Summer 8-10-2021

Does Math Anxiety Moderate the Effect of an Online Arousal Reappraisal Intervention on Math Performance?

Zachary Taber

ACCEPTANCE

This dissertation, DOES MATH ANXIETY MODERATE THE EFFECT OF AN ONLINE AROUSAL REAPPRAISAL INTERVENTION ON MATH PERFORMANCE?, by ZACHARY TABER, was prepared under the direction of the candidate's Dissertation Advisory Committee. It is accepted by the committee members in partial fulfillment of the requirements for the degree, Doctor of Philosophy, in the College of Education & Human Development, Georgia State University.

The Dissertation Advisory Committee and the student's Department Chairperson, as representatives of the faculty, certify that this dissertation has met all standards of excellence and scholarship as determined by the faculty.

Kenneth Rice, Ph.D.
Committee Chair

Jeff Ashby, Ph.D.
Committee Member

Maggie Renken, Ph.D.
Committee Member

Page Anderson, Ph.D.
Committee Member

June 30th, 2021
Date

Brian Dew, Ph.D.
Chairperson, Department of Counseling and Psychological Services

Paul A. Alberto, Ph.D.
Dean, College of Education & Human Development

AUTHOR'S STATEMENT

By presenting this dissertation as a partial fulfillment of the requirements for the advanced degree from Georgia State University, I agree that the library of Georgia State University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote, to copy from, or to publish this dissertation may be granted by the professor under whose direction it was written, by the College of Education & Human Development's Director of Graduate Studies, or by me. Such quoting, copying, or publishing must be solely for scholarly purposes and will not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential financial gain will not be allowed without my written permission.

Zachary Taber

NOTICE TO BORROWERS

All dissertations deposited in the Georgia State University library must be used in accordance with the stipulations prescribed by the author in the preceding statement. The author of this dissertation is:

Zachary Brenner Taber

Department of Counseling and Psychological Services

College of Education & Human Development

Georgia State University

The director of this dissertation is:

Kenneth Rice

Department of Counseling and Psychological Services

College of Education & Human Development

Georgia State University

Atlanta, GA 30303

CURRICULUM VITAE

ZACHARY B. TABER

554 Elbridge Dr. NW
Atlanta, GA 30318
631-805-6693

EDUCATION

2016-present	**Ph.D. in progress, Georgia State University** Counseling Psychology
2016	**M.A. University of Denver** Sport and Performance Psychology
2008	**B.A. University of Maryland, College Park** Dual Majors: Government and Politics, History

HONORS, AWARDS, & SCHOLARSHIPS

2020	Kenneth and Mary Metheny Scholarship, Georgia State University
2016-present	Dean's Research Doctoral Fellowship, Georgia State University
2014-2016	Professional Psychology Scholarship, University of Denver
2010-2011	Eli J. Segal AmeriCorps Education Award
2010	President's Volunteer Service Award
2006-2007	Long Island Area Council Student Activity Fund Recipient
2005-2008	Dean's List, College of Behavioral and Social Sciences
2004-2008	President's Scholarship, University of Maryland, College Park

PEER REVIEWED JOURNAL ARTICLES

Taber, Z., Rice, K.G., Carroll, J., Hong, J. (2019). *Does perfectionism moderate the effects of values affirmations on stress and performance?* Manuscript in preparation.

Liu, Y., Rice, K.G., **Taber, Z.**, Dennis, B., & Rice, F. (2019). Perfectionism and burnout in R&D teams. *Academy of Management Proceedings, 2019*(1), 17942. doi:10.5465/ambpp.2019.174

Rice, K.G., & **Taber, Z.** (2019). Measurement invariance and latent profiles of perfectionism in clients and non-clients. *Journal of Counseling Psychology, 66*(2), 2010-223. doi:10.1037/cou0000326

BOOK CHAPTERS

Rice, K. G, & **Taber, Z.** (2018) Perfectionism. In S. Pfieffer (Ed.), *Handbook of giftedness in children*: *Psychoeducational theory, research, and best practices* (pp. 227-254). Springer.

PRESENTATIONS

Hong, J., **Taber, Z.**, Aiello, M., Dennis, B.S., Rice, K.G. (August, 8, 2020). *The role of perfectionism between self-compassion and social emotional loneliness.* Poster presented at the APA Convention. Washington, DC, USA.

Taber, Z., Hong, J., Aiello, M., Dennis, B.S., Rice, K.G. (February, 28, 2020). *Mindfulness and perfectionism: Investigating associations between mindfulness facets, perfectionism factors, well-being, and psychological distress.* Poster presented at the SPSP Annual Meeting. New Orleans, LA, USA.

Taber, Z., Hong, J., Aiello, M., Dennis, B.S., Rice, K.G. (August, 10, 2019). *Relationships between perfectionism, mindfulness, and social-emotional loneliness in college students*. Poster presented at the APA Convention. Chicago, IL, USA.

Dennis, B.S., Hong, J., **Taber, Z.,** & Rice, K.G. (August, 11, 2019). *Race and income: Moderators of perfectionism and stress*. Poster presented at the APA Convention. Chicago, IL, USA.

Rice, K.G., & **Taber, Z.** (August 11, 2018). *Perfectionism in clients and nonclients: Classification and profiles of psychological distress.* Paper presented at the APA Convention. San Francisco, CA, USA.

Taber, Z., & Rice, K. G. (March 3, 2018). *Effects of a brief mindfulness intervention for maladaptive perfectionists*. Poster presented at the SPSP Annual Meeting. Atlanta, GA, USA.

Taber, Z., Carrol, J. L., Kelly, V. A., Hong, J., & Rice, K. G. (May 27, 2017). *Does perfectionism moderate the effects of values affirmations on stress and performance?* Poster presented at the APS Annual Convention, Boston, MA.

Taber, Z., Goffman, D., & Gearity, B. (November 4, 2016). *A trioethnography of existential post-structural sport psychology and sport coaching praxis.* Paper presented at the NASSS Annual Meeting. Tampa, FL, USA.

Ramesh, R., McDonald, J., & **Taber, Z.** (April 9, 2016). *Exploring the mental strengths and challenges of performing artists*. Paper presented at the AASP Southwest Regional Conference. Denver, CO, USA.

Goffman, D., & **Taber, Z.** (March 28, 2015). *The pervasive nature of the judging mind: Buddhist perspectives on nonjudgmental awareness and the performance experience* [.5 hr colloquium]. AASP Southwest Regional Conference. Denver, CO, USA.

DOES MATH ANXIETY MODERATE THE EFFECT OF AN ONLINE AROUSAL REAPPRAISAL INTERVENTION ON MATH PERFORMANCE?

by

ZACHARY B. TABER

Under the Direction of Kenneth G. Rice, Ph.D.

ABSTRACT

Math anxiety has been consistently associated with lowered math performance and achievement regardless of skills and abilities. These decrements in math performance and achievement and accompanying avoidance of opportunities to improve math skills can have outsize influence on the school and career path of college students. In the following chapter, we conducted a systematic review and meta-analysis of the last 30 years of intervention literature for math anxiety, focusing on math performance as an outcome of interest. For the nine studies that fit the inclusion criteria we summarize the effect sizes, though moderator analyses were not conducted due to low heterogeneity between studies and lack of power to detect effects. We discuss implications and directions for future research on math anxiety. In the following study, noting the importance of appraisal processes in math anxiety, we examined the effect of a brief online arousal reappraisal intervention on math performance and math anxiety. We also investigate whether pre-intervention math anxiety moderates the effect of arousal reappraisal on math performance, hypothesizing that highly math anxious individuals will show greater conditional effects than individuals with low or moderate levels of math anxiety. Results supported the moderation of the intervention effect by pre-intervention math anxiety, but in the opposite direction of the hypothesized conditional effect. More specifically, students in the intervention group with moderate or lower levels of pre-intervention math anxiety scored significantly higher than students with similar levels of pre-intervention math anxiety in the control group, while there were no significant group differences when students had higher levels of math anxiety.

INDEX WORDS: Math Anxiety, Arousal Reappraisal, Intervention, Treatment, Moderation.

DOES MATH ANXIETY MODERATE THE EFFECT OF AN ONLINE AROUSAL REAPPRAISAL INTERVENTION ON MATH PERFORMANCE?

by

ZACHARY B. TABER

A Dissertation

Presented in Partial Fulfillment of Requirements for the

Degree of

Doctor of Philosophy

in

Counseling Psychology

in

The College of Education

Georgia State University

Atlanta, GA

2021

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Meta-Analysis of the Effect of Psychological Interventions for Math Anxiety on Math Performance: A Review of the Past 30 Years

For years, U.S. educational policy-makers have sought to promote science, technology, engineering, and mathematics (STEM) literacy, skills, and education to supply the needs of the U.S. economy for an increasingly technically trained workforce (National Science & Technology Council Committee on STEM Education, 2018). Numeracy skills are central to this project, though students in the U.S. remain below the global mean in math performance despite executive branch commitments to investing in STEM education (Organization for Economic Co-operation & Development, 2018; Handelsman & Smith, 2016). While pedagogical and systemic issues are often the main focus of these efforts, and rightly so, addressing attitudinal and affective factors can also help promote the development and practice of numeracy skills in children, adolescents, and adults. One such affective factor that is the focus of this review is math anxiety.

Math Anxiety

Math anxiety "involves feelings of tension and anxiety that interfere with the manipulation of numbers and the solving of mathematical problems in a wide variety of ordinary life and academic situations" (Richardson & Suinn, 1972, p. 551). These feelings are somewhat common in the U.S., with estimates of prevalence in college populations, with whom most studies on math anxiety have been conducted, ranging from 11% (Richardson & Suinn, 1972) to 66% of students enrolled in math classes (Betz, 1978).

Research on math anxiety began in earnest with the development and validation of the 98-item Math Anxiety Rating Scale (MARS) in a paper by Richardson and Suinn (1972), which also found negative associations between MARS scores and math performance. The negative associations between math anxiety, math achievement, and math performance have been consistent across decades of studies. A meta-analysis by Hembree (1990) synthesizing the state of math anxiety research in 1990, reported summary correlations between math anxiety, math course grades, and measures of math aptitude or achievement in the -0.25 to -0.40 range. Also summarizing the effect of math anxiety on math performance for 13 studies of collegiate samples, Hembree reported that on average, individuals with high math anxiety (HMA) scored 0.61 *SD*s lower than individuals with low math anxiety (LMA) on math tests.

Operationalization of Math Anxiety

In the history of math anxiety research, HMA and LMA groups have been operationalized in a number of ways. Firstly, the unwieldy nature of the original 98-item version of the MARS led to numerous attempts to create more time-efficient measures of the construct, many of which use items from the original MARS. These include: a 30 item brief version of the Math Anxiety Rating Scale (MARS-30; Suinn & Winston, 2003); the 24-item Revised Math Anxiety Rating Scale (MARS-R; Plake & Parker) which was further shortened to 12 items in a factor analytic study by Hopko (2003); the Abbreviated Math Anxiety Scale (AMAS; Hopko, Mahadean, Bare, & Hunt, 2003), another factor analytic attempt to shorten the MARS-R that resulted in a 9-item measure validated in a large-scale (*N* = 1,239) study; and the 25-item short version of the Math Anxiety Rating Scale (sMARS; Alexander & Martray, 1989). The majority of psychometric studies (e.g. Rounds & Hendel, 1980; Alexander & Cobb, 1987; Hopko et al., 2003) of the MARS and its brevity-oriented progeny have found that a two-factor structure best fits the data (Pletzer, Wood, Scherndl, Kerschbaum, & Nuerk, 2016), though interpretations regarding what is captured by those factors and which factors are most salient to measurement of math anxiety have changed over time as measures have been refined.

Rounds and Hendel (1980) first reported on the two-factor structure of the MARS, identifying the factors as Mathematics Test Anxiety, or anxiety related to the demands of mathematics courses or evaluations (e.g. "Taking an examination (quiz) in a math course"), and Numerical Anxiety, or anxiety in everyday situations requiring partial arithmetic and numeracy skills (e.g. "Totaling up a dinner bill that you think overcharged you"). In their development of the MARS-R, Plake and Parker (1982) also settled on a two-factor structure, renaming the factors Learning Mathematics Anxiety, or anxiety associated with the activity or process of learning math (e.g. "Listening to a lecture in a math class"), and Mathematics Evaluation Anxiety, or anxiety associated with tests, quizzes, or other evaluations in math or statistics (e.g. "Being given a 'pop' quiz in a math class"). As Plake and Parker's (1982) original study was likely underpowered and methodologically flawed, Hopko (2003) submitted the MARS-R to a more rigorous test of construct validity, ultimately eliminating 12 items to arrive at a model that fit the two factor Learning Mathematics Anxiety and Mathematics Evaluation Anxiety structure well in both the validation sample and a replication sample. As Hopko (2003) explains, his revision of the MARS-R eliminated many items that were seemingly less central to the experience of doing math. The majority of the retained items directly relate to anxiety experienced in the process of academic math performance or calculation, or in the anticipation of that process. As Pletzer et al. (2016) notes, qualitatively, this revision seemingly aligns the two factors of the MARS-R more with the experiences captured by the original MARS Mathematics Test Anxiety factor than the Numerical Anxiety factor. Moving on from the MARS family of measures, a final measure of math anxiety to highlight is the Fennema-Sherman Mathematics Attitude Scales (Fennema & Sherman, 1976), particularly the Mathematics Anxiety Scale (MAS) included as a 12-item subscale in that measure. The MAS was designed to measure distressing arousal and negative affect (e.g. anxiety or dread) associated with doing math (e.g. "Mathematics usually makes me feel uncomfortable and nervous").

These measures have been utilized in a variety of mathematic procedures to identify HMA and LMA individuals for research, assessment, and intervention, with no particular consensus (Ramirez et al., 2018). Some techniques used by researchers include labeling individuals at ± 1 *SD* of the sample mean as HMA or LMA respectively (e.g. Maloney et al., 2015), using a median split with HMA above the sample median and LMA below the median (e.g. Brunyé et al., 2013), identifying the upper and lower quartiles of the sample as HMA and LMA (e.g. Henslee & Klein, 2017), respectively, or dividing the sample into HMA (+1 *SD*), LMA (-1 *SD*), and medium (mean) math anxiety groups based on the sample mean and *SD* (Ashcraft & Kirk, 2001).

Math Anxiety and Math Performance

Given the significant effect of math anxiety on math performance and intent to pursue further math education (Hembree, 1990), interventions to reduce math anxiety or reduce its deleterious effects on math performance are an important research priority. Hembree's meta-analysis, the only one to address interventions for math anxiety, synthesized 73 effect sizes, concluding that behavioral intervention such as systematic desensitization and others ($d = 0.60$), cognitive restructuring interventions ($d = 0.32$), and cognitive-behavioral interventions ($d = 0.50$), were the only interventions with summary effect sizes that were significantly different from null effects. In the approximately 30 years since his review was published, there have been significant developments in the science and practice of psychological interventions. As such, an update on the last 30 years of math anxiety intervention literature is warranted.

The Present Study

The purpose of the current study was to provide a quantitative meta-analytical review of the last 30 years of published research examining the effects of psychological interventions for math anxiety on math performance. Math performance was selected as the outcome variable of choice because, despite the robust negative correlation between math anxiety and math performance, some HMA individuals have shown the ability to use cognitive and emotion regulation strategies to perform well on math tasks despite high levels of math anxiety (Lyons & Beilock, 2012). As such, interventions may work through mechanisms other than math anxiety reduction to positively affect math performance.

Methods

Literature Search

An electronic database search was conducted using PsycINFO and ERIC with the following all text search terms: (intervention OR treatment OR Therapy OR prevention) AND (arithmetic anxiety OR math* anxiety OR calculation anxiety OR statistics anxiety). The search identified 331 studies, with 17 selected for further review due to abstracts fitting inclusion criteria. Of those studies, 8 were excluded (see Figure 1.1). Two studies were excluded because they did not measure math performance, three studies were excluded because they involved curriculum or tutoring rather than psychological interventions, one study was excluded because the intervention was not a treatment for math anxiety, and two studies were excluded because effect sizes could not be calculated from the data reported and the authors did not respond to requests for additional data.

Inclusion/ Exclusion Criteria

Studies were included in the meta-analysis if they (a) investigated psychological (as opposed to tutoring or curricular) interventions that specifically targeted math anxiety; (b) measured math performance; (c) included an effect size or sufficient information for computation or estimation of an effect size; (d) were written in English; (e) were a published journal article or pre-print; and (f) were conducted after Hembree's meta-analysis in 1990. The implementation of the criteria resulted in the final inclusion of 9 studies reporting 12 effect sizes quantifying group differences in math performance following interventions for math anxiety.

Recorded Variables

Each included study was coded by the primary author for the following variables: study design (within participants or randomized-controlled trial), sample math anxiety (HMA only, HMA and LMA, or undifferentiated), intervention orientation (cognitive, mindfulness, expressive writing, values affirmation), intervention modality (online, individual, group, written instructions, or written and online instructions), total intervention time in minutes, mean age of participants, participant grade level, predominant gender of participants, predominant race/ethnicity, country where study was conducted, math anxiety measure, and math performance task.

Study participants were primarily female postsecondary students aged 18 to 25. Five out of nine studies did not report the race or ethnicity of the participants. Of the three studies that reported race or ethnicity, two included predominantly White samples, with Jamieson et al. (2016) reporting on a majority Black or African-American sample. All studies were conducted in the U.S. Three out of nine studies measured math anxiety using the MARS, with two studies utilizing the MARS-30 and one study each administering the AMAS, sMARS, and the MAS, respectively. There was little overlap between studies in regards to math outcome measures, with two studies using algebra tasks and two studies using mental arithmetic with carrying operations, which have been shown to require greater working memory resources (Lyons & Beilock, 2012), and one study using computer-administered multistep arithmetic problems. Two studies used naturalistic methods, with students' in-class exams serving as the outcome measure, while two studies used validated measures of math achievement, the Wide Range Achievement Test 3 (Wilkinson, 1993) and the Differential Aptitude Tests (Bennett, Seashore, 7 Wesman, 1981). A third study used a validated abbreviated numeracy scale consisting of word problems testing probabilistic skills in combination with an objective numeracy scale of symbolic arithmetic problems that had been used in a previous study. Cognitive interventions were the most popular, with four primarily cognitive and 3 primarily behavioral interventions tested. Additionally, one mindfulness focused breathing intervention and one Acceptance and Commitment Therapy intervention was evaluated, and three brief social-psychological interventions (one study each testing stress reappraisal, expressive writing, and values affirmations) contributed effect sizes to the analysis. Total intervention time ranged from 7 minutes for a brief writing exercise to 360 minutes of group therapy across 6 sessions (M = 106.88 minutes). For further details on study characteristics, see Table 1.1.

Analysis Procedures

A priori power analyses were conducted using the **dmetar** (Harrer, Cuijpers, Furukawa, & Ebert, 2019) package in R, which utilizes equations published by Borenstein, Hedges, Higgins, and Rothstein (2009). Random effects models were used to account for between-studies variance, as studies differed considerably on a number of variables including design and interventions tested (Borenstein et al., 2009). Power analyses were conducted assuming low (t^2 = 0.33) and moderate (t^2 = 0.67) levels of heterogeneity, respectively (Hedges & Pigott, 2001), as 0.33 is a ratio of between-studies variance to within-studies variance commonly found in the psychological literature (Schmidt, 1992). For studies meeting the inclusion criteria (k = 9), average treatment group and control group samples sizes across studies were calculated (*n* = 34 and *n* = 35, respectively) and included in the power analysis. Following Bloom, Hill, Rebeck, Black, & Lipsey's (2008) recommendation that performance gaps be used as effect-size benchmarks for academic achievement or performance interventions, we set the effect size for the power analysis at $d = 0.305$, or half of the math performance gap that Hembree (1990) found between high and low math anxious individuals. Assuming a *p*-value of .05, power was acceptable to detect the $d = 0.305$ effect given both low (power = 90.59%) and moderate (power = 83.24%) levels of heterogeneity.

The meta-analyses were conducted using the **metafor** package (Viechtbauer, 2010) in R. In deriving effect sizes and confidence intervals, random-effects models were used. Random-effects models assume variation in effect sizes between studies due to both sampling error and true random variance arising from differences between study procedures and settings. The analysis included studies with randomly controlled pre-test post-test designs, randomly controlled post-test only designs, and within-subjects pre-test post-test designs without control groups. Following recommendations by Lakens (2013) for calculating and synthesizing effect sizes across study designs, for randomly controlled studies, the pre-treatment means and standard deviations for pre- and post-intervention measurements of the control and treatment groups were extracted from each study. Hedges' *g*, a measure of the standardized mean difference between treatment and control groups, was calculated to correct for baseline differences between groups on outcome variables, as, when compared to Cohen's *d*, it provides an unbiased effect size estimate when samples sizes are small or vary within studies (Hedges, 1989). Rules of thumb for interpretation of Hedges' *g* are as follows: small, $d = 0.2$; medium, $d = 0.5$; large, $d = 0.8$ (Borenstein, Hedges, Higgins, & Rothstein, 2009). In the case of studies using within-subjects designs or lacking a control group (e.g. Zettle, 2003), Hedges' g_{av} was calculated, as it is a less biased estimate of the effect size than Cohen's d_{av}, and it is considered comparable to Hedges' *g* in a between-subjects design, therefore facilitating synthesis across study designs (Lakens, 2013). When studies consisted of post-test only controlled designs (e.g. Vance & Watson, 1994), Hedges' *g* was calculated using the mean difference between treatment and control-group means (Hoyt & Del Re, 2018). Three studies (Schneider & Nevid, 1993; Vance & Watson, 1994; Zettle, 2003) utilized two treatment groups in addition to a control group. In those cases, to ensure each study only contributed one effect size for math performance, effect sizes were aggregated into a single composite effect size for each study using the Borenstein, Hedges, Higgins, and Rothstein (2009) procedure included as the default in the R package, **MAd** (Del Re & Hoyt, 2014)**.**

Results

Sensitivity Analyses

Sensitivity analyses (Viechtbauer & Cheung, 2010) were conducted using the METAFOR package in R for the random effects model tested. Inspection of studentized residuals and covariance ratios indicated that Vance and Watson (1994) was likely an outlier (standardized studentized residual = 3.19, DFFTS = 0.86) and omission of the study would greatly decrease heterogeneity and improve the precision of the model ($\Delta I^2 = 24.70$, covariance ratio = 0.73), therefore Vance and Watson (1994) was removed from the analysis.

Publication Bias

To account for publication bias, we visually inspected a funnel plot of study effect sizes and standard errors, finding little evidence of publication bias. This is further supported by Egger's test of the intercept (Egger, Smith, Schneider, & Minder, 1997), which was non-significant ($p = 0.61$) at the recommended $p < 0.1$ criterion, and Begg and Mazumdar's (1994) rank correlation test for publication bias, which was also non-significant (p = 0.18). Rosenthal's *Fail-safe N* (Rosenthal, 1979) indicated that 12 studies with effect sizes of 0 would have to be added to the meta-analysis to render the summary effect non-significant at the $p < .05$ level.

Overall Effect Sizes

The synthesized effect sizes illustrating the effect of math anxiety interventions on math performance are reported in Figure 1.2 along with the study weights (also visually represented by the size of the squares on the plot). All effects but one were in the positive direction, ranging from -0.07 to 0.56, such that interventions resulted in increases in math performance across studies. Relative to the control groups, interventions resulted in a small estimated effect on math performance equal to 0.23 *SD*s, 95% CI [0.07, 0.39], roughly a third of the -0.61 *SD* math performance gap found between low and high math-anxious individuals in Hembree's (1990) meta-analysis. Heterogeneity in the random effects model was low (I^2 = 18.3%) and the chi-square test for heterogeneity was non-significant, Q = 2.47, df = 7, p = 0.93 (Higgins & Thompson, 2002; Higgins et al., 2003). With little heterogeneity to explain, moderator analyses were not conducted (Borenstein et al., 2009).

Discussion

This systematic review and meta-analysis aimed to investigate the effectiveness of psychological interventions for math anxiety over the past 30 years across a variety of modalities, with modest, but encouraging findings. Meta-analyses indicated small effects for math anxiety interventions across empirical studies for enhancing math performance. Due to limited power because of the small number of intervention studies conducted since Hembree (1990), analyses evaluated treatment and control groups with HMAs and LMAs combined, which foreclosed on opportunities to meta-analytically examine differential treatment effects for

HMA and LMA individuals. Several studies incorporated level of math anxiety as a factor in the analysis or included *M*s and *SD*s for treatment and control group subsamples by math anxiety group membership, which should be encouraged to facilitate investigation into whether treatment effects are conditional on levels of math anxiety. It is possible that treatment effects could be even greater for HMA individuals than for the HMA and LMA individuals combined, though the opposite may also be true.

Even more promising, many of the interventions required considerably less resources than treatments associated with other trait-level psychological problems, with treatments ranging from 8 minutes interacting with written prompts or 7 minutes of expressive writing at the least resource-intensive end of the spectrum, to 360 minutes of group therapy at the most, a dose of therapy that would be considered brief by mental health clinicians. Additionally, many of the shortest social-psychological interventions reported effects comparable or often greater than individual or group therapy approaches that require skilled therapists, meeting spaces, and significant time commitments. This suggests that math anxiety may not require the more resource-intensive approaches often used with anxiety disorders, such as individual or group therapy. Also, many of the social-psychological interventions used only written instructions, making intervention delivery through online platforms likely feasible and worth investigating. These findings are considerably different from those of Hembree (1990), who reported non-significant summary effects for low-resource investment classroom psychological interventions, compared to moderate to large summary effects for cognitive, behavioral, and cognitive-behavioral therapeutic interventions.

It must be stated that these interpretations are premature and based on a limited sample of studies. It is also important to note that due to resource limitations, we restricted the sample to published articles or pre-prints, excluding unpublished studies, research reports, dissertations, conference proceedings, and other so-called *grey* literature. Because studies with significant effects are more likely to be published, our decision to focus only on the published literature may have resulted in an upward bias in our reported effect size (see Borenstein, 2009). While the tests of publication bias and *Fail-safe N* support our conclusions that interventions for math anxiety are effective in improving math performance, the inclusion of grey literature in our meta-analysis would likely have provided a more precise summary estimate of that effect that is different from the one we report here.

Regarding treatment and intervention recommendations, Hembree (1990) largely declined to advance any in his article, concluding only that "treatment can restore the performance of formerly high-anxious students to the performance level associated with low mathematics anxiety" (p. 44). Researchers, educators, and clinicians must derive their own recommendations from his summary in the results section, that the behavioral treatment modalities of systematic desensitization, anxiety management training, and conditioned inhibition were "highly successful in reducing mathematics anxiety levels," while cognitive restructuring treatments "produced a moderate reduction" (p. 43). Unfortunately, because our sample included only 1 or 2 studies investigating each intervention, we can offer little more than Hembree (1990) in the way of treatment and intervention recommendations. Cognitive interventions faired relatively poorly in our meta-analysis, especially when compared with Hembree's findings. However, our small sample urges caution in drawing any conclusions except that these interventions have shown effectiveness in the past and remain worthy of further study. While the studies detailing the effects of brief interventions such as stress reappraisal, values affirmations, and expressive writing are promising, only one study for each of these treatments was able to be included in our meta-analysis. As Maxwell et al. (2015) state, "The question of whether a pattern seemingly identified in an original study is in fact more than just noise can often be addressed by testing whether the pattern can be replicated in a new study" (p. 487). In their survey of research papers and meta-analyses across psychology, Stanley et al. (2018) found that individual studies were overwhelmingly underpowered, concluding that "the typical under powered study, individually or when simply combined into a meta-analysis, offers little information about the magnitude or significance of the phenomenon it examines" (p. 1342). They go beyond Maxwell et al., when it comes to establishing the convincing evidence for an effect that the best treatment recommendations are built on, calling for multisite, preregistered replication studies or, failing that, the careful meta-analytic synthesis of several unregistered replication studies. Overall, when aiming to provide recommendations for treatment, this requires an ongoing shift away from the previous criteria of the statistical significance of an effect in a single study or replication study and toward a precise estimate of an effect obtained via large multisite replications or through pooling statistical power across several studies. That said, we hope math anxiety researchers take up this approach with these promising brief interventions so evidenced-based recommendations can be offered in the future.

The small number of empirical studies conducted in the past 30 years indicates that the field of psychology is a long way from offering what is possible with respect to interventions for math anxiety. With this small set of studies, moderator analyses would likely be underpowered and at risk for type II error (Hedges & Pigott, 2001), making questions of moderation difficult to pursue. More research is needed on psychological interventions for math anxiety with diverse populations and in diverse settings (e.g. Jamieson et al., 2016) to better understand which interventions are most effective and for whom. Research on math anxiety would also benefit from the development of a consensus around which measures to use for quantifying both math anxiety and math performance scores. The studies we synthesized used several different math anxiety measures and none of the studies used the same math performance measures, making it difficult to pursue the comparisons across interventions needed to generate recommendations for education or mental health professionals. A precursor to this movement toward consensus may be basic measurement work investigating measurement invariance of math anxiety and math performance measures, to better understand how these measures perform when making the comparisons across groups that are essential to investigating conditional processes. While some invariance work on math anxiety measures has been conducted in European samples, it remains largely unexplored with participants in the U.S.A. This work could inform a thoughtful discussion between researchers on which measures may be most useful for the field to adopt moving forward. Another area in which convergence could benefit math anxiety researchers is in how to define high or low math anxious individuals. Many of the methods currently being used by researchers to classify people by their math anxiety scores rely on artificial cut points that are somewhat more grounded in statistical convenience than theoretical or empirical justification. Updated methods such as latent class or profile analysis could be used to identify more naturally occurring groups or to incorporate multiple variables, such as math anxiety, math performance, and math avoidance (Wang et al., 2018) to develop cut scores for math anxiety measures that are more grounded in students' experiences or level of risk for adverse educational outcomes.

Conclusion

The current study presents a meta-analysis of the effectiveness of interventions for math anxiety on math performance over the past 30 years. Across all studies, it was found that interventions resulted in small positive effects on math performance, reducing the performance gap due to math anxiety by approximately a third, suggesting that psychological interventions can be a useful adjunct to curricular or systemic interventions in efforts to improve numeracy in the U.S. Overall, the findings suggest that individuals with math anxiety may benefit from a range of interventions, and that more research on existing and novel interventions for math anxiety are needed.

References

Alexander, L., & Martray, C. (1989). The development of an abbreviated version of the Mathematics Anxiety Rating Scale. *Measurement and Evaluation in Counseling and Development, 22*, 143–150. doi:10.1080/07481756.1989.12022923

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General, 130*(2), 224–237. doi:10.1037/0096-3445.130.2.224

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics, 50*(4), 1088. doi:10.2307/2533446

Bennett, G. K., Seashore, H. G., & Wesman, A. G. (1981) *Differential Aptitude Tests (Forms V and W)*, San Antonio, Tx: The Psychological Corporation, Harcourt Brace Jovanovich.

Betz, N. (1978). Prevalence, distribution, and correlates of math anxiety in college students. *Journal of Counseling Psychology, 25,* 441–548.doi:10.1037/0022-0167.25.5.441

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness, 1*(4), 289–328. doi:10.1080/19345740802400072

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wiley & Sons: doi:10.1002/9780470743386

Brunyé, T. T., Mahoney, C. R., Giles, G. E., Rapp, D. N., Taylor, H. A., & Kanarek, R. B. (2013). Learning to relax: Evaluating four brief interventions for overcoming the negative emotions accompanying math anxiety. *Learning and Individual Differences, 27*, 1–7. doi:10.1016/j.lindif.2013.06.008

National Science & Technology Council Committee on STEM Education. (2018). *Charting a course for success: America's strategy for STEM education.* https://www.whitehouse.gov/wp-content/uploads/2018/12/STEM-Education-Strategic-Plan-2018.pdf

Del Re, A. C. & Hoyt, W. T. (2014). MAd: Meta-Analysis with mean differences. R package version 0.8-2. URL https://cran.r-project.org/package=MAd

Douglas, H. P., & LeFevre, J.A. (2018). Exploring the influence of basic cognitive skills on the relation between math performance and math anxiety. *Journal of Numerical Cognition, 3*(3), 642–666. doi:10.5964/jnc.v3i3.113

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. BMJ, *315*(7109), 629–634. doi:10.1136/bmj.315.7109.629

Fennema, E., & Sherman, J. A. (1976). Fennema-Sherman Mathematics Attitudes Scales: Instruments designed to measure attitudes toward the learning of mathematics by females and males. *Journal for Research in Mathematics Education, 7*(5), 324. doi:10.2307/748467

Handelsman, J. & Smith, M. (2016). Stem for all [Blog post]. Retrieved from https://obamawhitehouse.archives.gov/blog/2016/02/11/stem-all

Harrer, M., Cuijpers, P., Furukawa, T.A, & Ebert, D. D. (2019). Doing meta-analysis in R: A hands-on guide. https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/.

Hedges LV (1989). An unbiased correction for sampling error in validity generalization studies. *Journal of Applied Psychology, 74*(3), 469–477.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education, 21*(1), 33. doi:10.2307/749455

Henslee, A., & Klein, B. (2017). Using brief guided imagery to reduce math anxiety and improve math performance: A pilot study. *Journal of STEM Education*, *18*(4).

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in medicine*, *21*(11), 1539-1558.

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327*(7414), 557-560.

Hopko, D. R. (2003). confirmatory factor analysis of the Math Anxiety Rating Scale–Revised. *Educational and Psychological Measurement, 63*(2), 336–351. doi:10.1177/0013164402251041

Hopko, D. R., Mahadevan, R., Bare, R. L., & Hunt, M. K. (2003). The Abbreviated Math Anxiety Scale (AMAS). *Assessment, 10*(2), 178–182. doi:10.1177/1073191103010002008

Hoyt, W. T., & Del Re, A. C. (2018) Effect size calculation in meta-analyses of psychotherapy outcome research, *Psychotherapy Research, 28*(3), 379-388, doi:10.1080/10503307.2017.1405171

Jamieson, J. P., Peters, B. J., Greenwood, E. J., & Altose, A. J. (2016). Reappraising stress arousal improves performance and reduces evaluation anxiety in classroom exam situations. *Social Psychological and Personality Science*, *7*(6), 579–587. doi:10.1177/1948550616644656

Kim, Y., Thayne, J., & Wei, Q. (2016). An embodied agent helps anxious students in mathematics learning. *Educational Technology Research and Development, 65*(1), 219–235. doi:10.1007/s11423-016-9476-z

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*. doi:10.3389/fpsyg.2013.00863

Lyons, I. M., & Beilock, S. L. (2011). Mathematics anxiety: Separating the math from the anxiety. *Cerebral Cortex, 22*(9), 2102–2110. doi:10.1093/cercor/bhr289

Maloney, E. A., Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2015). Intergenerational effects of parents' math anxiety on children's math achievement and anxiety. *Psychological Science, 26*(9), 1480–1488. doi:10.1177/0956797615592630

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *American Psychologist, 70*(6), 487–498. doi:10.1037/a0039400

Organization for Economic Co-operation and Development. (2018). *PISA 2015 results in focus*. https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf

Park, D., Ramirez, G., & Beilock, S. L. (2014). The role of expressive writing in math anxiety. *Journal of Experimental Psychology: Applied, 20*(2), 103–111. doi:10.1037/xap0000013

Plake, B. S., & Parker, C. S. (1982). The development and validation of a revised version of the Mathematics Anxiety Rating Scale. *Educational and Psychological Measurement, 42*(2), 551–557. doi:10.1177/001316448204200218

Pletzer, B., Wood, G., Scherndl, T., Kerschbaum, H. H., & Nuerk, H.-C. (2016). Components of mathematics anxiety: Factor modeling of the MARS30-Brief. *Frontiers in Psychology, 7*. doi:10.3389/fpsyg.2016.00091

Peters, E., Shoots-Reinhard, B., Tompkins, M. K., Schley, D., Meilleur, L., Sinayev, A., … Crocker, J. (2017). Improving numeracy through values affirmation enhances decision and STEM outcomes. *PLOS ONE, 12*(7), e0180674. doi:10.1371/journal.pone.0180674

Pizzie, R., & Kraemer, D. J. M. (2018, January 20). The influence of emotion regulation on arousal and performance in math anxiety. OSF Preprints. https://doi.org/10.31219/osf.io/f3d59

Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2013). Math anxiety, working memory, and math achievement in early elementary school. *Journal of Cognition and Development*, *14*(2), 187–202. doi:10.1080/15248372.2012.664593

Ramirez, G., Shaw, S. T., & Maloney, E. A. (2018). Math anxiety: Past research, promising interventions, and a new interpretation framework. *Educational Psychologist*, *53*(3), 145–164. doi:10.1080/00461520.2018.1447384

Richardson, F. C., & Suinn, R. M. (1972). The Mathematics Anxiety Rating Scale: Psychometric data. *Journal of Counseling Psychology*, *19*(6), 551–554. doi:10.1037/h0033456

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin*, *85*, 638-641

Rounds, J. B., & Hendel, D. D. (1980). Measurement and dimensionality of mathematics anxiety. *Journal of Counseling Psychology*, *27*(2), 138–149. doi:10.1037/0022-0167.27.2.138

Schmidt, F. L. (1992). What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology. *American Psychologist*, *47*(10), 1173–1181. doi:10.1037/0003-066x.47.10.1173

Schneider, W. J., & Nevid, J. S. (1993). Overcoming math anxiety: A comparison of stress inoculation training and systematic desensitization. *Journal of college student development*.

Suinn, R. M., & Winston, E. H. (2003). The Mathematics Anxiety Rating Scale, a brief version: Psychometric data. *Psychological Reports, 92*(1), 167–173. doi:10.2466/pr0.2003.92.1.167

Vance, W. R., & Watson, T. S. (1994). Comparing anxiety management training and systematic rational restructuring for reducing mathematics anxiety in college students. *Journal of College Student Development 35*(4), 261-264.

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor package. *Journal of Statistical Software, 36*(3). doi:10.18637/jss.v036.i03

Viechtbauer, W., & Cheung, M. W.-L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods, 1*(2), 112–125. doi:10.1002/jrsm.11

Wang, Z., Shakeshaft, N., Schofield, K., & Malanchini, M. (2018). Anxiety is not enough to drive me away: A latent profile analysis on math anxiety and math motivation. *PLOS ONE, 13*(2), e0192072. doi:10.1371/journal.pone.0192072

Wilkinson, G. S. (1993). *The Wide Range Achievement Test administration manual*. Wilmington, DE: Wide Range, Inc.

Young, C. B., Wu, S. S., & Menon, V. (2012). The neurodevelopmental basis of math anxiety. *Psychological Science, 23*(5), 492–501. doi:10.1177/0956797611429134

Zettle, R. D. (2003). Acceptance and Commitment Therapy vs. Systematic Desensitization in treatment of mathematics anxiety. *The Psychological Record, 53*(2), 197–215. doi:10.1007/bf03395440

Table 1.1

Summary of Study Characteristics

Study	Tx Group Size	Sample	Design	Intervention	Math Anxiety Measure	Math Performance Task
Brunyé et al. (2013)	18	Undergraduate students	Within subjects	15 minutes of focused breathing	MARS-30	Mental arithmetic
Jamieson et al. (2016)	46	Community college developmental math students	RCT	8 minutes of stress reappraisal	AMAS	Developmental math in-class exam
Kim et al. (2017)	59	9th grade Algebra I students	RCT	240 minutes of math lessons with cognitive restructuring	MARS-R	Computer-based algebra problems
Park et al. (2014)	40	HMA and LMA undergraduate students	RCT	7 minutes of expressive writing	sMARS	Mental arithmetic
Peters et al. (2017)	112	Undergraduate students	RCT	2 15-minute values affirmations	MAS	Abbreviated objective numeracy scale
Pizzie et al. (2018)	52	HMA and LMA undergraduates	Within subjects	30 minutes of cognitive reappraisal training	MARS-30	Multi-step computer-based arithmetic
Schneider & Nevid (1993)	30*	HMA undergraduate students	3-arm RCT	360 minutes of Stress Inoculation Training or systematic desensitization	MARS	Differential Aptitude Test
Vance &Watson (1994)	39*	HMA undergraduate students	3-arm RCT, post-test only	165 minutes of Anxiety Management Training or Systematic Rational Restructuring	MARS	Departmental algebra exam
Zettle (2003)	24*	HMA undergraduate students	2-arm uncontrolled randomized trial, post-test only	360 minutes of individual Acceptance and Commitment Therapy or systematic desensitization	MARS	WRAT3 arithmetic subtest

*Indicates total number of intervention group participants across intervention arms.

Note. Tx = treatment, MARS = Math Anxiety Rating Scale (Richardson & Suinn, 1972), MARS-30 = 30 item version of the MARS (Suinn & Winston, 2003), AMAS = Abbreviated Math Anxiety Scale (Hopko et al., 2003), MARS-R Revised Math Anxiety Rating Scale (Plake & Parker, 1982), MAS = Fennema-Sherman Mathematics Anxiety Scale (Fennema & Sherman, 1976), sMARS = Short Math Anxiety Rating Scale (Alexander & Martray, 1989), HMA = high math anxiety, LMA = low math anxiety, RCT = randomized controlled trial.

Figure 1.1

Study Flow Diagram

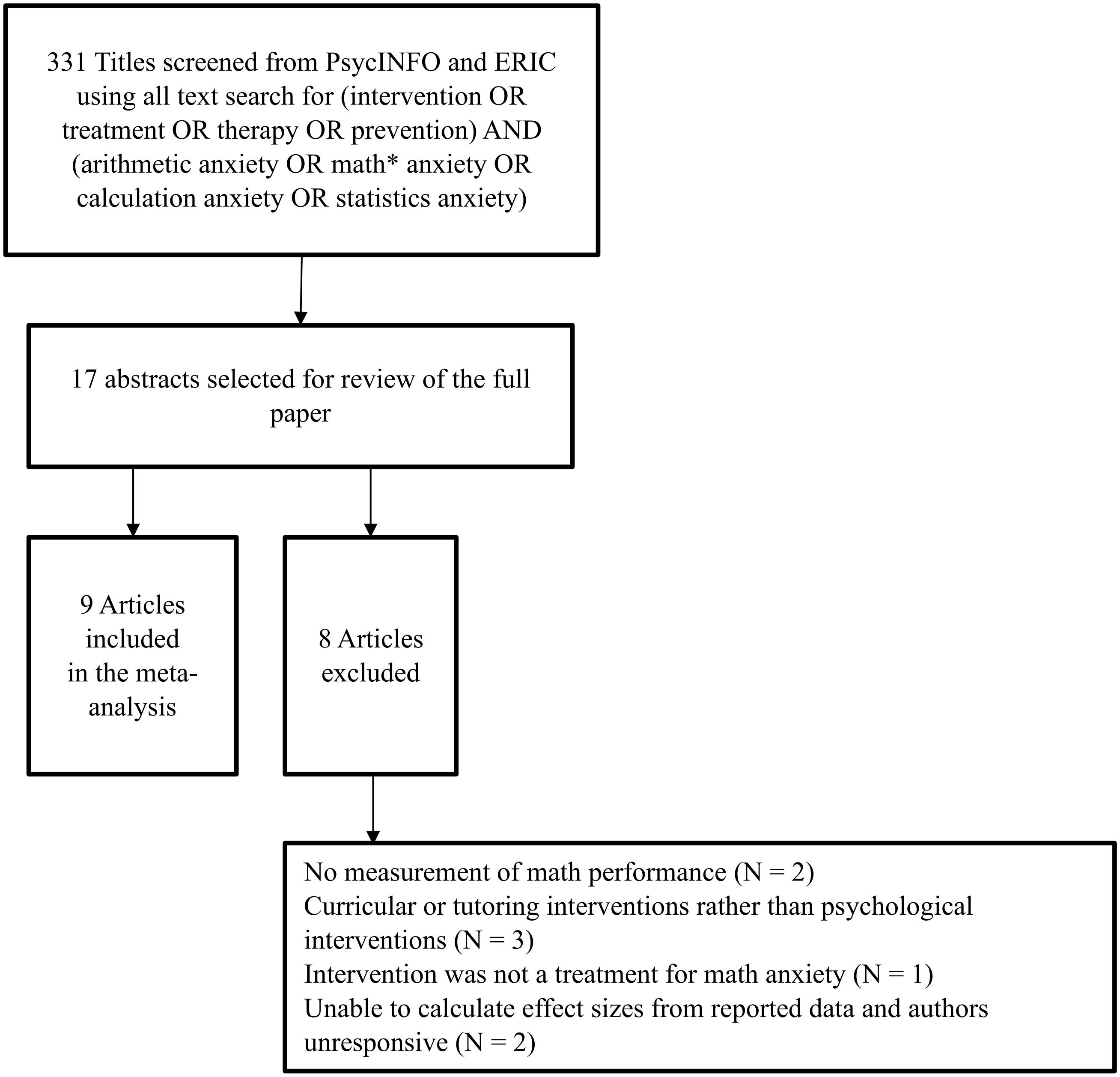
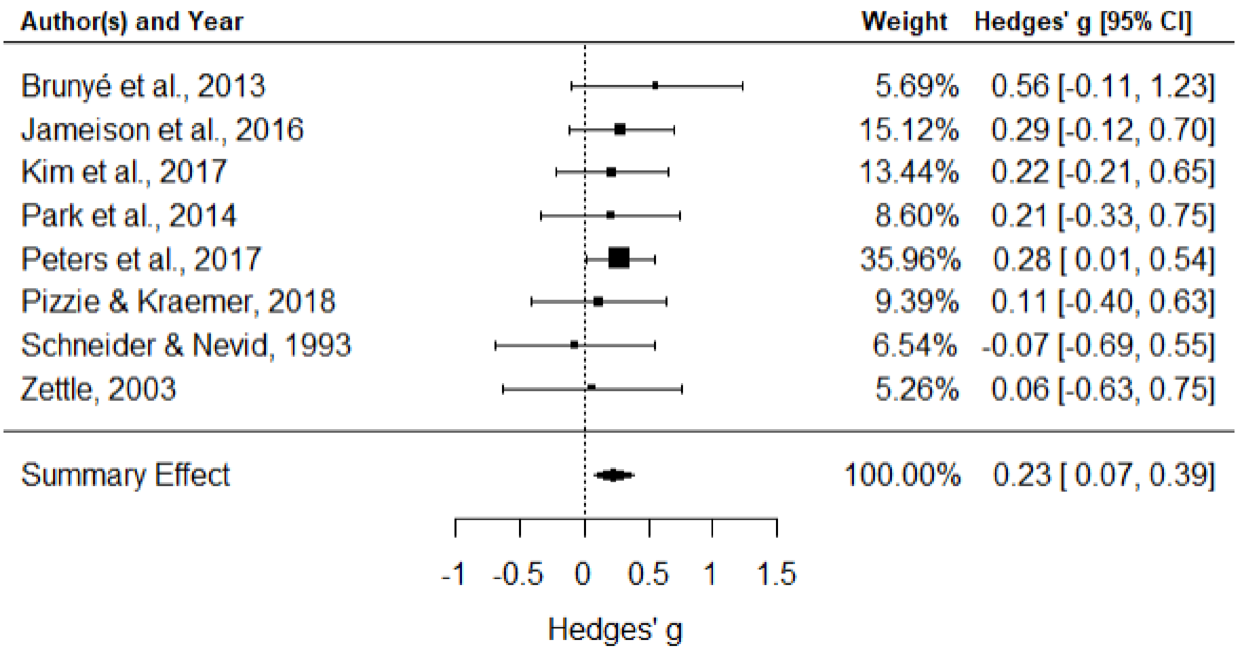


Figure 1.2

Forest Plot of Random Effects Model of Math Performance



2 The Effect of an Online Arousal Reappraisal Intervention on Math Performance: Examining the Moderating Role of Math Anxiety

For many college and high school students, feelings about math can have an outsize effect on significant life events, such as college choice and acceptance, college major, degree completion, career path, and income over the lifespan. College students who completed at least one advanced math course in high school, a time when many opt out of advanced math due to affective or attitudinal factors rather than skill deficits (Ma & Willms, 1999), have over twice the likelihood of graduating with a bachelor's degree than those who did not, independent of reading level and math ability (Trusty & Niles, 2003). One reason for this may be that collegiate developmental math classes, often a mandated destination for students who do not complete advanced math coursework, have the highest failures and withdrawal rates in higher education (Noel-Levitz, 2006). This can be a significant barrier toward degree completion, as 70% of these courses do not count toward degree credit and frequently have to be taken more than once, as only 30% of students pass all of their developmental math courses and roughly 20% of students simply avoid enrolling in the required courses at all (Bonham & Boylan, 2011). While many individual, systemic, and environmental factors interact to produce these outcomes, math anxiety has emerged as especially salient, as it has been repeatedly associated with math avoidance and negative developmental trajectories in math learning and achievement (Ahmed, 2018).

Math Anxiety and its Associations

Math anxiety, or feelings of tension, apprehension, or anxiety that inhibit one's ability to successfully engage in mathematical tasks, is most frequently operationalized as a multidimensional personality construct consisting of two higher order factors. The first factor, Mathematics Test Anxiety, defined as anxiety associated with thinking about or performing math in evaluative situations, is incorporated in virtually every two-factor model of math anxiety. While they tend to agree on the inclusion of a Mathematics Test Anxiety factor, operationalizations of math anxiety differ in their interpretation of the second factor. Some scales measure a Numerical Anxiety factor, defined as tension and fear related to having to use math or numeracy skills in everyday life (Rounds & Hendel, 1980). Other scales operationalize the second factor as Learning Math Anxiety, or anxiety due to engaging in or thinking about activities related to learning math (Hopko, 2003). Math anxiety is widely considered a persistent, trait-level construct that is associated with, but distinct from both test anxiety and state and trait anxiety (Hembree 1990; Kazelskis et al., 2000). Decades of research on math anxiety has illustrated its negative associations with variables such as math self-concept, math self-confidence, math enjoyment, math achievement, and math performance (Hembree, 1990; Ma, 1999). Students with higher levels of math anxiety also have a greater tendency to avoid mathematics, taking fewer high school math courses, showing lower intentions in high school and college to take additional math courses, and indicating lower openness to careers in physical science (Hembree 1990; Chipman et al., 1992). The negative relationships between math anxiety and math performance and math anxiety and math achievement are at the heart of why math anxiety may be a crucial element to address to improve student outcomes.

Accounts of Math Anxiety

In their narrative review of the math anxiety literature, Ramirez et al. (2018) discuss two primary accounts of math anxiety's relationship with student performance and achievement. The disruption account posits that math anxiety causes decrements in math performance through temporary reductions in working memory (WM) capacity. Working memory capacity is important because math problem solving becomes more and more dependent on working memory as number values increase and individuals must turn to strategies rather than simply retrieving the answer from long term storage (Ashcraft & Krause, 2007). In the disruption account, math anxiety, when activated by a math task, is thought to increase the potential for negative and ruminative cognitions, often related to the consequences of failure at math. Cognitive and affective engagement with these cognitions and ruminative processes act like a secondary task, drawing away working memory resources needed to solve higher order mathematical operations (e.g. carrying operations; Ashcraft & Kirk, 2001). Alternatively, the reduced competency account theorizes that math anxiety is a product of lower numerical/spatial skills, which cause underperformance and failure experiences in math and resultant anxiety. Similar to other accounts of anxiety and avoidance, math anxiety, conditioned by negative experiences with math, then leads to avoidance of math classes and other opportunities to improve math skills. This account is supported by research with math anxious students, who report taking fewer math courses and having lower intent to take math courses (Hembree, 1990). The authors argue that math anxiety develops in relation to individuals' appraisals of previous experiences with math and the outcomes of those experiences. Rather than simply arising from avoidance of math, math ability deficits, exposure to societal narratives regarding who is good at math, or unhelpful cognitions that siphon off WM resources, students' interpretations of their lived experience with math and associated internal states are integral to determining who develops math anxiety and the effect of math anxiety on performance.

The Role of Appraisal in Math Anxiety

Empirical studies provide support for Ramirez et al.'s (2018) contention that appraisals are an important component of math anxiety. Shi and Liu (2016) illustrated the negative effect that intrusive thoughts arising from negative appraisals can have on working memory in math anxious students. The authors divided their sample of Chinese undergraduate students into HMA and LMA students and randomized them into two different conditions in which WM capacity was tested with a reading span task. In the math-related condition, participants read sentences related to dysfunctional beliefs, appraisals, or responses regarding math. In the neutral condition, the sentences involved neutral stimuli. Shi and Lui found that in the math-related condition, when compared to the neutral condition, WM was reduced in math-anxious students, providing support for negative appraisals and anxious thoughts as a causal factor in WM deficits in math performance situations.

Lyons and Beilock (2012) investigated neural differences between high and low performing individuals with significant math anxiety using functional magnetic resonance imaging (fMRI) during problem-solving tasks. Participants answered sets of word problems and sets of mental arithmetic problems and were cued before each set as to whether the next set would consist of word or math problems. The fMRI results indicated that when cued to expect math problems, high performing math anxious participants showed greater activation of a frontoparietal network associated with the control of negative emotions, suggesting that emotion regulation strategies such as appraisal may be a factor in higher levels of performance in math anxious individuals.

Arousal Reappraisal

While the high-performing HMA individuals in Lyons and Beilock's (2012) study may have learned to attenuate their negative affective responses to math through trial and error, interventions that aim to efficiently teach individuals similar skills with regards to their physiological stress responses have been developed and are in the process of being tested. These interventions, known as arousal reappraisal, rely on the biopsychosocial model of challenge and threat for their theoretical underpinnings (Jamieson et al., 2013). In their biopsychosocial model of challenge and threat, Blascovich and Mendes explained that challenge states and threat states are two types of acute stress responses highly influenced by situational appraisals. Threat states are experienced when demands are appraised to outstrip available resources. Challenge states are experienced when resources are appraised to be greater than situational demands. These states have differential patterns of physiological responses. Challenge states tend to result in greater cardiac efficiency and increased vasodilation for greater blood flow, while threat states exhibit less efficient cardiac responses and constriction of blood vessels, making it more difficult to move blood throughout the periphery of the body. These states also place different demands on cognitive resources. Relative to challenge states, threat states have been associated with greater negative affect and increased attentional bias toward negative information and threat cues (Jamieson et al., 2012; Jamieson et al, 2013). By teaching recipients to change cognitions regarding their physiological arousal and view it as an adaptive coping response, Jamieson et al. (2013) theorize that arousal reappraisal interventions can make recipients more likely to respond to acute stressors through challenge rather than threat states.

Previous work on arousal reappraisal has demonstrated positive effects on math performance. Jamieson, Mendes, Blackstock, and Schmader (2010) tested an arousal reappraisal intervention with undergraduate students planning to take the Graduate Record Examination (GRE). Before taking a practice GRE in the lab, students were randomized into two groups. The intervention group received a prompt informing them that research showed that their physiological arousal could help, rather than harm, their performance on the test and that, if they felt anxious, they should remind themselves that their arousal could enhance their performance. It is important to note that this intervention is relatively brief, with its longer iterations taking no more than 8 minutes for student to complete. After completing the practice test and a one to three month waiting period, the students returned to the lab when they had completed an in vivo GRE attempt and their score reports were incorporated into the dataset. Jamieson et al. (2010) found that both in the practice and naturalistic testing situations, students who had received the reappraisal intervention scored higher on the math section of the GRE, but showed no effect on verbal performance. This is likely due to the executive functioning and WM demands placed on cognitive systems by GRE math problems. As detailed in the disruption account, viewing one's physiological stress responses as ego-syntonic and facilitative may help free up cognitive resources that had been focused on negative cognitions or self-monitoring the perceived threat of distressing levels of arousal.

A variation of this intervention was tested in a 2 x 2 randomized controlled design with female college students. Half the students experienced a laboratory stereotype threat manipulation in which a math exam modeled after the GRE was described as an indicator of intellectual ability that produced gender-differences in performance, while the other half received gender-fair instructions (John-Henderson, Rheinschmidt, & Mendoza-Denton, 2015). The authors found that female students in the stereotype threat condition that did not receive the arousal reappraisal intervention performed significantly worse than students who did, with the intervention effectively closing the stereotype threat performance gap. Also, participants were tested for cytokine Interleukin-6 (IL-6), an indicator of inflammation and physiological stress, before and after the exam. Participants who received the arousal reappraisal intervention had lower levels of inflammation regardless of testing condition. There were no significant differences in the no-threat condition between the intervention and control groups. As temporary reductions in WM have been identified as a causal factor in math underperformance due to stereotype threat effects (Maloney, Schaeffer, & Beilock, 2013), this study provides further support for arousal reappraisal as an effective intervention when WM capacity is potentially attenuated due to cognitive load.

A further test of arousal reappraisal interventions was conducted by Jamieson, Peters, Greenwood, and Altose (2016) with students enrolled in a community college developmental math course over several semesters. Using a naturalistic design, students were randomized into a control or arousal reappraisal condition. The first in-class exam served as a pretest, with a second in-class exam acting as the posttest, before which the intervention or control condition was delivered. Students who received the intervention reported reductions in math anxiety at posttest, with greater math exam performance than the control group. The authors examined mechanisms behind this effect, finding that students who reappraised their arousal had higher perceived coping resources than those in the control condition.

Areas for Future Research

Jamieson and co-authors (2016, p. 7) shared their hopes for further research on arousal re-appraisal and math anxiety, stating, “This research also calls for future research on the moderating role of math anxiety. Given the relatively low (in an absolute sense) math anxiety group means… it is possible that the reappraisal manipulation examined here is only effective for individuals exhibiting moderate levels of anxiety.” Ostensibly, with interventions such as arousal reappraisal, the objective is to buffer the effects of stress or math anxiety on the most vulnerable individuals, i.e., the most math anxious. Without examining whether math anxiety moderates the effect of arousal reappraisal interventions on math performance or other outcomes of interest, there is no way to know that the highly math anxious individuals being targeted are benefiting from the intervention, because the effects may be mainly due to performance gains in individuals with low or moderate math anxiety. Therefore, a necessary next step in evaluating arousal reappraisal interventions in relation to math anxiety and performance is to address the question of moderation.

Jamieson and co-authors also asserted, “it is our hope that this and other intervention approaches can be distilled, scaled, and disseminated to potentially improve students’ lives at near zero cost” (2016, p. 7). While we believe that more research is needed on the effectiveness of arousal reappraisal with math anxious students before committing to large-scale scale implementation in college courses, one way to test the scalability of this intervention is to administer it in an online format, an approach not yet tested by researchers. Though higher education enrollments decreased by 3.2% overall between 2012 and 2015, the proportion of students in higher education taking at least one online course has increased by 11% during the same period, to 32% of students (Allen & Seaman, 2017). During this period, online-only enrollment has increased by 9.6% percent for undergraduate students, with a 30% increase in online-only undergraduate enrollment in public institutions. This growth in online learning is in contrast to a 5% decline in on-campus enrollment from 2012-2015 across institutions of higher learning and a 4% decline in public institutions. These trends have been rapidly accelerated by the widespread adoption of online and hybrid learning formats due to social distancing requirements necessitated by the COVID-19 pandemic beginning during the Spring 2020 semester and continuing into the Fall 2020 semester (College Crisis Initiative, 2021). While, as of April 2021, most four-year colleges in the U.S. have announced plans to return to majority in-person classes, community colleges are not following suit (Burke, 2021). Also, college students have indicated that they prefer to take some of their courses in an online format, pointing to a continuation of the shift toward greater online learning (McKenzie, 2021). To summarize, enrollment trends indicate that higher education is increasingly being conducted in online formats, arguably at the expense of traditional in-class enrollment. Therefore, to keep pace with trends in educational enrollment, it is important to test psychological interventions targeting underperforming students in online as well as in-person formats.

The Present Study

To address these gaps, the current study examines the effects of a brief online arousal reappraisal intervention on math anxiety and math performance, testing whether levels of math anxiety prior to the intervention moderate the effects of arousal reappraisal on math performance and whether levels of math anxiety after the intervention mediate the effect of the intervention on math performance.

Hypothesis 1. We predict that the intervention/control condition will significantly predict math performance, such that participants receiving the arousal reappraisal intervention will perform better on a math posttest than those in the control group.

Hypothesis 2. We predict that the intervention/control condition will significantly predict math anxiety at Time 2, such that participants receiving the arousal reappraisal intervention will endorse less math anxiety than those in the control group when controlling for math anxiety at Time 1.

Hypothesis 3. We predict that math anxiety at Time 1 will moderate the effect of the intervention/control condition on math performance. We hypothesize that due to the increased tendency of HMA individuals to experience negative emotional reactions to arousal in math performance contexts and associated cognitions that interfere with WM, there will be a greater conditional effect of arousal reappraisals on math performance at higher levels of math anxiety. Specifically, we predict that HMA individuals who receive the intervention will show greater increases in math performance relative to the control group than moderate or LMA individuals who receive the intervention.

Hypothesis 4. We predict that Time 2 Math Evaluation Anxiety will mediate the effect of the intervention/control condition on posttest math performance scores. Specifically, we predict that participants receiving the intervention will show significant reductions in Math Evaluation Anxiety when compared to the control group and that a significant amount of the variance in predicted intervention group increases in Time 2 math performance scores will be accounted for by reductions in Time 2 Math Evaluation Anxiety.

Method

Participants

Participants were college students recruited from undergraduate courses at an urban university in the southeastern United States. The sample included 268 USA college students (179 women [66.7%], 85 men, 4 missing gender data). Ages ranged from 18 to 67 years (M = 23.1; SD = 5.9), with about 83% of the students aged 18-25 years. The racial/ethnic distribution was 39.9% Black or African American, 22.4% Asian or Asian American, 20.1% White, Non-Hispanic, 8.6% Hispanic or Latin-x, 7.8% Multi-racial, 0.4% Native American or Alaska Native, and 0.7% in other categories. Comparatively, in 2019, the racial/ethnic distribution of the undergraduate student body (59.2% women, 40.7% men) at that university was approximately 42.1% Black or African American, 15.2% Asian or Asian American, 22.8% White, Non-Hispanic, 12.6% Hispanic or Latin-x, 6.5% Multi-racial, 0.1% Native American or Alaska Native and 0.6% in other categories or missing (U.S. Department of Education, 2021). For their participation, some students received credit toward a research requirement in their courses while others obtained nominal extra credit per instructor preferences. The diversity of the sample is a strength. Though Hembree's (1990) meta-analysis of five studies (*N* = 804) comparing levels of math anxiety in White and Black/African American college students found no significant differences between the groups, as illustrated by the meta-analysis we undertook earlier, there is a dearth of intervention research with non-White samples, with Jamieson (2016) being the only study meeting inclusion criteria within the past 30 years.

Procedure

To address the research question, a randomized pre-test post-test with control design was utilized with data collected at two time points, the assessment stage (Time 1) and treatment stage (Time 2). All surveys and interventions were administered through Qualtrics, an online survey platform. Students who signed up for the study through the university's research participation portal and consented to participate first completed the Time 1 measures. One week later, participants who successfully completed the Time 1 measures were invited through an email prompt to complete the treatment stage of the study. At Time 2, to better approximate the evaluative threat of a real testing situation, all participants were prompted that, "This study is concerned with measurement of math and reasoning abilities. You will be working on reasoning problems as part of a test designed to measure math intelligence and predict success in collegiate mathematics courses. Your scores will be compared with those of other participants to rank your chances of success in collegiate math courses. Please make a strong and genuine effort on the test to allow for accurate evaluation of your abilities and limitations." Similar prompts have been shown to induce evaluative threat responses in undergraduate students (Martens, Johns, Greenberg, & Schimel, 2006). Participants were then randomly assigned to the arousal reappraisal or control condition. Specifically, we used the randomizer feature in Qualtrics to randomly present either the arousal reappraisal or control component to the participants. Next, participants in the arousal reappraisal condition received an intervention, adapted from Jamieson, et al. (2016) with only minor changes (see Appendix G). Participants in the arousal reappraisal condition read a series of summaries of faux scientific articles that reflect the theme that arousal should be viewed as adaptive and facilitative in challenging situations, rather than distressing or harmful, and that arousal aids performance. After each article, participants were required to provide brief written responses of at least 20 characters to questions that served to reinforce the information provided and act as a manipulation check. To match the reading and writing time of the intervention group participants as closely as possible, the control group read a series of summaries of nature articles about how birds have evolved to meet the demands of flight (see Appendix H). Control group participants read the same number of summaries as participants in the arousal reappraisal condition and were also required to provide brief written responses of at least 20 characters to questions about each article as a manipulation check. The control group article summaries and the question prompts that followed were matched to counterparts in the arousal reappraisal condition such that the passages were identical in word count and within the same Flesch Kinkaid Grade reading level. All passages were on or below the 8th grade reading level. Participants in the intervention group averaged 36 words per response while those in the control group typed an average of 47 words per response. After the intervention or control, participants completed the Time 2 math anxiety measures followed by the post-test math performance measures. Concealment was used in several ways throughout the course of the study. To reduce potential selection effects, in the portal the participants used to access the study, we described the study as related to personality, reasoning, and collegiate success and did not inform participants of the quantitative nature of the tasks they would be completing or that the study would be focused on math anxiety. Our threat-induction prompt informed students that their scores would be compared with those of other students to rank their chances of collegiate success, while no actual comparisons were made, nor would it be likely that those ranks would be predictive of their collegiate success. Upon completion of the study, participants were given more information about the study, including the elements of concealment. Participants were also informed that "math achievement tests such as the ones you took as a part of this study cannot account for the amount of effort you will put into your future studies, which is an extremely important factor in academic success."

Measures

Math anxiety. Math anxiety was measured using the Abbreviated Math Anxiety Scale (AMAS; Hopko, Mahadevan, Bare, & Hunt, 2003). The AMAS is comprised of two subscales, LMA (5 items), assessing anxiety related to learning math (e.g. "Listening to a lecture in math class") and MEA (4 items), capturing anxiety regarding performing math in an evaluative context (e.g. "Being given a 'pop' quiz in math class"). It uses a 5 point response scale ranging from 1 (*low anxiety*) to 5 (*high anxiety*) with higher scores indicating greater levels of math anxiety. Studies incorporating the AMAS have reported high internal consistency in college samples, with coefficients α in the mid 0.80 to low 0.90 range for the full scale and both subscales (Douglas & LeFevre, 2018; Hopko et al., 2003; Schillinger, Vogel, Diedrich, & Grabner, 2018). Hopko et al. (2003) reported excellent test-retest reliability over a two-week period for the AMAS ($r = 0.85$) and the LMA ($r = 0.78$) and MEA subscales (($r = 0.83$). The AMAS has shown good psychometric qualities and while, like other math anxiety measures, measurement invariance across gender and race in U.S. samples has not been reported, there has been support for gender invariance in Italian samples (Primi et al, 2014).

Math Performance. Math performance pre-test and post-test measures were created by dividing items of the CFT (Sowinski, Dunbar, & LeFevre, 2014) and the BMA-3 (Steiner & Ashcraft, 2012) into two equal forms of 95 items each. The CFT, developed to improve on existing measures of speed and accuracy of multi-digit calculation, consists of one 60-item page each of double-digit addition, subtraction, and multiplication problems. For each operation, following the completion of two practice items, participants are given one minute to complete as many problems as possible. Estimates of the internal consistency of the CFT have been high, with Cronbach's α of 0.90 or greater (Douglas & LeFevre, 2018; Sowinski et al., 2014) and approximately normal distributions of scores. The CFT has shown good convergent validity with other arithmetic skill measures (Bourassa, 2014) and correlations in the expected directions with math anxiety in previous studies (Douglas & LeFevre, 2018; Sowinski et al., 2014). To create a split form with different items on the pre and posttest to prevent practice effects, odd numbered items were assigned to the pretest and even numbered items were assigned to the posttest. As a result, each test consisted of one page of 30-items for double-digit addition, one page of 30-items for double-digit subtraction, and one page of 30 items for double digit multiplication. Participants were given 30 seconds rather than the typical 60 seconds per page to complete as many items as possible due to the reduced number of items. Scoring of this measure is based on the total number of correct items.

The BMA-3, based on items from the Wide Range Achievement Test 3 (WRAT3; Wilkinson, 1993), was created by Steiner and Ashcraft (2012) as a 10-item brief assessment of math achievement for university students that could substitute for the full version of the WRAT3 in studies of math anxiety. BMA-3 items increase in difficulty as the assessment proceeds. The BMA-3 has an overall correlation of 0.66 with the fourth edition of the WRAT (WRAT4;

Wilkinson & Robertson, 2006), which the authors related to the decreased difficulty of the WRAT4 when compared to the WRAT3. Steiner and Ashcraft (2012) reported adequate internal consistency in the initial sample, with Cronbach's $\alpha = 0.69$. Scores on the BMA-3 have correlated negatively with math anxiety with correlations ranging from $r = -0.27$ to $r = -0.41$ (Douglas & LeFevre, 2018; Steiner & Ashcraft, 2012), which is consistent with past research on math anxiety and math achievement (Hembree, 1990). Scoring of the BMA-3 is based on the total number of correct items.

To construct a split form of the BMA-3 to minimize practice effects, values used in individual items were modified slightly to create 10 additional items requiring the same procedures as the original items, but with different item solutions. This mirrored the process reported by Steiner and Ashcraft of creating the BMA-3 by modifying WRAT3 items. The items generated in this fashion were sequenced from easiest to hardest, corresponding to the arrangement of the original items of the BMA-3, and used as a pre-test measure of math achievement.

Data Analysis

Power analyses. Consistent with recommendations by Frazier, Tix, and Barron (2004), a priori power calculations were conducted in G*Power (version 3.1.9.2; Faul, Erdfelder, Lang, Buchner; 2007) to provide guidance on appropriate sample size to detect hypothesized main and interaction effects. To achieve power of 0.80 to detect the $d = 0.61$ effect size consistently reported in arousal reappraisal interventions (Jamieson et al, 2016) for math performance, given an alpha level of .05 and assuming equal sized treatment and control groups, a minimum total sample size of 68 (34 participants per group) is required. For math anxiety, achieving power of 0.80 to detect the $d = 0.49$ effect of arousal reappraisal on math anxiety found in Jamieson et al.

(2016), the only arousal reappraisal study to measure changes in math anxiety, requires a minimum total sample size of 106 (53 participants per group). The test of the moderation effect involves a null hypothesis test of whether the percentage of variance explained by the intervention/control condition x math anxiety interaction term is significantly different from zero (Jaccard, Turrisi, & Wan, 1990). To achieve 0.80 power, given a relatively small (Cohen, 1988), but meaningful effect of $f^2 = .053$, or ~5% of the variance in math performance accounted for, a minimum total sample size of 230 (115 participants per group) is required.

For the mediation analyses, power was calculated using correlations and standard deviations calculated from Jamieson et al. (2016) and Hembree (1990). Monte Carlo simulations were run using Schoemann, Boulton, and Short’s (2017) web-based power analysis tool. For analyses to achieve .80 power to detect an indirect effect of arousal reappraisal on math performance through math anxiety given a correