task ratings would be higher for social images relative to nonsocial images and among people with social anxiety symptom scores above the clinical cutoff relative to those with social anxiety symptom scores below the clinical cutoff. Image types included social images that were also presented in the controlled block (auto familiar), social images that were only included in the automatic block (auto novel), and control images. All social images elicited a 1st person perspective. The assumption of sphericity was violated so Greenhouse-Geisser corrected values are reported for within-subjects variables (image type). Consistent with the hypothesis and with results from the pilot study as well as the controlled block in the main study, there were significant main effects of image type (F (1.42, 204.33) = 321.35, $p < 0.01$, partial $\eta^2 = 0.69$) and social anxiety symptom score (F (1, 144) = 26.04, $p < 0.01$, partial $\eta^2 = 0.15$). In contrast, and inconsistent with our hypothesis, the main effect of block order was not significant ($p > 0.05$). Regarding the main effect of image type, Bonferroni corrected pairwise comparisons revealed significantly higher outcome probability bias task ratings in response to both novel and familiar social images when compared to control images (*M* differences = 41.77 and 48.01, respectively, both *p* values < 0.01). The difference between novel and familiar social images was also significant (*M* difference = 6.24, $p > 0.05$). Mean outcome probability bias task ratings were higher for auto familiar images compared to auto new images. Regarding the main effect of social anxiety symptom score, outcome probability bias task ratings were significantly higher among those with symptoms of social anxiety above the clinical cutoff.

These main effects were qualified by significant interactions between image type and social anxiety symptom score (F (1.42, 204.33) = 5.27, $p = 0.01$, partial $\eta^2 = 0.04$) and between image type and block order (F (1.42, 204.33) = 13.70, $p < 0.01$, partial $\eta^2 = 0.09$). The interaction between block order and social anxiety symptom score was not significant, nor was the three-way interaction (*p* values > 0.05). Regarding the interaction between image type and block order, outcome probability bias task ratings were significantly higher among participants with symptoms of social anxiety above the clinical cutoff for social images, specifically (F (1, 144) = 5.37, $p = 0.02$, partial $\eta^2 = 0.04$; see Figure 4.3). Regarding the interaction between image type and block order, those who completed the controlled block first reported higher outcome probability bias task ratings relative to those who completed the automatic block first, but only for social images (see Figure 4.4). This interaction effect was consistent with the order effect detected in the pilot study.

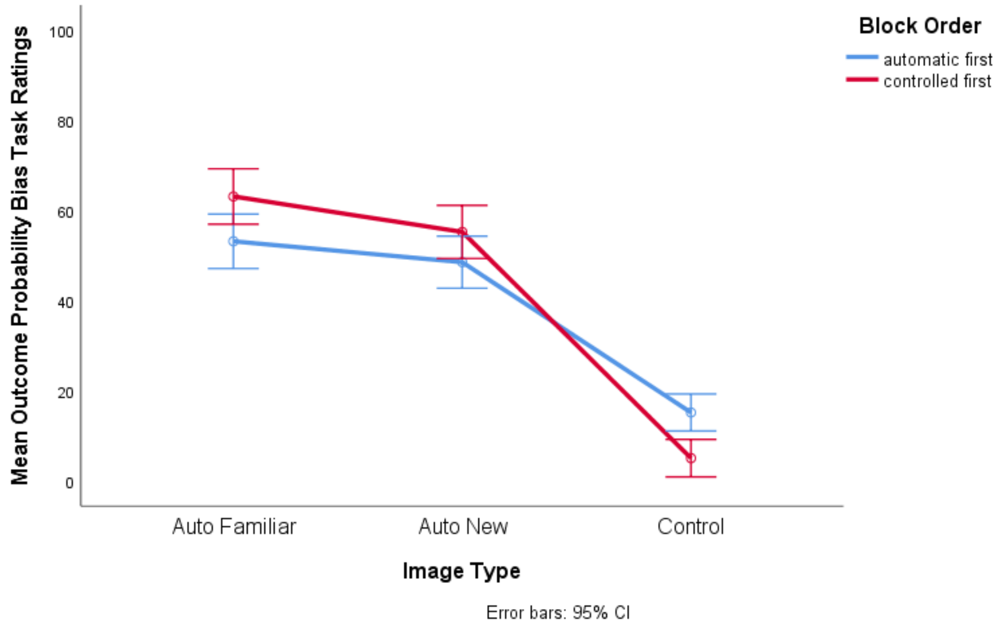


Figure 4.3 Main Study Outcome Probability Bias Task Ratings by Image Type and Block Order for the Automatic Block Collapsing Across Social Anxiety Symptom Score

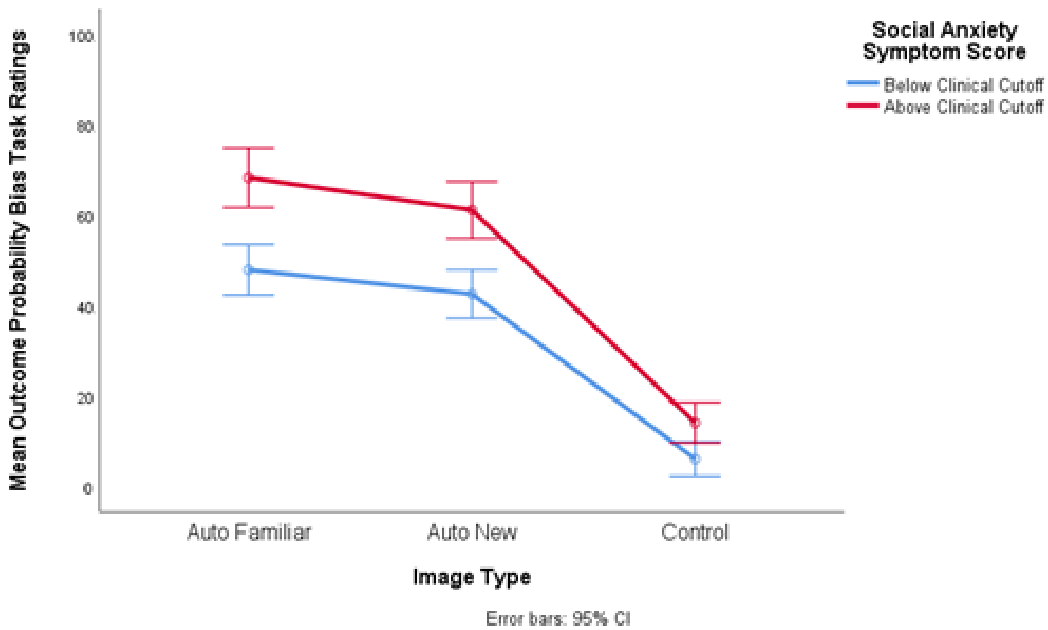


Figure 4.4 Main Study Outcome Probability Bias Task Ratings by Image Type and Social Anxiety Symptom Score for the Automatic Block Collapsing Across Block Order

A follow-up one-way ANOVA revealed no significant difference in outcome probability bias task ratings for novel images based on block order (F (1, 147) = 1.27, p = 0.26, partial η^2 = 0.01). Although outcome probability bias task ratings for familiar images were higher among those who completed the controlled block first relative to those who completed the automatic block first, the difference was not statistically significant (F (1, 147) = 3.09, p = 0.08, partial η^2 = 0.02).

Accuracy and confidence ratings were also examined to evaluate if the automatic block elicited a greater degree of automatic processing relative to controlled processing. The average number of accurate responses to comprehension questions was 4.69 out of eight (SD = 1.15), which is just above the level of chance, supporting a high degree of automatic processing and a minimum of controlled processing. The average confidence rating was 1.85 out of four (SD = 0.36). Low ratings reflect low confidence in the accuracy of responses to comprehension questions about the automatic trials and also support the high degree of influence of automatic processing on responses.

4.3.6 Criterion Validity

A series of multiple regression analyses were used to test the hypothesis that social items from the outcome probability bias task would be associated with self-reported symptoms of social anxiety, self-reported safety behaviors during the virtual reality speech task (BAT), duration of the speech, and peak subjective units of distress during the speech. Manipulation checks were implemented to assess the virtual reality speech task's ability to induce state anxiety.

4.3.6.1 Manipulation Checks for the Virtual Reality Behavioral Avoidance Test

All participants completed self-report measures of state anxiety (STAI), presence (TPI) and cybersickness (SSQ). I expected participants with symptoms of social anxiety above the clinical cutoff would report higher state anxiety following the virtual reality speech task relative to those with symptoms of social anxiety below the clinical cutoff.

Due to unequal sample sizes, a Mann-Whitney U analysis was used to evaluate post-speech state anxiety among those with social anxiety symptom scores above and below the clinical cutoff. The analysis was significant ($p < 0.01$), indicating state anxiety was significantly higher after the virtual reality speech task among participants with high social anxiety relative to those with low social anxiety.

Presence ratings, measured by the TPI, were generally high, with a mean score of 89.06 (SD = 21.26) out of a possible 140. Items with frequently low ratings (rated a three or less out of seven) included “how much did it seem as if you could reach out and touch the objects or people in the virtual environment” (42.4%), “how often did you want to try to touch something in the virtual environment” (68.5%), “to what extent did you feel you could interact with the people in the virtual environment” (44.4%), and “how much control over the interaction with the people in the virtual environment did you feel you had?” (51.7%). This was not surprising given that there were no tactile stimuli as part of the virtual reality environment and the avatars did not verbally respond to the participant.

Symptoms of cybersickness, measured by the SSQ, were somewhat common; however, it is important to note that many of the items in the questionnaire overlap with fear symptoms (e.g. general discomfort, difficulty focusing, sweating, nausea, difficulty concentrating, stomach awareness). The mean SSQ score was 21.66 (SD = 5.47) out of a possible 64. Taken together, evaluations of the state anxiety, presence, and cybersickness support the validity of the virtual reality speech task in the present study.

4.3.6.2 Regression Analyses

Multiple regression analyses were used to evaluate the criterion validity of the outcome probability bias task by testing the hypothesis that outcome probability bias task ratings would significantly predict self-reported symptoms of social anxiety, self-reported safety behaviors during a speech to a virtual audience, speech length, and subjective distress during the speech. Notably, a high degree of correlation among measures, such as the different types of social images included in the outcome probability bias task, may contribute to problems with multicollinearity (Field, 2013). As such, social images from the outcome probability bias task were collapsed across social 1st and social 3rd image types for the controlled block and across novel and familiar image types for the automatic block.

Symptoms of depression and stress, measured by the DASS, were controlled for in all analyses, which would provide further support for the discriminant validity of the outcome probability bias task. Interpretation bias, measured by the WSAP, was also controlled for in all analyses to determine whether outcome probability bias produces effects above and beyond interpretation bias. This would promote dissociation between outcome probability bias and interpretation bias. As such, DASS stress, DASS depression, WSAP benign bias, and WSAP threat bias scores were entered in the first step and mean outcome probability bias task ratings in response to social and control images were entered in the second step for all analyses. WSAP benign bias scores were calculated by subtracting reaction times for benign rejections from reaction times for benign endorsements. WSAP threat bias scores were calculated by subtracting reaction times for threat target word rejections from reaction times for threat target word endorsements. Results from regression analyses are reported in Table 4.4.

Table 4.4 Results from Regression Analyses Assessing the Relations Among the Outcome Probability Bias Task in the Main Study and Symptoms of Social Anxiety, Safety Behaviors, Speech Length, and Peak Subjective Units of Distress Controlling for Symptoms of Depression and Stress and Interpretation Bias

DV	Predictor	Model	β	*P*	95% CI
LSAS	Controlled Block	Step 1: F (4, 144) = 14.48, p < 0.01, R^2 = 0.29 Step 2: F (6, 144) = 19.87, p < 0.01, ΔR^2 = 0.17			
	Social Images		0.45	<0.01	0.43 – 0.81
	Control Images		-0.02	0.77	-0.46 – 0.34
	Automatic Block	Step 1: F (4, 144) = 14.48, p < 0.01, R^2 = 0.29 Step 2: F (6, 144) = 15.21, p < 0.01, ΔR^2 = 0.11			
	Social Images		0.30	<0.01	0.16 – 0.44
	Control Images		0.11	0.12	-0.04 – 0.34
SAFE	Controlled Block	Step 1: F (4, 144) = 7.15, p < 0.01, R^2 = 0.17 Step 2: F (6, 144) = 6.95, p < 0.01, ΔR^2 = 0.06			
	Social Images		0.27	<0.01	0.05 – 0.24
	Control Images		0.02	0.82	-0.17 – 0.21
	Automatic Block	Step 1: F (4, 144) = 7.15, p < 0.01, R^2 = 0.17 Step 2: F (6, 144) = 6.38, p < 0.01, ΔR^2 = 0.05			
	Social Images		0.20	0.01	0.02 – 0.14
	Control Images		0.07	0.35	-0.05 – 0.13
BAT Speech Length	Controlled Block	Step 1: F (4, 144) = 2.46, p = 0.048, R^2 = 0.07 Step 2: F (6, 144) = 1.70, p > 0.05, ΔR^2 = 0.00			
	Automatic Block	Step 1: F (4, 144) = 2.46, p = 0.048, R^2 = 0.07 Step 2: F (4, 144) = 1.66, p > 0.05, ΔR^2 = 0.00			

Peak SUDS	Controlled Block	Step 1: F (4, 143) = 3.94, p < 0.01, R^2 = 0.10 Step 2: F (6, 143) = 6.81, p < 0.01, ΔR^2 = 0.13			
	Social Images		0.41	<0.01	0.03 – 0.08
	Control Images		-0.14	0.09	-0.09 – 0.01
	Automatic Block	Step 1: F (4, 143) = 3.94, p < 0.01, R^2 = 0.10 Step 2: F (4, 143) = 4.78, p < 0.01, ΔR^2 = 0.07			
	Social Images		0.28	0.01	0.01 – 0.04
	Control Images		-0.03	0.67	-0.03 – 0.02

Note. LSAS = Liebowitz Social Anxiety Scale, SAFE = Subtle Avoidance Frequency Examination, BAT = Behavioral Avoidance Test, SUDS = Subjective Units of Distress.

All models were significant for both the automatic and controlled blocks of the outcome probability bias task across measures except for speech length. Among the significant models, social images were identified as significant predictors and control images were not.

4.3.7 Factor Structure

Principal axis factoring with orthogonal (varimax) rotation was used to evaluate the factor structure of both blocks of the outcome probability bias task. Varimax rotation maximizes the dispersion of factor loadings and is particularly useful for interpretation of factors (Field, 2009).

4.3.7.1 Controlled Block

Data from both the pilot study and the main study were combined for the exploratory factor analysis of the controlled block (n = 203). Evaluation of the scree plot and eigenvalues indicated the presence of 3 factors. Rotated factor loadings are reported in Table 4.5. The first factor was primarily comprised of images relevant to public speaking anxiety. The images with the highest loadings included images containing audiences, podiums, and microphones. Several images of job interviews, shaking hands with someone in a suit, and meetings also loaded on this factor. The second factor was composed of social gatherings and parties. Most of the images in the third factor were close up images of people's faces. The rest were images where people seem to be making negatively judgmental facial expressions. None of the control images loaded significantly on any of the factors.

Table 4.5 Rotated Factor Loadings from Exploratory Factor Analysis of Outcome Probability Bias Task Controlled Block

Image Filename	Image Description	Factor 1	Factor 2	Factor 3
Social009e	public speaking	**0.87**	0.10	0.12
Canva37e	publlic speaking	**0.83**	0.16	0.18
Canva4	public speaking	**0.83**	0.18	0.09
Canva34e	public speaking	**0.82**	0.11	0.11
Canva39e	public speaking	**0.79**	0.03	0.14
Canva42e	public speaking	**0.78**	0.15	0.22
Social002	public speaking	**0.78**	0.13	0.13
Canva35e	public speaking	**0.77**	0.25	0.06
Social045e	public speaking	**0.77**	0.19	0.21
Canva27	public speaking	**0.76**	0.15	0.17
Canva40e	interview	**0.76**	0.14	0.16
Canva44e	interview	**0.76**	0.09	0.19
Canva33e	public speaking	**0.75**	0.11	0.11
Canva28	public speaking	**0.74**	0.18	0.07
Canva23	interview	**0.74**	0.07	0.26
Canva36e	public speaking	**0.73**	0.08	0.20
Canva5	public speaking	**0.70**	0.18	*0.31*
Social044e	interview	**0.70**	0.18	0.22
Canva2	interview	**0.68**	0.10	0.22
Canva79	interview	**0.67**	0.26	0.00

Canva25	interview	**0.67**	0.19	0.20
Canva67e	business meeting	**0.66**	0.23	0.14
Social024e	asking teacher a question	**0.62**	0.29	0.04
Canva1	handshake	**0.61**	0.29	0.07
Canva58e	business meeting	**0.61**	0.13	0.27
Canva69e	raising hand in talk	**0.61**	0.26	0.21
Canva68e	in front of class	**0.60**	0.23	0.13
Canva41e	interview	**0.60**	0.00	*0.43*
Canva29e	handshake	**0.58**	*0.30*	-0.01
Canva24	handshake	**0.54**	0.22	0.02
Canva64e	public setting	**0.54**	0.21	*0.34*
Canva31e	handshake	**0.53**	0.29	-0.03
Social037	audience	**0.52**	0.08	0.19
Canva62e	meeting	**0.52**	0.18	0.21
Canva63e	working with teacher	**0.51**	0.25	0.02
Social004	audience	**0.50**	0.22	0.22
Canva76e	meeting	**0.47**	*0.36*	0.01
Canva73e	working with teacher	**0.41**	0.22	0.11
Canva21	eye-contact man working	**0.38**	*0.34*	0.17
Canva12	business meeting	**0.38**	0.26	*0.35*
Canva14	meeting	**0.36**	*0.34*	0.09
Social039	audience	**0.30**	0.14	0.14
Canva47e	party	0.26	**0.72**	-0.03
Canva13	bar	0.25	**0.69**	0.14
Canva50e	party	0.24	**0.68**	0.08

Canva56e	party	0.25	**0.67**	0.08
Canva55e	bar	0.26	**0.66**	0.06
Canva48e	party	0.20	**0.65**	0.08
Canva43e	party	*0.31*	**0.62**	0.01
Canva75e	3 people drinking	0.25	**0.52**	*0.31*
Canva8	students working, smiling	*0.32*	**0.51**	0.05
Faces234	party	0.14	**0.49**	0.04
Canva54e	party	0.29	**0.46**	0.21
Canva6	two women in elevator	0.21	**0.45**	0.27
Canva51e	bar	0.14	**0.45**	*0.34*
Canva49e	party	0.29	**0.37**	0.25
Canva18	eating in public	0.18	**0.37**	0.24
Canva20	face	0.22	0.01	**0.77**
Social010	face	0.24	0.16	**0.74**
Canva15	face	*0.36*	0.03	**0.71**
Canva19	face	*0.30*	0.09	**0.70**
Canva3	bored panel	*0.43*	-0.04	**0.57**
Canva10	gossiping girls	*0.39*	0.22	**0.57**
Canva9	3 people strong eye-contact	*0.34*	0.20	**0.51**
Faces300	face	0.12	0.10	**0.47**
Canva78e	meeting	*0.31*	0.22	**0.47**
Canva7	girls whispering	*0.36*	*0.35*	**0.40**
Faces142	face	0.17	0.11	**0.39**
Social059	face	0.24	0.10	**0.38**
Social029	face	0.16	0.19	**0.36**

Canva22	walking into a talk	0.18	0.20	**0.32**
N172	control	0.23	0.05	0.08
N497	control	0.12	0.07	0.07
N496	control	0.10	-0.01	0.06
N606	control	0.10	0.04	0.00
N197	control	0.08	0.22	0.05
N566	control	0.07	0.05	0.03
N200	control	0.07	0.12	0.07
N323	control	0.06	0.09	-0.06
N706	control	0.05	0.07	0.05
N499	control	0.05	0.17	0.10
N132	control	0.04	-0.05	0.08
N498	control	0.04	0.23	0.01
N293	control	0.03	0.06	-0.06
N574	control	0.02	0.03	0.15
N221	control	0.02	0.10	0.07
N170	control	0.01	0.09	0.14
N572	control	0.00	0.03	0.11
N296	control	-0.01	0.12	0.09
N477	control	-0.01	0.09	0.03
N101	control	-0.01	0.01	0.14
N654	control	-0.01	0.03	0.01
N282	control	-0.03	0.10	0.02
N144	control	-0.06	0.02	0.16

Factor correlations			
Factor 1	1		
Factor 2	.64*	1	
Factor 3	.65*	.56*	1

Note. Extraction done via principal axis factoring with varimax rotation. The strongest loading for each image is in boldface. Loadings at or above .30 are shown in italics if the loading is not the strongest for that particular image.

Means and standard deviations for the factors are reported in Table 4.6. Means and standard deviations for participants with social anxiety symptom scores above and below the clinical cutoff (60; Mennin et al., 2012) are also reported.

Table 4.6 Means and Standard Deviations of Outcome Probability Bias for Each Factor Identified in Exploratory Factor Analysis of Controlled Block: Overall and Separated by Social Anxiety Symptom Clinical Cutoff

Factor	Overall M (SD)	Below LSAS Cutoff M (SD)	Above LSAS Cutoff M (SD)
1	47.94 (22.38)	37.91 (22.47)	58.39 (18.43)
2	28.61 (19.66)	20.32 (16.13)	35.10 (18.66)
3	50.72 (20.93)	41.92 (22.29)	58.29 (16.18)

Note. LSAS = Liebowitz Social Anxiety Scale

A 3 (factors) x 2 (social anxiety symptom score above or below the clinical cutoff) mixed-design ANOVA revealed significant main effects of factor F (2, 292) = 123.22, $p < 0.01$, partial $\eta^2 = 0.46$) and social anxiety symptom score (F (1, 146) = 41.51, $p < 0.01$, partial $\eta^2 = 0.22$). Contrasts revealed mean outcome probability bias task ratings were significantly lower for the second factor relative to the first factor (F (1, 146) = 173.92, $p < 0.01$, partial $\eta^2 = 0.54$) and the third factor (F (1, 146) = 196.05, $p < 0.01$, partial $\eta^2 = 0.57$). Outcome probability bias task ratings did not significantly differ between the first and the third factor ($p > 0.05$). For the main effect of social anxiety symptom score, mean outcome probability bias task ratings were higher among participants with social anxiety symptom scores above the clinical cutoff compared to those with scores below the clinical cutoff across all factors. The interaction between factors and social anxiety symptom scores was not significant ($p > 0.05$; see Figure 4.3).

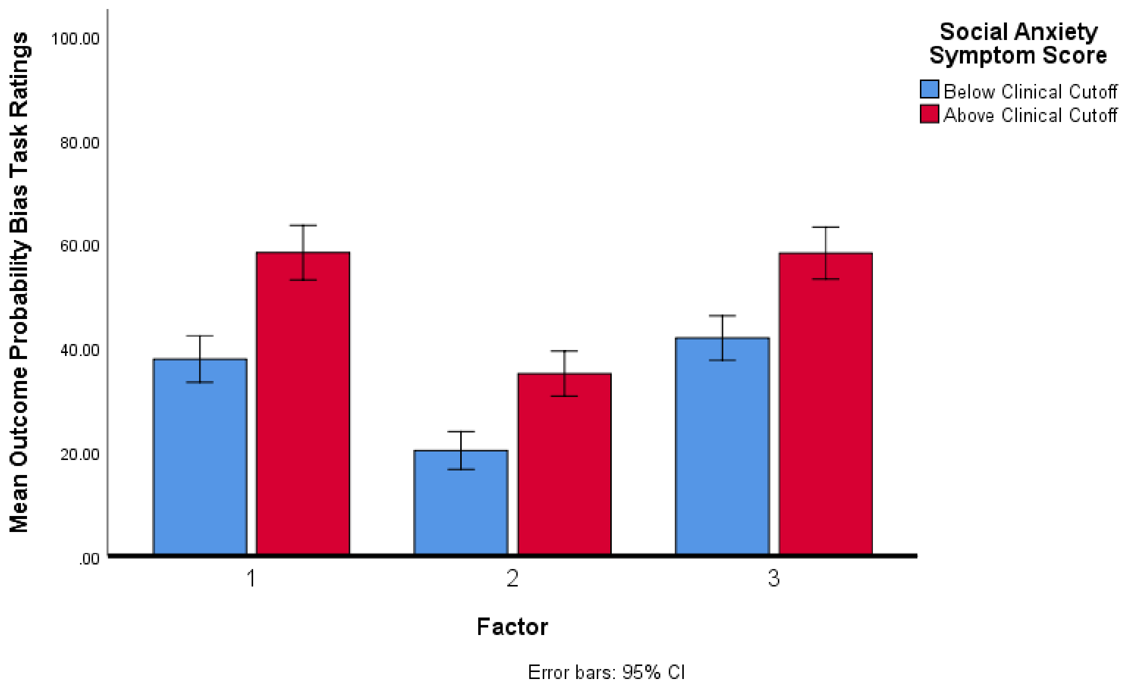


Figure 4.5 Mean Outcome Probability Bias Task Ratings for Factor Structures Identified in Exploratory Factor Analysis of Controlled Block Separated by Social Anxiety Symptom Clinical Cutoff

4.3.7.2 Automatic Block

Only data from the main study were used for the exploratory factor analysis of the automatic block (n = 148). Again, evaluation of the scree plot and eigenvalues indicated the presence of 3 factors (see Table 4.7). Because automatic trials contained three images of the same type, factor loadings are reported for trial image type rather than individual images. The first factor in the present analysis was similar to the first factor from the controlled block and was composed of images of public speaking and interviews. The second factor included trials with images of threatening facial expressions, consistent with the third factor from the controlled block. The two control trials loaded on factor three.

Table 4.7 Rotated Factor Loadings from Exploratory Factor Analysis of Automatic Block

Image Type	Factor 1	Factor 2	Factor 3
Interview (Novel)	**0.84**	0.11	0.11
Public Speaking (Familiar)	**0.79**	0.28	0.07
Interview (Familiar)	**0.78**	0.26	0.13
Public Speaking (New)	**0.68**	0.17	0.20
Faces (New)	0.20	**0.87**	0.00
Faces (Familiar)	*0.31*	**0.80**	-0.05
Control Buildings	0.09	-0.03	**0.91**
Control Objects	0.19	-0.01	**0.74**

Note. Extraction done via principal axis factoring with varimax rotation. The strongest loading for each trial is in boldface. Loadings at or above .30 are shown in italics if the loading is not the strongest for that particular trial.

4.4 Main Study Discussion

4.4.1 Reliability, Convergent Validity, Discriminant Validity, Construct Validity, and Criterion Validity

The present pattern of results supports the reliability, convergent validity, discriminant validity, construct validity, and criterion validity of the outcome probability bias task. Internal consistency ranged from good to excellent for all stimulus types ($\alpha = 0.82 – 0.96$), supporting reliability. Social images included in the outcome probability bias task demonstrated moderate and significant positive correlations with a standardized paper-and-pencil measure of outcome probability bias ($r = 0.33 – 0.48$), supporting convergent validity. The outcome probability bias task demonstrated stronger correlations with conceptually similar measures (paper-and-pencil questionnaires of outcome probability bias and outcome cost bias; controlled block: $r = 0.42 – 0.48$; automatic block: $r = 0.29 – 0.33$) relative to conceptually dissimilar measures (self-report questionnaires of depression and stress; controlled block: $r = 0.28 - 0.36$; automatic block: $r = 0.18 – 0.21$), supporting convergent and discriminant validity (DeVellis, 2011). Outcome probability bias task ratings were higher in response to social images relative to nonsocial images, supporting construct validity. Ratings were also higher among participants with self-reported symptoms of social anxiety above a clinical cutoff score than those with scores below a clinical cutoff score, also supporting construct validity. In addition, social images in the task demonstrated positive and significant associations with self-reported symptoms of social anxiety (LSAS; $\beta = 0.30 – 0.45$, $p < 0.05$), consistent with findings from other studies that relied on paper-and-pencil questionnaires to measure outcome probability bias (Calamaras et al., 2015; Hoffart et al., 2009; O'Toole et al., 2015; Smits et al., 2006), supporting criterion validity. Social images in the task were also positively associated with self-reported safety behaviors (SAFE; $\beta = 0.20 – 0.27$, $p < 0.05$) and subjective units of distress during a virtual reality speech task (SUDS; $\beta = 0.28 – 0.41$, $p < 0.05$). These relations were significant while controlling for interpretation bias and self-reported symptoms of depression and stress, further supporting criterion and discriminant validity of the outcome probability bias task.

Two findings were not consistent with the proposed hypotheses. First, I hypothesized that there would be a positive, significant relation between the outcome probability bias task and speech length during the virtual reality behavioral avoidance test (BAT), but there was none. Ceiling effects may have contributed to this finding, as 55 participants (37%) spoke for the full 10 minutes, the modal speech length. This contributed to significant negative skew in the data and could have increased the likelihood of a false negative result (Fayers, 2011), as limited variability may contribute to limited power to detect an effect. In support of this view, all significant correlations between speech length and other theoretically related measures (interpretation bias, self-reported symptoms of social anxiety, self-reported safety behaviors) were weaker than expected.

Alternatively, this null result could indicate that the virtual reality speech task failed to induce anxiety. Though a virtual reality behavioral avoidance test has been shown to be feasible for assessment of specific phobia (Mühlberger et al., 2008), the feasibility of a virtual reality behavioral avoidance test designed to induce public speaking anxiety has never been evaluated in a non-clinical sample. This alternative explanation is unlikely, however, as post-speech state anxiety ratings were significantly higher among participants with social anxiety symptom scores above the clinical cutoff compared to participants with scores below the clinical cutoff, suggesting the virtual reality speech task was effective at inducing anxiety among people for whom it would likely be a fear-relevant social situation.

Another possible interpretation of this null finding is that outcome probability bias may not relate to behavioral avoidance. In support of this view, speech length was also unrelated to a paper-and-pencil questionnaire of outcome probability bias (OPQ) in the present study. Lack of a relation between outcome probability bias and behavioral avoidance is inconsistent with theories of social anxiety disorder (Clark, 2001; Clark & Wells, 1995; Heimberg et al., 2010; Rapee & Heimberg, 1997), so these unexpected findings warrant additional research. For example, the outcome probability bias task could be modified to induce either a positive or negative outcome probability bias. Following outcome probability bias training, participants would complete an unmodified version of the task to verify a bias had been induced successfully. Subsequently, participants would be instructed to engage in a behavioral avoidance test. Higher levels of avoidance (measured by speech length and/or safety behaviors) among participants in the negative outcome probability bias condition relative to those in the positive bias condition would indicate that outcome probability bias does, in fact, contribute to behavioral avoidance.

Second, I hypothesized there would be higher ratings in the automatic block for familiar images relative to novel images among participants who completed the controlled block first, but this was not the case. One of the goals of the main study was to evaluate the extent to which the availability heuristic may be contributing to probability assessments at an automatic level of processing. In the pilot study, outcome probability bias task ratings were higher, on average, when participants completed the controlled block first as compared to completing the automatic block first. Because some of the images were used in both blocks, participants who completed the controlled block first (during which the images were processed at a slower rate) were likely to be familiar with the images when they appeared for the second time during the automatic block. Increased familiarity has been shown to enhance the availability heuristic (Tversky & Kahneman, 1974). As such, this order effect potentially supported the effect of the availability heuristic on outcome probability bias.

Alternatively, higher ratings could have resulted from increased experience with the task among those who had already completed the controlled block. This interpretation was supported by the presence of an order effect in the controlled block in the main study: participants who completed the automatic block first gave higher outcome probability bias task ratings during the controlled block relative to those who completed the controlled block first. The automatic block in the main study was modified in order to rule out increased experience as an alternative explanation for higher outcome probability bias task ratings among participants who completed the controlled block first. In the main study, automatic trials containing novel images from the same category (i.e. public speaking, interview) that were not included in the controlled block were added. I hypothesized that order effects would be present for familiar images, but not for novel images. Though outcome probability bias task ratings were higher for familiar images relative to novel images among those who completed the controlled block first relative to those who completed the automatic block first, this difference was not statistically significant. Notably, familiar images were rated higher than novel images regardless of block order, suggesting that the novel images may not have been comparable to the familiar images despite the fact that they depicted the same social themes. Additional research using the automatic block with more trials and using familiar and novel images with previously established outcome probability ratings that are comparable is needed to clarify the role of the availability heuristic versus demand bias in outcome probability bias.

The fact that the outcome probability bias task significantly predicted symptoms of social anxiety while controlling for interpretation bias should be interpreted with caution. The aim of this analysis was to test discriminant validity. More specifically, I hypothesized that the outcome probability bias task would positively predict symptoms of social anxiety above and beyond variability accounted for by interpretation bias, and this would support the dissociation between these two constructs. Notably, the interpretation bias measure included in the present study (WSAP) was weakly correlated with social anxiety symptoms measured by the LSAS (r = 0.17). Other researchers have found that this measure of interpretation bias is positively related to social anxiety symptoms with medium to large effect sizes (Gonsalves et al., 2019), so the weak correlation in the present study is puzzling. It is possible that method variance could have led to poor precision of the measure in this study. Fatigue may have contributed to poor precision because the study procedure typically took one to two hours to complete and the interpretation bias task was the final task. Another source of method variance is that the outcome probability bias task is an imagery-based task and the interpretation bias task is a verbal-based task.

4.4.2 Factor Structure

I designed the outcome probability bias task with the intent of measuring a single latent variable: outcome probability bias. Images included in the task, however, were quite diverse and reflected a wide variety of social situations that have the potential to result in negative judgment, embarrassment, or humiliation (e.g. public speaking, going to a party, being stared at). An exploratory factor analysis was used to investigate the extent to which there may be different subfactors of outcome probability bias that, to date, have yet to be identified.

Analysis of the controlled and automatic blocks yielded three factors. The first factor was primarily comprised of images most commonly identified by pilot participants as anxiety provoking in their anecdotal feedback (public speaking, interviews). The second factor contained images of parties, social gatherings, and drinking in public. The third factor included images that emphasized negative, judgmental, or threatening facial expressions. Analyses revealed mean outcome probability bias task ratings were significantly higher for the first and third factors relative to the second factor. Given that the second factor is composed of images of parties, this second factor may only elicit outcome probability bias among those with more generalized symptoms of social anxiety.

Alternatively, the distinction between the second factor and the other two factors may reflect differences in outcome probability bias when responding to social images with positive, neutral, and negative valences. Many of the images in the first factor are ambiguous. Images in the second factor commonly feature people who are smiling and appear to be enjoying themselves.

Images in the third factor include people with negative or threatening facial expressions. As such, outcome probability bias may be differentially relevant to anticipation of positive and negative social encounters. Relatedly, researchers have found that people with social anxiety experience fear of positive and negative evaluation (Weeks, Heimberg, Rodebaugh, & Norton, 2008). Going forward, it may be important for researchers to specify the nature of outcome probability bias (negative or positive) and the type of social situation that elicits this bias (commonly anticipated to be threatening, ambiguous, or rewarding).

Future research is needed to explore potential differences in the relations among the factors identified in the controlled block of the outcome probability bias task and symptoms of social anxiety. Higher levels of outcome probability bias in response to the second factor, for example, could potentially be associated with more generalized, more severe, and/or more treatment resistant symptoms of social anxiety. For example, expectation violations necessary for effective exposure therapy (Craske et al., 2008) may be more difficult to achieve if outcome probability bias is more generalized and applies to social situations when people without social anxiety disorder tend to predict a positive outcome is more likely. A longitudinal exposure-based treatment study involving regular administration of the outcome probability bias task to a clinical sample of people with specific and generalized symptoms of social anxiety disorder could shed light on this possibility. Before such research is undertaken, a confirmatory factor analysis will be needed to demonstrate stability of the three factors identified in the current study. Evaluation of the outcome probability bias task in an independent sample is recommended for more valid hypothesis testing of factor structure (DeVellis, 2011).

4.4.3 Recommended Changes for Future Use

The outcome probability bias task would benefit from a significant reduction in items. The high internal consistency of social items and very strong correlation between the social 1st and social 3rd images in the controlled block indicate redundancy among the images. Further, exploratory factor analyses did not identify separate factors between social 1st and social 3rd images, indicating the distinction between these image types may have been unnecessary. Replication of the present findings with a clinical sample may be needed to verify this, however, because utilization of mental imagery of the self as seen by the audience is more common among people with social anxiety disorder relative to those without the disorder (Hackmann et al., 1998; Wells et al., 1998). Outcome probability bias task ratings may therefore be higher in response to social 3rd images relative to social 1st images among people with social anxiety disorder but not for those without the disorder.

The Spearman-Brown prophecy formula (Brown, 1910; Spearman, 1910) was used to indicate how many items should be kept to maintain an internal consistency of 0.92. The result indicated 27%, or approximately 20 images would be sufficient. As such, it is feasible to divide the images included in the controlled block of the outcome probability bias task in half and develop two “forms” of the task. This would allow for further evaluation of the task’s reliability.

Another recommendation for change is to make the stimulus presentation duration for the controlled block constant rather than self-paced. Notably, some of the participants rapidly advanced to the rating screen during the controlled block. It is possible, therefore, that they viewed the images just as quickly as they did the automatic trials during the automatic block, limiting the construct validity of the controlled block. Based on research from Calvo et al. (1999), controlled block images should be set to 1000ms.

5 DISCUSSION

The current study presents the first computer task designed to measure outcome probability bias in social anxiety. Results demonstrate this task's good to excellent internal consistency, support the task's validity, and identify a tripartite factor structure. It is also the first study to evaluate outcome probability bias from automatic and controlled levels of processing and the first to attempt to dissociate outcome probability bias from interpretation bias.

Researchers have suggested that outcome probability bias results from the availability heuristic (Butler & Mathews, 1983) and may manifest at both controlled and automatic levels of processing (Beck & Clark, 1997). To date, studies of outcome probability bias have only measured this construct at a controlled level of processing. Better understanding of the role of automatic processing in outcome probability bias can inform theoretical models of cognitive mechanisms in anxiety. The present study used rapid presentation of images to assess whether outcome probability bias is manifest at an automatic level of processing.

It can be challenging, however, for researchers to experimentally isolate automatic processes. In fact, McNally (1995) has claimed that it may not be possible to develop an experimental task that fully isolates automatic processing. Consistent with this, Jacoby (1998) has indicated that all tasks measure levels of processing on a spectrum from mostly automatic to mostly controlled. Therefore, the automatic block of the outcome probability bias task likely elicited responding that incorporated some controlled processing. Jacoby's (1998) process dissociation procedure can be used to assess the degree of automatic processing measured in a task. In this procedure, estimates of controlled and automatic processing are estimated by assessing responses that are similar to or different from a previously made response. The present study's design was not compatible with this procedure. Instead, automaticity was evaluated through accuracy in response to comprehension questions about the images and confidence ratings. In the main study, mean accuracy of responses to comprehension questions was just above the level of chance and average confidence ratings were fairly low, supporting a high degree of automatic processing in the automatic block.

Controlled processing has been shown to interfere with biased judgments that result from the availability heuristic. In Menon and Raghubir's (2003) study, for example, the availability heuristic did not affect participants' judgments when they were aware of the potential for this bias, but only if participants were aware of the potential for biased processing *before* they made a judgment and if they were under conditions of low cognitive load. This may inform the mechanism through which cognitive restructuring contributes to reductions in outcome probability bias in treatment. Increased awareness of cognitive errors and monitoring of thoughts may attune people to the potential for unrealistic thinking and prevent it before it starts. This approach may only be effective, however, when people are not experiencing high cognitive load (e.g. from stress, distractions, or even symptoms). Experimental manipulation of outcome probability bias is needed to test this possibility.

For example, a modified version of the outcome probability bias task could be used to induce negative outcome probability bias among people without social anxiety disorder. After a test phase demonstrating successful induction of bias, participants could receive psychoeducation about the impact of the availability heuristic on judgments of probability before viewing and rating a new set of images, half of which are novel and half of which are familiar. Reminders of the effect of familiarity on judgment could be given before or after participants viewed the images. Participants would then give their ratings of outcome probability. Participants could also respond to the new set of images under conditions of high or low cognitive load. Outcome probability bias would be expected to persist among participants who received reminders about the impact of familiarity on judgment after viewing the images and among participants under high cognitive load. Outcome probability bias ratings may be reduced in response to reminders about the impact of familiarity on judgment before viewing the images and among participants under low cognitive load. This pattern of results would support the role of the availability heuristic in outcome probability bias. Inclusion of measures of symptoms of social anxiety before and after this procedure could also be used to clarify the conditions under which cognitive restructuring may reduce symptoms through reduction of outcome probability bias.

Evaluation of outcome probability bias at an automatic level of processing may also inform the dissociation of outcome probability bias from interpretation bias. Interpretation bias and outcome probability bias have sometimes been conflated in the literature (Mathews and MacLeod, 1994). Though others have argued that outcome probability bias and interpretation bias are distinct (Amir et al., 1998; Blanchette & Richards, 2010), this is the first study to test this possibility. Beck and Clark's (1997) information processing model of anxiety posits that outcome probability bias manifests in the immediate preparation stage, which involves a combination of automatic and controlled processing, and interpretation bias manifests in the secondary elaboration stage, which only involves controlled processing. In support of this model, previous research indicates that interpretation bias is not detected when stimuli are presented at shorter durations, conditions under which automatic processing predominates (e.g. Calvo & Castillo, 1997; Richards & French, 1992). Evidence of outcome probability bias following rapid stimulus presentations in the present study may therefore indicate that this bias is distinguishable from interpretation bias because it occurs during an earlier stage of cognitive processing (i.e. the immediate preparation stage, which is influenced by a combination of automatic and controlled processing; Beck & Clark, 1997).

Notably, all known studies that have evaluated interpretation bias resulting from automatic processing have utilized verbal tasks. It is possible that automatic effects on interpretation bias aren't detectible for lexical tasks but are detectible for visual ones. As such, interpretation bias may, in fact, result from automatic processing but this has yet to be detected due to sole reliance on verbal tasks to assess it. Further research on interpretation bias using visual stimuli is needed to explore this possibility. The present task format could potentially be modified for this purpose. For example, participants could view rapidly presented images or short videos of socially ambiguous situations that elicit a third-person perspective. Participants could then quickly choose from two images that disambiguate the previous stimulus in either a benign or threatening manner. Higher proportions of threatening images would reflect greater interpretation bias.

Though replication and additional reliability and validity testing is needed, the present results show promise that the outcome probability bias task could be used for experimental manipulation of this construct. Currently, the causal role of outcome probability bias in social anxiety disorder is unknown. Though theories of social anxiety disorder posit that outcome probability bias causes symptoms (Clark, 2001; Clark & Wells, 1995; Heimberg et al., 2010; Rapee & Heimberg, 1997), it could be produced by them. Similar to manipulation of measures of attention and interpretation bias to allow for cognitive bias modification (Hertel & Mathews, 2011), it may be possible to alter the present task to simulate outcome probability bias in people with low levels of social and public speaking anxiety. Outcome probability bias may potentially be induced using the present task by incorporating feedback about the actual likelihood of social embarrassment. Alternatively, participants could be presented with images showing a negative outcome from the social situation that was rated previously. Participants would complete many trials with consistent feedback that the likelihood of social embarrassment was high. I would expect outcome probability ratings to increase over the course of training. Participants would rate new images with higher levels of negative outcome probability relative to participants who received no training or positive outcome probability bias training. I would also expect higher levels of symptoms of social anxiety following completion of negative outcome probability bias induction relative to those who received positive outcome probability bias induction.

These methods could also be used in the development of a cognitive bias modification paradigm designed to reduce outcome probability bias and improve symptoms of social anxiety. This could be achieved by providing feedback indicating the likelihood of a feared outcome is low over multiple trials. Notably, it may be important to only use first-person imagery in an outcome probability bias modification paradigm due to the negative impact of mental representations of the self as seen by the audience on symptoms and performance (Hirsch et al., 2003; Hirsch et al., 2004; Spurr & Stopa, 2003; Stopa & Jenkins, 2007; Vassilopoulos, 2005). Research shows that cognitive biases and self-imagery interact and may contribute to symptoms (Hirsch et al., 2007). Engagement in third-person imagery while anticipating a social encounter may therefore interfere with potential ameliorative effects of outcome probability bias modification training. In contrast, training participants to engage in more first-person imagery may enhance therapeutic effects of outcome probability bias modification training. This could be tested by randomly assigning a sample of people with social anxiety disorder to receive outcome probability bias modification training containing only first- or third-person images. Training with first-person images may predict a greater reduction in outcome probability bias and in symptoms of social anxiety compared to training with third-person images. Alternatively, training with third-person images may be akin to exposure therapy, where greater anxiety induced by the images due to the third-person perspective may enhance expectancy violation (Craske et al., 2008) and therefore convey a greater therapeutic effect.

Outcome probability bias is present in a variety of other anxiety and related disorders in addition to social anxiety disorder, including specific phobia (Menzies & Clarke, 1995), generalized anxiety disorder (Butler & Mathews, 1983), agoraphobia (McNally & Foa, 1987; Poulton & Andrews, 1996), panic disorder (Uren et al., 2004) and acute stress disorder (Warda & Bryant, 1998). The present task may be adapted to assess, and potentially manipulate, outcome probability bias associated with these disorders as well. Given that bias has been found to be specific to the type of disorder (Gilboa-Schechtman et al., 2000; Poulton & Andrews, 1996; Uren et al., 2004), the social images included in the present outcome probability bias task may not be applicable to other disorders and would need to be replaced with images (or even short video clips) of situations in which escape is not possible (agoraphobia), people displaying symptoms of a heart attack (panic disorder), situations that commonly induce worry such as financial ruin (generalized anxiety disorder), or stimuli commonly used in exposure therapy for trauma (acute stress disorder). The SFIP contains several high-quality images of blood and spiders that could be used to modify the task for outcome probability bias in specific phobia (Michałowski et al., 2017). Regarding acute stress disorder, research has shown that early exposure-based intervention can improve prognosis and prevent development of posttraumatic stress disorder (Kearns et al., 2012). As such, an outcome probability bias modification program containing images from a recent trauma affecting many individuals (such as the COVID-19 pandemic) could potentially be an intervention for acute stress geared towards prevention of posttraumatic stress disorder.

A notable limitation of the present study is generalizability of the outcome probability bias task to other populations due to use of a nonclinical sample of college students. The use of a convenience sample of college students can lead to unreliable estimates of behavior and limit replicability of findings (Peterson & Merunka, 2014). Generalizability of findings to a clinical sample is unknown. Another limitation to the present study is that the construct validity of the controlled block was limited by the self-paced stimulus presentation duration, as some participants may have viewed the images too quickly to allow for controlled processing. In addition, it was not possible to measure the test-retest reliability of the outcome probability bias task because the task was only administered once. It is possible that outcome probability bias is stable over time, but this has not been examined in the literature. Researchers have found that participants with social anxiety disorder who completed a randomized controlled trial approximately 5 years prior reported similar (improved) levels of outcome probability bias at post-treatment and long-term follow-up (Benbow & Anderson, 2019). Longitudinal research is needed, however, to truly establish the course of outcome probability bias over time. Assuming outcome probability bias is stable, evaluation of test-retest reliability will be important for the internal validity of any studies that may use the outcome probability bias task in the future.

Another important limitation to consider is that it was not possible to determine if participants were rating the images based on their perceptions of outcome probability or simply based on the level of anxiety evoked by the image. Several studies investigating the role of the availability heuristic in probability judgments indicate that probability judgments are impacted by trait anxiety and memory for relevant outcomes more so than priming or state anxiety (Butler & Mathews, 1987; Johnson & Tversky, 1983; MacLeod & Campbell, 1992; Zelenski & Larsen, 2002); however, the effect of priming and state anxiety on probability judgments using the present paradigm is unknown.

Despite these limitations, the present study details the development of a new tool that could enhance our understanding of outcome probability bias in social anxiety. Pending replication of the psychometric properties found in the current study, researchers may potentially use this task for multimodal assessment, experimental manipulation, and even treatment of outcome probability bias in social anxiety and other related disorders. In addition, further evaluation of the factor structure of this task could reveal new subfactors of outcome probability bias and inform theoretical models of social anxiety disorder.

REFERENCES

Acosta, A., & Vila, J. (1990). Emotional imagery: Effect of autonomic response information on physiological arousal. *Cognition and Emotion, 4*(2), 145-160. doi: 10.1080/02699939008407144

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Amir, N., Foa, E. B., & Coles, M. E. (1998). Negative interpretation bias in social phobia. *Behaviour Research and Therapy*, *36*(10), 945-957. doi: 10.1016/S0005-7967(98)00060-6

Anderson, P. L., Price, M., Edwards, S. M., Obasaju, M. A., Schmertz, S. K., Zimand, E., & Calamaras, M. R. (2013). Virtual reality exposure therapy for social anxiety disorder: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 81*(5), 751–760. doi: 10.1037/a0033559

Anderson, P. L., Zimand, E., Hodges, L. F., & Rothbaum, B. O. (2005). Cognitive behavioral therapy for public-speaking anxiety using virtual reality for exposure. *Depression and Anxiety, 22*(3), 156–158. doi: 10.1002/da.20090

Andrews, G., Freed, S., & Teesson, M. (1994). Proximity and anticipation of a negative outcome in phobias. *Behaviour Research and Therapy, 32*(6), 643–645. doi: 10.1016/0005-7967(94)90020-5

Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychological Assessment, 10*(2), 176–181. doi: 10.1037/1040-3590.10.2.176

Barlow, D. H., Farchione, T. J., Sauer-Zavala, S., Latin, H. M., Ellard, K. K., Bullis, J.R., Bentley, K. H., Boettcher, H. T., & Cassiello-Robbins, C. (2017). *Unified protocol for transdiagnostic treatment of emotional disorders: Therapist guide (Second Edition).* Oxford University Press.

Beard, C., & Amir, N. (2008). A multi-session interpretation modification program: Changes in interpretation and social anxiety symptoms. *Behaviour Research and Therapy*, *46*(10), 1135–1141. doi: 10.1016/j.brat.2008.05.012

Beard, C., & Amir, N. (2009). Interpretation in social anxiety: When meaning precedes ambiguity. *Cognitive Therapy and Research*, *33*(4), 406–415. doi: 10.1007/s10608-009-9235-0

Beck, A. T. (1976) *Cognitive Therapy and the Emotional Disorders.* International Universities Press.

Beck, A. T., & Clark, D. A. (1997). An information processing model of anxiety: Automatic and strategic processes. *Behaviour Research and Therapy, 35*(1), 49-58. doi: 10.1016/S0005-7967(96)00069-1

Beck, A. T., Emery, G., & Greenberg, R. L. (2005). *Anxiety disorders and phobias: A cognitive perspective.* Basic Books.

Beidel, D. C., Turner, S. M., Jacob, R. G., & Cooley, M. R. (1989). Assessment of social phobia: Reliability of an impromptu speech task. *Journal of Anxiety Disorders*, *3*(3), 149-158. doi: 10.1016/0887-6185(89)90009-1

Benbow, A. A., & Anderson, P. L. (2019). Long-term improvements in probability and cost biases following brief cognitive behavioral therapy for social anxiety disorder. *Cognitive Therapy and Research, 43*(2), 412-418. doi: 10.1007/s10608-018-9947-0

Birnboim, S. (2003). The automatic and controlled information-processing dissociation: Is it still relevant?. *Neuropsychology Review, 13*(1), 19-31. doi: 10.1023/A:1022348506064

Blanchette, I., & Richards, A. (2010). The influence of affect on higher level cognition: A review of research on interpretation, judgement, decision making and reasoning. *Cognition & Emotion, 24*(4), 561-595. doi: 10.1080/02699930903132496

Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology, 1904- 1920, 3*(3), 296-322. doi: 10.1111/j.2044-8295.1910.tb00207.x

Butler, G., & Mathews, A. (1983). Cognitive processes in anxiety. *Advances in Behaviour Research and Therapy, 5*(1), 51–62. doi: 10.1016/0146-6402(83)90015-2

Butler, G., & Mathews, A. (1987). Anticipatory anxiety and risk perception. *Cognitive Therapy and Research, 11*(5), 551-565. doi: 10.1007/BF01183858

Calamaras, M. R., Tully, E. C., Tone, E. B., Price, M., & Anderson, P. L. (2015). Evaluating changes in judgmental biases as mechanisms of cognitive-behavioral therapy for social anxiety disorder. *Behaviour Research and Therapy, 71*, 139-149. doi: 10.1016/j.brat.2015.06.006

Calvo, M. G. (2000). The time course of predictive inferences depends on contextual constraints. *Language and Cognitive Processes, 15*(3), 293–319. doi: 10.1080/016909600386066

Calvo, M. G., & Castillo, M. D. (1997). Mood congruent bias in interpretation of ambiguity strategic processes and temporary activation. *The Quarterly Journal of Experimental Psychology, 50A*(1), 163-182. doi: 10.1080/713755684

Calvo, M. G., & Castillo, M. D. (1998). Predictive inferences take time to develop. *Psychological Research, 61*, 249–260. doi: 10.1007/s004260050029

Calvo, M. G., & Castillo, M. D. (2001). Bias in predictive inferences during reading. *Discourse Processes, 32*(1), 43-71. doi: 10.1207/S15326950DP3201_03

Calvo, M. G., Castillo, M. D., & Estevez, A. (1999). On-line predictive inferences in reading: Processing time during versus after the priming context. *Memory & Cognition, 27*(5), 834-843. doi: 10.3758/BF03198536

Calvo, M. G., Eysenck, M. W., & Castillo, M. D. (1997). Interpretation bias in test anxiety: The time course of predictive inferences. *Cognition & Emotion, 11*(1), 43-64. Doi: 10.1080/026999397380023

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81-105. doi: 10.1037/h0046016

Carr, A. T. (1974). Compulsive neurosis: A review of the literature. *Psychological Bulletin, 81*(5), 311-318. doi: 10.1037/h0036473

Carroll, J. S. (1978). The effect of imagining an event on expectations for the event: An interpretation in terms of the availability heuristic. *Journal of Experimental Social Psychology, 14*(1), 88-96. doi: 10.1016/0022-1031(78)90062-8

Clark, D. M. (2001). A cognitive perspective on social phobia. In W. R. Crozier, & L. E. Alden (Eds.), *International handbook of social anxiety: Concepts, research and interventions relating to the self and shyness* (pp. 405-430). John Wiley & Sons.

Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 69-93). The Guilford Press.

Coles, M. E., & Heimberg, R. G. (2000). Patterns of anxious arousal during exposure to feared situations in individuals with social phobia. *Behaviour Research and Therapy, 38*(4), 405–424. doi: 10.1016/S0005-7967(99)00092-3

Cooper, A. A., Clifton, E. G., & Feeny, N. C. (2017). An empirical review of potential mediators and mechanisms of prolonged exposure therapy. *Clinical Psychology Review, 56*, 106-121. doi: 10.1016/j.cpr.2017.07.003

Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy*, *46*(1), 5–27. doi: 10.1016/j.brat.2007.10.003

Cronbach, L. J. (1970). *Essentials of psychological testing* (3rd ed.). Harper & Row.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281-302. doi: 10.1037/h0040957

Cuming, S., Rapee, R. M., Kemp, N., Abbott, M. J., Peters, L., & Gaston, J. E. (2009). A self-report measure of subtle avoidance and safety behaviors relevant to social anxiety: Development and psychometric properties. *Journal of Anxiety Disorders, 23*(7), 879-883. doi: 10.1016/j.janxdis.2009.05.002

DeVellis, R. F. (2011). *Scale development: Theory and applications.* Sage.

Do, C. B., & Batzoglou, S. (2008). What is the expectation maximization algorithm? *Nature Biotechnology, 26*(8), 897-899. doi: 10.1038/nbt1406

Donaldson, T. S. (1968). Robustness of the *F*-test to errors of both kinds and the correlation between the numerator and denominator of the *F*-ratio. *Journal of the American Statistical Association*, *63*(322), 660-676. doi: 10.2307/2284037

Embretson (Whitely), S. E. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin*, *93*(1), 179–197. doi: 10.1037/0033-2909.93.1.179

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272–299. doi: 10.1037/1082-989X.4.3.272

Fayers, P. (2011). Alphas, betas and skewy distributions: Two ways of getting the wrong answer. *Advances in Health Sciences Education, 16*(3), 291-296. doi: 10.1007/s10459-011-9283-6

Field, A. (2013). *Discovering statistics using IBM SPSS statistics.* Sage.

Foa, E. B., Franklin, M. E., Perry, K. J., & Herbert, J. D. (1996). Cognitive biases in generalized social phobia. *Journal of Abnormal Psychology, 105*(3), 433-439. doi: 10.1037/0021-843X.105.3.433

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin, 99*(1), 20-35. doi: 10.1037/0033-2909.99.1.20

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review*, *107*(2), 289–344. doi: 10.1037/0033-295X.107.2.289

Gilboa-Schechtman, E., Franklin, M. E., & Foa, E. B. (2000). Anticipated reactions to social events: Differences among individuals with generalized social phobia, obsessive compulsive disorder, and nonanxious controls. *Cognitive Therapy and Research*, *24*(6), 731-746. doi: 10.1023/A:1005595513315

Gold, M. S., & Bentler, P. M. (2000). Treatments of missing data: A Monte Carlo comparison of RBHDI, iterative stochastic regression imputation, and expectation-maximization. *Structural Equation Modeling, 7*(3), 319-355. doi: 10.1207/S15328007SEM0703_1

Gonsalves, M., Whittles, R. L., Weisberg, R. B., & Beard, C. (2019). A systematic review of the word sentence association paradigm (WSAP). *Journal of Behavior Therapy and Experimental Psychiatry, 64,* 133–148. doi: 10.1016/j.jbtep.2019.04.003

Gregory, B., Peters, L., Abbott, M. J., Gaston, J. E., & Rapee, R. M. (2015). Relationships between probability estimates, cost estimates, and social anxiety during CBT for social anxiety disorder. *Cognitive Therapy and Research*, *39*(5), 636–645. doi: 10.1007/s10608-015-9692-6

Hackmann, A., Surawy, C., & Clark, D. M. (1998). Seeing yourself through others' eyes: A study of spontaneously occurring images in social phobia. *Behavioural and Cognitive Psychotherapy, 26*(1), 3-12. doi: 10.1017/S1352465898000022

Hallion, L. S., & Ruscio, A. M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychological Bulletin, 137*(6), 940-958. doi: 10.1037/a0024355

Heimberg, R. C., Brozovich, F. A., & Rapee, R. M. (2010). A cognitive behavioral model of social anxiety disorder: Update and extension. In S. G. Hofmann & P. M. DiBartolo (Eds.), *Social anxiety: Clinical, developmental, and social perspectives* (pp. 395-422). Elsevier Academic Press. doi: 10.1016/B978-0-12-375096-9.00015-8

Heimberg, R. G., Liebowitz, M. R., Hope, D. A., Schneier, F. R., Holt, C. S., Welkowitz, L. A., Juster, H. R., Campeas, R., Bruch, M. A., Cloitre, M., Fallon, B., & Klein, D. F. (1998). Cognitive behavioral group therapy vs phenelzine therapy for social phobia: 12-week outcome. *Archives of General Psychiatry, 55*(12), 1133–1141. doi: 10.1001/archpsyc.55.12.1133

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. *British Journal of Clinical Psychology, 44*(2), 227–239. doi: 10.1348/014466505X29657

Hertel, P. T., & Mathews, A. (2011). Cognitive bias modification: Past perspectives, current findings, and future applications. *Perspectives on Psychological Science, 6*(6), 521-536. doi: 10.1177/1745691611421205

Hirsch, C. R., Clark, D. M., Mathews, A., & Williams, R. (2003). Self-images play a causal role in social phobia. *Behaviour Research and Therapy, 41*(8), 909-921. doi: 10.1016/S0005-7967(02)00103-1

Hirsch, C. R., Mathews, A., & Clark, D. M. (2007). Inducing an interpretation bias changes self-imagery: A preliminary investigation. *Behaviour Research and Therapy, 45*(9), 2173-2181. doi: 10.1016/j.brat.2006.11.001

Hirsch, C., Meynen, T., & Clark, D. (2004). Negative self-imagery in social anxiety contaminates social interactions. *Memory, 12*(4), 496-506. doi: 10.1080/09658210444000106

Hoffart, A., Borge, F., Sexton, H., & Clark, D. (2009). Change processes in residential cognitive and interpersonal psychotherapy for social phobia: A process-outcome study. *Behavior Therapy, 40*, 10–22. doi: 10.1016/j.beth.2007.12.003

Hofmann, S.G. (2004). Cognitive mediation of treatment change in social phobia. *Journal of Consulting and Clinical Psychology, 72*, 392-399. doi: 10.1037/0022-006X.72.3.392

Holmes, E. A., & Mathews, A. (2005). Mental imagery and emotion: A special relationship? *Emotion, 5*(4), 489-497. doi: 10.1037/1528-3542.5.4.489

Holmes, E. A., Mathews, A., Dalgleish, T., & Mackintosh, B. (2006). Positive interpretation training: Effects of mental imagery versus verbal training on positive mood. *Behavior Therapy, 37*(3), 237-247. doi: 10.1016/j.beth.2006.02.002

Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy, 48*(1), 9–16. doi: 10.1037/a0022186

Ionescu, M. R. (2016). Subliminal perception of complex visual stimuli. *Romanian Journal of Ophthalmology, 60*(4), 226 - 230.

Jacoby, L. L. (1998). Invariance in automatic influences of memory: Toward a user's guide for the process-dissociation procedure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(1), 3-26. doi: 10.1037/0278-7393.24.1.3

Johnson, E. J., & Tversky, A. (1983). Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology, 45*(1), 20-31. doi: 10.1037/0022-3514.45.1.20

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*, 1–27. doi: 10.1146/annurev.clinpsy.3.022806.091432

Kearns, M. C., Ressler, K. J., Zatzick, D., & Rothbaum, B. O. (2012). Early interventions for PTSD: A review. *Depression and Anxiety, 29*(10), 833-842. doi: 10.1002/da.21997

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology, 3*(3), 203-220. doi: 10.1207/s15327108ijap0303_3

Kim, D., Bae, H., & Chon Park, Y. (2008). Validity of the subjective units of disturbance scale in EMDR. *Journal of EMDR Practice and Research, 2*(1), 57–62. doi: 10.1891/1933-3196.2.1.57

Lang, P. J., Levin, D. N., Miller, G. A., & Kozak, M. J. (1983). Fear behavior, fear imagery, and the psychophysiology of emotion: The problem of affective response integration. *Journal of Abnormal Psychology, 92*(3), 276-306. doi: 10.1037/0021-843X.92.3.276

Lee, K. M. (2004). Presence, explicated. *Communication Theory, 14*(1), 27-50. doi: 10.1111/j.1468-2885.2004.tb00302.x

Liebowitz, M. R. (1987). Social phobia. *Modern Problems in Pharmacopsychiatry, 22*, 141–173. doi: 10.1159/000414022

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*(404), 1198-1202. doi: 10.2307/2290157

Lombard, M., Ditton, T. B., & Weinstein, L. (2009). Measuring presence: The temple presence inventory. *Proceedings of the 12th Annual International Workshop on Presence* (1-15).

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. doi: 10.1016/0005-7967(94)00075-U

Lucock, M. P., & Salkovskis, P. M. (1988). Cognitive factors in social anxiety and its treatment. *Behaviour Research and Therapy, 26*(4), 297–302. doi: 10.1016/0005-7967(88)90081-2

MacLeod, C., & Campbell, L. (1992). Memory accessibility and probability judgments: An experimental evaluation of the availability heuristic. *Journal of Personality and Social Psychology, 63*(6), 890-902. doi: 10.1037//0022-3514.63.6.890

Marchewka, A., Żurawski, Ł., Jednoróg, K., & Grabowska, A. (2014). The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Research Methods, 46*(2), 596-610. doi: 10.3758/s13428-013-0379-1

Mathews, A., & MacLeod, C. (1994). Cognitive approaches to emotion and emotional disorders. *Annual Review of Psychology, 45*(1), 25-50. doi: 10.1146/annurev.ps.45.020194.000325

Mathews, A., & MacLeod, C. (2005). Cognitive vulnerability to emotional disorders. *Annual Review of Clinical Psychology*, *1*, 167–195. doi: 10.1146/annurev.clinpsy.1.102803.143916

McManus, F., Clark, D., & Hackmann, A. (2000). Specificity of cognitive biases in social phobia and their role in recovery. *Behavioural and Cognitive Psychotherapy*, *28*(3), 201–209. doi:10.1017/S1352465800003015

McNally, R. J. (1995). Automaticity and the anxiety disorders. *Behaviour Research and Therapy, 33*(7), 747-754. doi: 10.1016/0005-7967(95)00015-P

McNally, R. J., & Foa, E. B. (1987). Cognition and agoraphobia: Bias in the interpretation of threat. *Cognitive Therapy and Research, 11*(5), 567-581. doi: /10.1007/BF01183859

Mennin, D. S., Fresco, D. M., Heimberg, R. G., Schneier, F. R., Davies, S. O., & Liebowitz, M. R. (2002). Screening for social anxiety disorder in the clinical setting: Using the Liebowitz Social Anxiety Scale. *Journal of Anxiety Disorders, 16*(6), 661–673. doi: 10.1016/S0887-6185(02)00134-2

Menon, G., & Raghubir, P. (2003). Ease-of-retrieval as an automatic input in judgments: A mere-accessibility framework? *Journal of Consumer Research, 30*(2), 230-243. doi: 10.1086/376804

Menzies, R. G., & Clarke, J. C. (1995). Danger expectancies and insight in acrophobia. *Behaviour Research and Therapy, 33*(2), 215-221. doi: 10.1016/0005-7967(94)P4443-X

Michałowski, J. M., Droździel, D., Matuszewski, J., Koziejowski, W., Jednoróg, K., & Marchewka, A. (2017). The Set of Fear Inducing Pictures (SFIP): Development and validation in fearful and nonfearful individuals. *Behavior Research Methods, 49*(4), 1407-1419. doi: 10.3758/s13428-016-0797-y

Moscovitch, D. A., Gavric, D. L., Senn, J. M., Santesso, D. L., Miskovic, V., Schmidt, L. A., McCabe, R.E., & Antony, M. M. (2012). Changes in judgment biases and use of emotion regulation strategies during cognitive-behavioral therapy for social anxiety disorder: Distinguishing treatment responders from nonresponders. *Cognitive Therapy and Research*, *36*(4), 261–271. doi: 10.1007/s10608-011-9371-1

Mühlberger, A., Sperber, M., Wieser, M. J., & Pauli, P. (2008). A virtual reality behavior avoidance test (VR-BAT) for the assessment of spider phobia. *Journal of CyberTherapy and Rehabilitation*, *1*(2), 147-158.

O'Toole, M. S., Mennin, D. S., Hougaard, E., Zachariae, R., & Rosenberg, N. K. (2015). Cognitive and emotion regulation change processes in cognitive behavioural therapy for social anxiety disorder. *Clinical Psychology and Psychotherapy*, *22*(6), 667–676. doi: 10.1002/cpp.1926

Peterson, R. A., & Merunka, D. R. (2014). Convenience samples of college students and research reproducibility. *Journal of Business Research, 67*(5), 1035-1041. doi: 10.1016/j.jbusres.2013.08.010

Poulton, R. G., & Andrews, G. (1996). Change in danger cognitions in agoraphobia and social phobia during treatment. *Behaviour Research and Therapy, 34*(5–6), 413–421. doi: 10.1016/0005-7967(96)00009-5

Raghubir, P., & Menon, G. (1998). AIDS and me, never the twain shall meet: The effects of information accessibility on judgments of risk and advertising effectiveness. *Journal of Consumer Research, 25*(1), 52-63. doi: 10.1086/209526

Rapee, R. M., Gaston, J. E., & Abbott, M. J. (2009). Testing the efficacy of theoretically derived improvements in the treatment of social phobia. *Journal of Consulting and Clinical Psychology, 77*(2), 317–327. doi: 10.1037/a0014800

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

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black, & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). Appleton-Century-Crofts. doi: 10.1101/gr.110528.110

Richards, A., & French, C. C. (1992). An anxiety-related bias in semantic activation when processing threat/neutral homographs. *Quarterly Journal of Experimental Psychology: Section A, 45*(3), 503-525. doi: 10.1080/02724989208250625

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*(1), 1-66. doi: 10.1037/0033-295X.84.1.1

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize?. *Journal of Research in Personality, 47*(5), 609-612. doi: 10.1016/j.jrp.2013.05.009

Schoth, D. E., & Liossi, C. (2017). A systematic review of experimental paradigms for exploring biased interpretation of ambiguous information with emotional and neutral associations. *Frontiers in Psychology, 8*. doi: 10.3389/fpsyg.2017.00171

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments, 10*(3), 266-281. doi: 10.1162/105474601300343603

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology*, *61*(2), 195-202. doi: 10.1037/0022-3514.61.2.195

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*(2), 127-190. doi: 10.1037/0033-295X.84.2.127

Sinclair, S. J., Siefert, C. J., Slavin-Mulford, J. M., Stein, M. B., Renna, M., & Blais, M. A. (2012). Psychometric evaluation and normative data for the depression, anxiety, and stress scales-21 (DASS-21) in a nonclinical sample of US adults. *Evaluation & the Health Professions*, *35*(3), 259-279. doi: 10.1177/0163278711424282

Smits, J. A. J., Rosenfield, D., McDonald, R., & Telch, M. J. (2006). Cognitive mechanisms of social anxiety reduction: An examination of specificity and temporality. *Journal of Consulting and Clinical Psychology, 74*(6), 1203-1212. doi: 10.1037/0022-006X.74.6.1203

Spearman, C. (1910). Correlation calculated with faulty data. *British Journal of Psychology, 1904- 1920, 3*(3), 271-295. doi: 10.1111/j.2044-8295.1910.tb00206.x

Spielberger, C. D., & Gorsuch, R. L. (1983). *State-Trait Anxiety Inventory for Adults: Sampler Set: Manual, Test Booklet and Scoring Key*. Consulting Psychologists Press.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *State-Trait Anxiety Inventory manual.* Consulting Psychologists Press.

Spurr, J. M., & Stopa, L. (2003). The observer perspective: Effects on social anxiety and performance. *Behaviour Research and Therapy, 41*, 1009-1028. doi: 10.1016/S0005-7967(02)00177-8

Stopa, L., & Jenkins, A. (2007). Images of the self in social anxiety: Effects on the retrieval of autobiographical memories. *Journal of Behavior Therapy and Experimental Psychiatry, 38*(4), 459-473. doi: 10.1016/j.jbtep.2007.08.006

Taylor, C., & Alden, L. (2008). Self-related and interpersonal judgment biases in social anxiety disorder: Changes during treatment and relationship to outcome. *International Journal of Cognitive Therapy, 1,* 125–137. doi:10.1521/ijct.2008.1.2.125

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*(2), 207-232. doi: 10.1016/0010-0285(73)90033-9

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131. doi: 10.1126/science.185.4157.1124

Uren, T. H., Szabó, M., & Lovibond, P. F. (2004). Probability and cost estimates for social and physical outcomes in social phobia and panic disorder. *Journal of Anxiety Disorders, 18*(4), 481–498. doi: 10.1016/S0887-6185(03)00028-8

Vassilopoulos, S. P. (2005). Social anxiety and the vigilance-avoidance pattern of attentional processing. *Behavioural and Cognitive Psychotherapy, 33*(1), 13-24. doi: 10.1017/S1352465804001730

Voncken, M. J., & Bögels, S. M. (2006). Changing interpretation and judgmental bias in social phobia: A pilot study of a short, highly structured cognitive treatment. *Journal of Cognitive Psychotherapy, 20*(1), 59-73. doi: 10.1891/jcop.20.1.59

Voncken, M. J., Bögels, S. M., & de Vries, K. (2003). Interpretation and judgmental biases in social phobia. *Behaviour Research and Therapy, 41*(12), 1481-1488. doi: 10.1016/S0005-7967(03)00143-8

Vrana, S. R., Cuthbert, B. N., & Lang, P. J. (1986). Fear imagery and text processing. *Psychophysiology, 23*(3), 247-253. doi: 10.1111/j.1469-8986.1986.tb00626.x

Warda, G., & Bryant, R. A. (1998). Thought control strategies in acute stress disorder. *Behaviour Research and Therapy, 36*(12), 1171-1175. doi: 10.1016/S0005-7967(98)00082-5

Weeks, J. W., Heimberg, R. G., Rodebaugh, T. L., & Norton, P. J. (2008). Exploring the relationship between fear of positive evaluation and social anxiety. *Journal of Anxiety Disorders, 22*(3), 386-400. doi: 10.1016/j.janxdis.2007.04.009

Wells, A., Clark, D. M., & Ahmad, S. (1998). How do I look with my minds eye: Perspective taking in social phobic imagery. *Behaviour Research and Therapy, 36*(6), 631-634. doi: 10.1016/S0005-7967(98)00037-0

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence, 7*(3), 225-240. doi: 10.1162/105474698565686

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition.* Stanford University Press.

Wolpe, J. (1973). *The practice of behavior therapy* (2nd ed.). Pergamon.

Wolpe, J., & Lazarus, A. A. (1966). *Behavior therapy techniques: A guide to the treatment of neuroses.* Pergamon Press.

Zelenski, J. M., & Larsen, R. J. (2002). Predicting the future: How affect-related personality traits influence likelihood judgments of future events. *Personality and Social Psychology Bulletin, 28*(7), 1000-1010. doi: 10.1177/01467202028007012

APPENDICES

Appendix A: Outcome Probability Bias Task Stimuli

Appendix A.1 List of Images Included in Outcome Probability Bias Task

Stimulus Type	Image Source	Image Name
Control (25 images)	**SFIP**	**neutral_197**
	SFIP	neutral_282
	SFIP	**neutral_101**
	SFIP	**neutral_144**
	SFIP	neutral_323
	SFIP	neutral_022
	SFIP	neutral_132
	SFIP	neutral_170
	SFIP	neutral_172
	SFIP	neutral_293
	SFIP	neutral_496
	SFIP	neutral_572
	SFIP	neutral_574
	SFIP	neutral_200
	SFIP	neutral_221
	SFIP	neutral_284
	SFIP	neutral_497
	SFIP	neutral_498
	SFIP	**neutral_477**
	SFIP	**neutral_566**
	SFIP	neutral_606
	SFIP	**neutral_654**
	SFIP	neutral_296
	SFIP	neutral_499
	SFIP	neutral_706
Social 1st (37 images)	**SFIP**	**social_010**
	SFIP	social_029
	NAPS	**Faces_300_h**
	Canva	Canva15
	Canva	**Canva20**
	SFIP	social_059
	NAPS	Faces_142_h
	Canva	Canva19
	Canva	Canva21

	Canva	**Canva79**
	Canva	**Canva2**
	Canva	Canva23
	Canva	Canva24
	Canva	**Canva25**
	Canva	Canva1
	Canva	Canva3
	Canva	Canva26
	SFIP	social_039
	NAPS	Faces_234_h
	Canva	Canva6
	Canva	Canva7
	Canva	Canva8
	Canva	Canva9
	Canva	Canva10
	Canva	Canva12
	Canva	Canva13
	Canva	Canva14
	Canva	Canva18
	Canva	Canva22
	Canva	**Canva4**
	Canva	**Canva5**
	Canva	**Canva27**
	Canva	Canva28
	SFIP	social_002
	SFIP	social_004
	SFIP	social_023
	SFIP	social_037
Social 3rd (38 images)	Canva	Canva36e
	SFIP	social_024e
	SFIP	social_043e
	Canva	Canva43e
	SFIP	social_045e
	Canva	Canva29e
	Canva	Canva31e
	Canva	Canva33e
	Canva	Canva34e
	Canva	Canva35e
	Canva	Canva37e

	SFIP	social_009e
	Canva	Canva39e
	Canva	Canva40e
	Canva	Canva41e
	Canva	Canva42e
	Canva	Canva44e
	SFIP	social_044e
	Canva	Canva48e
	Canva	Canva49e
	Canva	Canva50e
	Canva	Canva51e
	Canva	Canva47e
	Canva	Canva56e
	Canva	Canva54e
	Canva	Canva55e
	Canva	Canva58e
	Canva	Canva59e
	Canva	Canva62e
	Canva	Canva63e
	Canva	Canva64e
	Canva	Canva67e
	Canva	Canva68e
	Canva	Canva69e
	Canva	Canva73e
	Canva	Canva75e
	Canva	Canva76e
	Canva	Canva78e
Auto New Images (All Social 1st)	**Canva**	**AutoNew2**
	Canva	**AutoNew10**
	Canva	**AutoNew3**
	Canva	**AutoNew5**
	Canva	**AutoNew6**
	Canva	**AutoNew11**
	Canva	**AutoNew4**
	Canva	**AutoNew9**
	Canva	**AutoNew8**

Note. Images highlighted in grey were only used as practice trial images for the controlled block. Images highlighted in blue were used in the controlled block and as practice trial images for the automatic block. Images in boldface were used in the main study automatic block.

Appendix A.2 Main Study Automatic Block Items

Practice Trials:

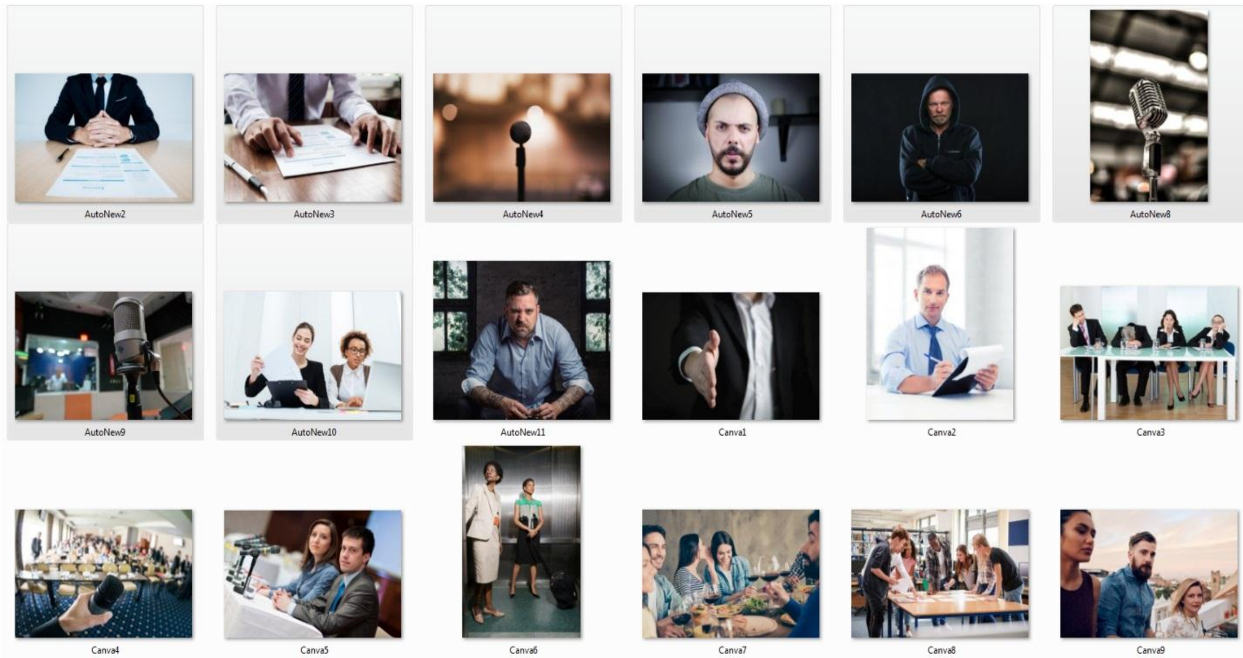
PracImage1	PracImage2	PracImage3	PracImage4	PracImage5	Was1	Was2	Was3	CorrectAns
pics/neutral_170.jpg	pics/neutral_132.jpg	pics/neutral_497.jpg	pics/neutral_499.jpg	pics/neutral_706.jpg	Boots	Window	Grass	1
pics/Canva7.jpg	pics/Canva9.jpg	pics/Canva10.jpg	pics/Faces_234_h.jpg	pics/Canva13.jpg	Job interview	Girls talking	Handshake	2

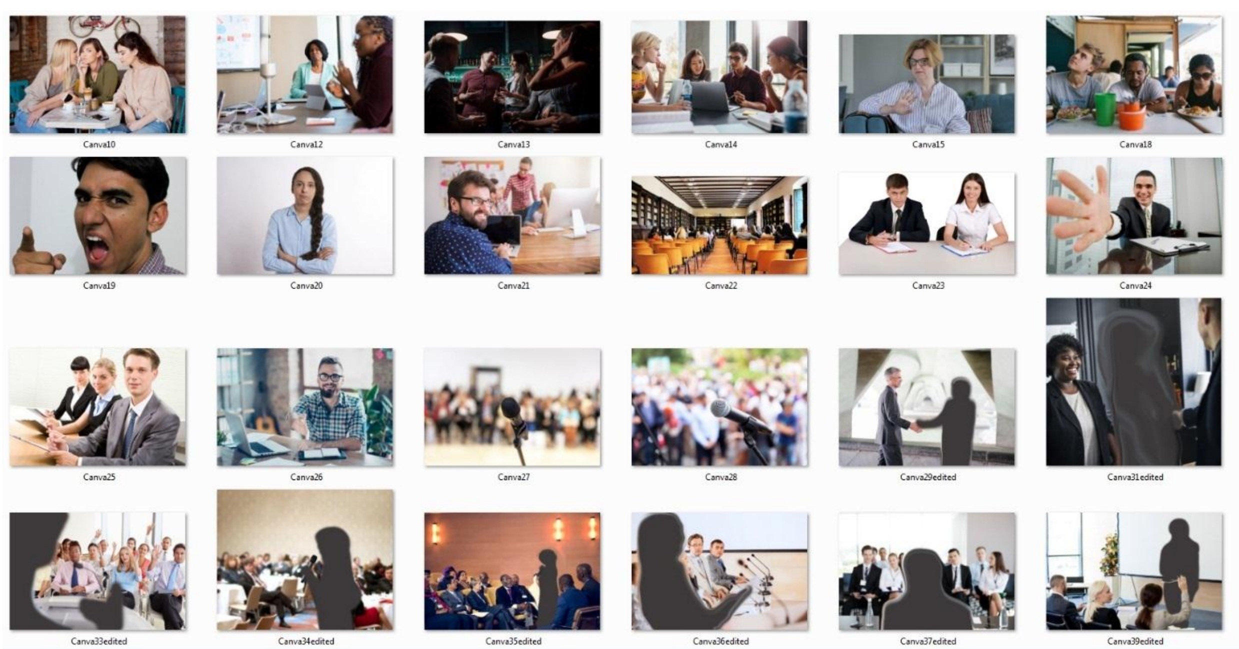
Task:

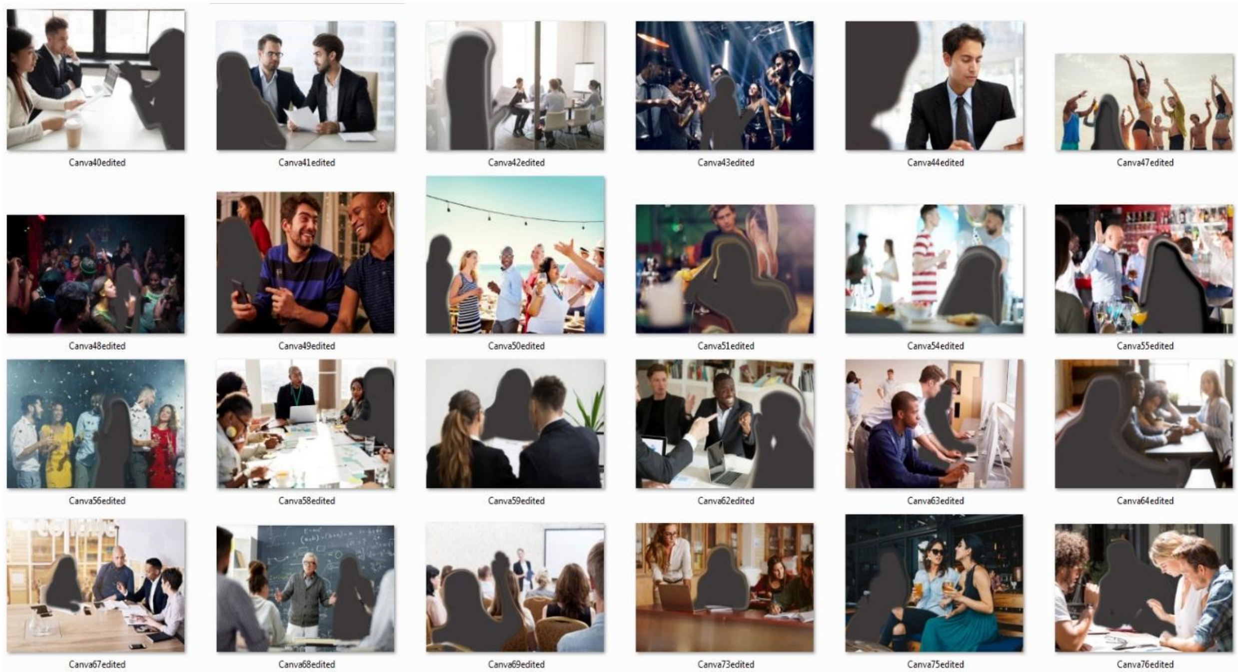
RealImage1	RealImage2	RealImage3	Was1	Was2	Was3	CorrectAns
pics/neutral_197.jpg	pics/neutral_101.jpg	pics/neutral_14	Window	Car	Building	3
pics/neutral_477.jpg	pics/neutral_566.jpg	pics/neutral_65	Money	Leaves	Green apple	2
pics/social_010.jpg	pics/Faces_300_h.jp	pics/Canva20.jp	Old man	Baby	Old woman	1
pics/Canva2.jpg	pics/Canva25	pics/Canva79.p	Two women in an elevator	Three people in suits	Man teaching a class	2
pics/Canva4.jpg	pics/Canva5.jpg	pics/Canva27.jp	People asking questions	Large auditorium	Man and woman	3
pics/AutoNew2.jpg	pics/AutoNew10.jpg	pics/AutoNew3.j	Man holding out hand	Two women	Man and woman	2
pics/AutoNew5.jpg	pics/AutoNew6.jpg	pics/AutoNew7.j	Sweaty man	Old woman	Old man in hoodie	3
pics/AutoNew4.jpg	pics/AutoNew9.jpg	pics/AutoNew8.j	Sound booth	People raising their	Audience and a camera	1

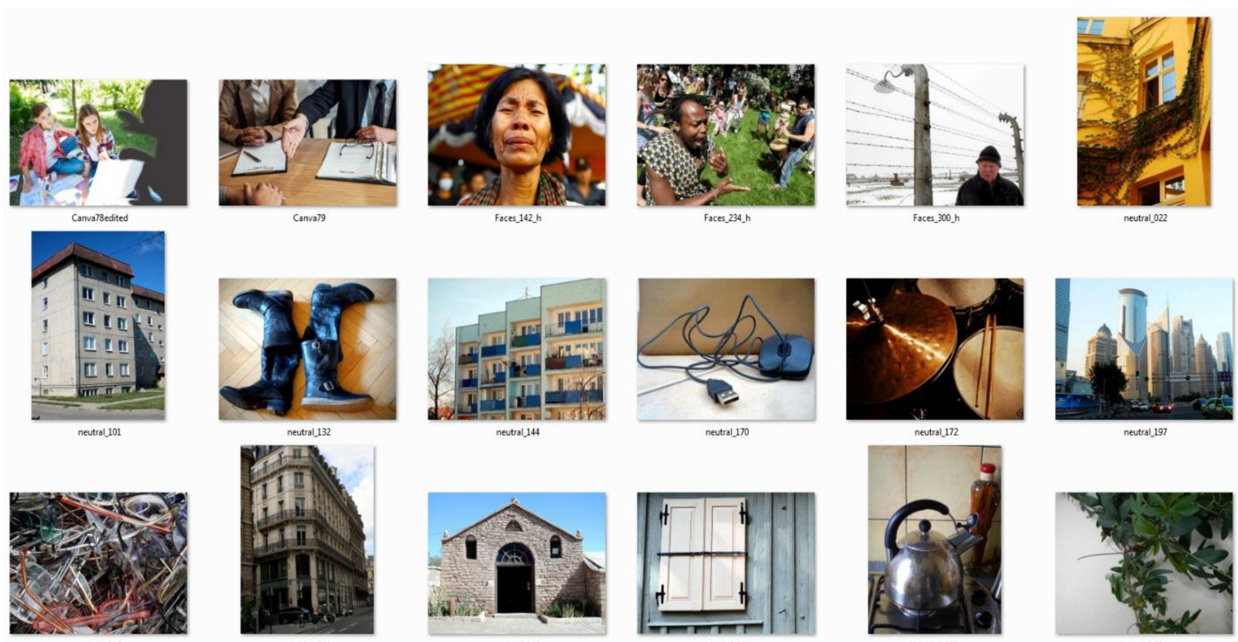
Note. Was1, Was2, and Was3 refer to comprehension question items. CorrectAns identifies the correct comprehension question item.

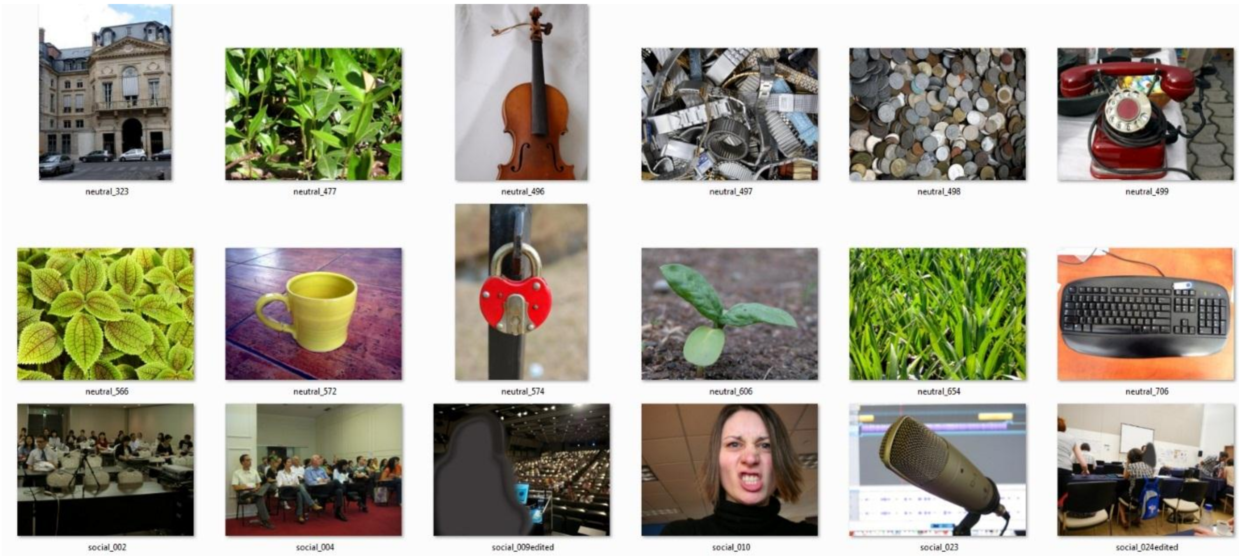
Appendix A.3 Outcome Probability Bias Task Images

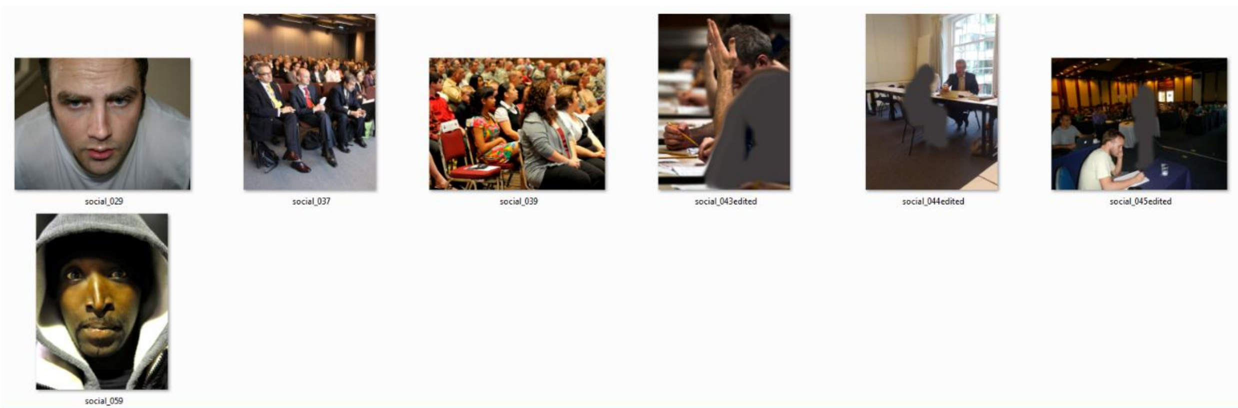





Canva40edited
Canva41edited
Canva42edited
Canva43edited
Canva44edited
Canva47edited
Canva48edited
Canva49edited
Canva50edited
Canva51edited
Canva54edited
Canva55edited
Canva56edited
Canva58edited
Canva59edited
Canva62edited
Canva63edited
Canva64edited
Canva67edited
Canva68edited
Canva69edited
Canva73edited
Canva75edited
Canva76edited


Canva78edited
Canva79
Faces_142_h
Faces_234_h
Faces_300_h
neutral_022
neutral_101
neutral_132
neutral_144
neutral_170
neutral_172
neutral_197
neutral_200
neutral_221
neutral_282
neutral_284
neutral_293
neutral_296


neutral_323
neutral_477
neutral_496
neutral_497
neutral_498
neutral_499
neutral_566
neutral_572
neutral_574
neutral_606
neutral_654
neutral_706
social_002
social_004
social_009edited
social_010
social_023
social_024edited


social_029
social_037
social_039
social_043edited
social_044edited
social_045edited
social_059

Appendix B: Self Report Questionnaires

Appendix B.1 Outcome Probability Questionnaire

Please rate how likely it is that the following outcomes will happen to you in a public speaking situation within the next year. Use the "0-8" scale below to indicate your answer, with "0" indicating that the outcome listed is not at all likely and "8" indicating the outcome listed is extremely likely.

	Not at all likely								Extremely likely
	0	1	2	3	4	5	6	7	8
You will feel embarrassed by something you did	○	○	○	○	○	○	○	○	○
You will sound dumb while talking to others	○	○	○	○	○	○	○	○	○
You will feel flustered in front of others	○	○	○	○	○	○	○	○	○
People will think that you are boring	○	○	○	○	○	○	○	○	○
At a party, others will notice that you are nervous	○	○	○	○	○	○	○	○	○
During a job interview or evaluation, you will freeze	○	○	○	○	○	○	○	○	○
While you are talking with several people, one of them will leave	○	○	○	○	○	○	○	○	○
You will be ignored by someone you know	○	○	○	○	○	○	○	○	○
You will do something foolish in public	○	○	○	○	○	○	○	○	○
You will fail to accomplish an important goal	○	○	○	○	○	○	○	○	○
You will fail to cope in your day-to-day living	○	○	○	○	○	○	○	○	○
You will be unexpectedly called in to see your supervisor at work	○	○	○	○	○	○	○	○	○

Appendix B.2 Outcome Cost Questionnaire

Outcome Cost Questionnaire

Please rate how bad or distressing the following outcomes would be for you if they were to occur in a social situation? Use the "0-8" scale below to indicate your answer, with "0" indicating that the outcome listed would be not at all distressing and "8" indicating the outcome listed would be extremely distressing.

	0-Not at all distressing	1	2	3	4	5	6	7	8-Extremely distressing
1.You were embarrassed by something you did	○	○	○	○	○	○	○	○	○
2.You sounded dumb to others	○	○	○	○	○	○	○	○	○
3.You felt flustered in front of others	○	○	○	○	○	○	○	○	○
4.People thought that you were boring	○	○	○	○	○	○	○	○	○
	0-Not at all distressing	1	2	3	4	5	6	7	8-Extremely distressing
5.At a party, others noticed that you were nervous	○	○	○	○	○	○	○	○	○
6.During a job interview or evaluation, you froze	○	○	○	○	○	○	○	○	○
7.While you were talking with several people, one of them left	○	○	○	○	○	○	○	○	○
8.You were ignored by someone you knew	○	○	○	○	○	○	○	○	○
	0-Not at all distressing	1	2	3	4	5	6	7	8-Extremely distressing
9. You did something foolish in public	○	○	○	○	○	○	○	○	○
10.You failed to accomplish an important goal	○	○	○	○	○	○	○	○	○
11.You failed to cope in your day-to-day living	○	○	○	○	○	○	○	○	○
12.You were unexpectedly called in to see your supervisor at work	○	○	○	○	○	○	○	○	○

Appendix B.3 Liebowitz Social Anxiety Scale

Please read each situation carefully and provide an answer to each column for that situation. The first column refers to how anxious or fearful you feel in the situation. The second refers to how often you avoid the situation. If you come across a situation that you ordinarily do not experience, we ask that you imagine "what if you were faced with that situation," and then, rate the degree to which you would fear this hypothetical situation and how often you would tend to avoid it. Please base your ratings on the way that the situations have affected you in the last week.

		Fear or Anxiety 0 = None 1 = Mild 2 = Moderate 3 = Severe	Avoidance 0 = Never (0%) 1 = Occasionally (1-33%) 2 = Often (33-67%) 3 = Usually (67-100%)
1.	Telephoning in public (P)		
2.	Participating in small groups (P)		
3.	Eating in public places (P)		
4.	Drinking with others in public places (P)		
5.	Talking to people in authority (S)		
6.	Acting, performing or giving a talk in front of an audience (P)		
7.	Going to a party (S)		
8.	Working while being observed (P)		
9.	Writing while being observed (P)		
10.	Calling someone you don’t know very well (S)		
11.	Talking with people you don’t know very well (S)		
12.	Meeting strangers (S)		
13.	Urinating n a public bathroom (P)		
14.	Entering a room when others are already seated (P)		

15.	Being the center of attention (S)		
16.	Speaking up at a meeting (P)		
17.	Taking a test (P)		
18.	Expressing a disagreement or disapproval to people you don't know very well (S)		
19.	Looking at people you don't know very well in the eyes (S)		
20.	Giving a report to a group (P)		
21.	Trying to pick up someone (P)		
22.	Returning goods to a store (S)		
23.	Giving a party (S)		
24.	Resisting a high pressure salesperson (S)		

Appendix B.4 State-Trait Anxiety Inventory

State-Trait Anxiety Inventory
STAI Form Y-1

Name…………………………………………….Date………….Age......... Sex: Male

DIRECTIONS: A number of statements which people have used to describe themselves are given below. Read each statement and then write the number in the blank at the end of the statement that indicates how you feel right now, that is, at this moment. There is no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe your present feelings best

S. No.		Not at all	Some What	Moderately so	Very much so
1.	I feel calm	1	2	3	4
2.	I feel secure	1	2	3	4
3.	I am tense	1	2	3	4
4.	I feel Strained	1	2	3	4
5.	I feel at ease	1	2	3	4
6.	I feel upset	1	2	3	4
7.	I am presently worrying over possible misfortunes	1	2	3	4
8.	I feel satisfied	1	2	3	4
9.	I feel frightened	1	2	3	4
10.	I feel comfortable	1	2	3	4
11	I feel self confident	1	2	3	4
12.	I feel nervous	1	2	3	4
13.	I am Jittery	1	2	3	4
14.	I feel indecisive	1	2	3	4
15.	I am relaxed	1	2	3	4
16.	I feel content	1	2	3	4
17.	I am worried	1	2	3	4
18.	I feel confused	1	2	3	4
19.	I feel steady	1	2	3	4
20.	I feel pleasant	1	2	3	4

Appendix B.5 Depression Anxiety Stress Scale

DASS21

Name: Date:

Please read each statement and circle a number 0, 1, 2 or 3 which indicates how much the statement applied to you **over the past week**. There are no right or wrong answers. Do not spend too much time on any statement.

The rating scale is as follows:

0 Did not apply to me at all
1 Applied to me to some degree, or some of the time
2 Applied to me to a considerable degree or a good part of time
3 Applied to me very much or most of the time

1 (s)	I found it hard to wind down	0	1	2	3
2 (a)	I was aware of dryness of my mouth	0	1	2	3
3 (d)	I couldn't seem to experience any positive feeling at all	0	1	2	3
4 (a)	I experienced breathing difficulty (e.g. excessively rapid breathing, breathlessness in the absence of physical exertion)	0	1	2	3
5 (d)	I found it difficult to work up the initiative to do things	0	1	2	3
6 (s)	I tended to over-react to situations	0	1	2	3
7 (a)	I experienced trembling (e.g. in the hands)	0	1	2	3
8 (s)	I felt that I was using a lot of nervous energy	0	1	2	3
9 (a)	I was worried about situations in which I might panic and make a fool of myself	0	1	2	3
10 (d)	I felt that I had nothing to look forward to	0	1	2	3
11 (s)	I found myself getting agitated	0	1	2	3
12 (s)	I found it difficult to relax	0	1	2	3
13 (d)	I felt down-hearted and blue	0	1	2	3
14 (s)	I was intolerant of anything that kept me from getting on with what I was doing	0	1	2	3
15 (a)	I felt I was close to panic	0	1	2	3
16 (d)	I was unable to become enthusiastic about anything	0	1	2	3
17 (d)	I felt I wasn't worth much as a person	0	1	2	3
18 (s)	I felt that I was rather touchy	0	1	2	3
19 (a)	I was aware of the action of my heart in the absence of physical exertion (e.g. sense of heart rate increase, heart missing a beat)	0	1	2	3
20 (a)	I felt scared without any good reason	0	1	2	3
21 (d)	I felt that life was meaningless	0	1	2	3

Appendix B.6 Temple Presence Inventory

SPATIAL PRESENCE:

VAR NAME	LDNG	ITEM
PLACE	.89	How much did it seem as if the objects and people you saw/heard had come to the place you were? (Not at all - Very much [7 points])
TOUCH	.88	How much did it seem as if you could reach out and touch the objects or people you saw/heard? (Not at all - Very much [7 points])
OBJECT	.83	How often when an object seemed to be headed toward you did you want to move to get out of its way? (Never - Always [7 points])
BETHERE	.79	To what extent did you experience a sense of being there inside the environment you saw/heard? (Not at all – Very much [7 points])
SNDLOCAL	.72	To what extent did it seem that sounds came from specific different locations? (Not at all - Very much [7 points])
TOUCHSMG	.68	How often did you want to or try to touch something you saw/heard? (Never - Always [7 points])
WINDOW	.58	Did the experience seem more like looking at the events/people on a movie screen or more like looking at the events/people through a window? (Like a movie screen – Like a window [7 points])

EIGENVALUE:	4.19
VARIANCE EXPLAINED:	59.85
STANDARDIZED CRONBACH'S ALPHA:	.91

SOCIAL PRESENCE - ACTOR W/I MEDIUM (PARASOCIAL INTERACTION):

VAR NAME	LDNG	ITEM
PPLSEEU	.83	How often did you have the sensation that people you saw/heard could also see/hear you? (Never - Always [7 points])
INTERACT	.82	To what extent did you feel you could interact with the person or people you saw/heard? (None - Very much [7 points])
LEFTPLCE	.79	How much did it seem as if you and the people you saw/heard both left the places where you were and went to a new place? (Not at all - Very much [7 points])
TOGETHER	.78	How much did it seem as if you and the people you saw/heard were together in the same place? (Not at all - Very much [7 points])

TALKTOYU	.77	How often did it feel as if someone you saw/heard in the environment was talking directly to you? (Never - Always [7 points])
EYECONT	.68	How often did you want to or did you make eye-contact with someone you saw/heard? (Never - Always [7 points])
CONTRINT	.67	Seeing and hearing a person through a medium constitutes an interaction with him or her. How much control over the interaction with the person or people you saw/heard did you feel you had? (None - Very much [7 points])

EIGENVALUE:	4.08
VARIANCE EXPLAINED:	58.24
STANDARDIZED CRONBACH'S ALPHA:	.90

SOCIAL PRESENCE - PASSIVE INTERPERSONAL:

VAR NAME	LDNG	ITEM
FACEEXPR	.89	During the media experience how well were you able to observe the facial expressions of the people you saw/heard? (Not well - Very well [7 points])
TONEVOIC	.85	During the media experience how well were you able to observe the changes in tone of voice of the people you saw/heard? (Not well - Very well [7 points])
STYLDRES	.79	During the media experience how well were you able to observe the style of dress of the people you saw/heard? (Not well - Very well [7 points])
BODYLANG	.69	During the media experience how well were you able to observe the body language of the people you saw/heard? (Not well - Very well [7 points])

EIGENVALUE:	2.61
VARIANCE EXPLAINED:	65.27
STANDARDIZED CRONBACH'S ALPHA:	.88

SOCIAL PRESENCE - ACTIVE INTERPERSONAL:

VAR NAME	LDNG	ITEM
MKSOUND	.84	How often did you make a sound out loud (e.g. laugh or speak) in response to someone you saw/heard in the media environment? (Never - Always [7 points])
SMILE	.73	How often did you smile in response to someone you saw/heard in the media environment? (Never - Always [7 points])

SPEAK	.61	How often did you want to or did you speak to a person you saw/heard in the media environment? (Never - Always [7 points])

EIGENVALUE:	1.61
VARIANCE EXPLAINED:	53.51
STANDARDIZED CRONBACH'S ALPHA:	.77

ENGAGEMENT (MENTAL IMMERSION):

VAR NAME	LDNG	ITEM
MENTALIM	.86	To what extent did you feel mentally immersed in the experience? (Not at all - Very much [7 points])
INVOLVNG	.80	How involving was the experience? (Not at all - Very much [7 points])
SENSEENG	.79	How completely were your senses engaged? (Not at all - Very much [7 points])
SENSREAL	.79	To what extent did you experience a sensation of reality? (Not at all - Very much [7 points])
EXCITING	.75	How relaxing or exciting was the experience? (Very relaxing - Very exciting [7 points])
ENGSTORY	.65	How engaging was the story? (Not at all - Very much [7 points])

EIGENVALUE:	3.61
VARIANCE EXPLAINED:	60.10
STANDARDIZED CRONBACH'S ALPHA:	.90

SOCIAL RICHNESS:

VAR NAME	LDNG	ITEM
REMOTE	.85	Please circle the number that best describes your evaluation of the media experience: Remote - Immediate (7 points)
UNEMOTNL	.83	Please circle the number that best describes your evaluation of the media experience: Unemotional - Emotional (7 points)
UNRESPON	.82	Please circle the number that best describes your evaluation of the media experience: Unresponsive - Responsive (7 points)
DEAD	.80	Please circle the number that best describes your evaluation of the media experience: Dead - Lively (7 points)
IMPERSNL	.78	Please circle the number that best describes your evaluation of the media experience: Impersonal - Personal (7 points)

INSENSTV	.76	Please circle the number that best describes your evaluation of the media experience: Insensitive - Sensitive (7 points)
UNSOCBLE	.76	Please circle the number that best describes your evaluation of the media experience: Unsociable - Sociable (7 points)

EIGENVALUE:	4.48
VARIANCE EXPLAINED:	63.99
STANDARDIZED CRONBACH'S ALPHA:	.93

SOCIAL REALISM:

VAR NAME	LDNG	ITEM
WOULDOCR	.87	The events I saw/heard <u>would</u> occur in the real world (Strongly disagree - Strongly agree [7 points])
COULDOCR	.76	The events I saw/heard <u>could</u> occur in the real world (Strongly disagree - Strongly agree [7 points])
OCRWORLD	.53	The way in which the events I saw/heard occurred is a lot like the way they occur in the real world (Strongly disagree - Strongly agree [7 points])

EIGENVALUE:	1.60
VARIANCE EXPLAINED:	53.34
STANDARDIZED CRONBACH'S ALPHA:	.75

PERCEPTUAL REALISM:

VAR NAME	LDNG	ITEM
FEELLIKE	.80	Overall how much did touching the things and people in the environment you saw/heard feel like it would if you had experienced them directly? (Not at all - Very much [7 points])
TEMPERAT	.74	How much did the heat or coolness (temperature) of the environment you saw/heard feel like it would if you had experienced it directly? (Not at all - Very much [7 points])
SMELLIKE	.70	Overall, how much did the things and people in the environment you saw/heard smell like they would had you experienced them directly? (Not at all - Very much [7 points])
LOOKLIKE	-	Overall, how much did the things and people in the environment you saw/heard look they would if you had experience them directly (Not at all - Very much [7 points])
SOUNDLKE	-	Overall, how much did the things and people in the environment you saw/heard sound like they would if you had experienced them directly? (Not at all - Very much [7 points])

EIGENVALUE:	1.67
VARIANCE EXPLAINED:	55.71
STANDARDIZED CRONBACH'S ALPHA:	.79

Appendix B.7 Simulator Sickness Questionnaire

Check the items that apply to you at this time:

Device	**None**	**Slight**	**Moderate**	**Severe**
General discomfort				
Fatigue				
Headache				
Eye strain				
Difficulty focusing				
Increased salivation				
Sweating				
Nausea				
Difficulty concentrating				
"Fullness of the head"				
Blurred vision				
Dizzy (eyes open)				
Dizzy (eyes closed)				
Vertigo				
Stomach awareness				
Burping				
Other (describe):				

"Stomach awareness" indicates a feeling of discomfort which is just short of nausea.

Appendix B.8 Subtle Avoidance Frequency Examination

Please rate how frequently you engaged in the following behaviors during the virtual reality speech.

1-Never 2 3 4 5-Always

Avoided eye contact

Remained silent

Spoke softly

Spoke in soft sentences

Kept still to avoid drawing attention to yourself

Tried to keep tight control over your behavior

Blanked out or switched off mentally

Spent time thinking of good excuses for escaping (stopping the speech)

Avoided speaking certain words (e.g. to prevent stuttering)

Rehearsed sentences in your mind

Was reserved about what you said

Think or say 'I'm not usually like this'

Asked or wanted to ask about your performance

Imagined you were somewhere else

Held your stop card or hands tightly

Accounted for poor performance by saying you didn't have adequate time to prepare

Held your arms still

Tried to think about other things

Avoided pauses in your speech

Was there anything in the virtual environment that made you feel particularly anxious?

____

If you had trouble making eye-contact, where were you looking?

____

Comments: ____

Appendix B.9 Study Feedback Questionnaire

We are interested in getting feedback about your reactions to the different images you saw in the computer task you completed before the virtual speech.

	1-Not at all difficult	2	3	4	5-Extremely difficult
Overall, how difficult was it for you to imagine yourself in the social situations shown in the pictures (those **without** a blurred figure)?	○	○	○	○	○

	1-Not at all difficult	2	3	4	5-Extremely difficult
Overall, how difficult was it for you to imagine yourself in the social situations shown in the pictures (those **with** a blurred figure)?	○	○	○	○	○

	1-Not at all difficult	2	3	4	5-Extremely difficult
Overall, how difficult was it for you to tell what was happening in the social images?	○	○	○	○	○

	1-Not at all difficult	2	3	4	5-Extremely difficult
Overall, how difficult was it for you to focus on or pay attention to the images?	○	○	○	○	○

	1- Not very imaginative	2	3	4	5-Extremely imaginative
How imaginative are you?	○	○	○	○	○

Were you able to tell the gender, approximate age, or race of any of the blurry figures?

Did any of the images make you feel anxious?

- ○ Yes
- ○ No

Which images made you feel the most anxious?

What were you afraid might happen in response to the images that made you feel anxious?

Which images made you feel the least anxious?

For the task where the images flashed very quickly, were you able to tell what the images were showing?

- Yes, nearly all of them
- Sometimes
- No, nearly none of them

Did you use any strategies when answering questions in response to the images? If so, what was your strategy?

Did you notice you pressed the spacebar faster or slower when viewing certain images? If so, which ones?

Please share any general feedback you have about your experience with the images: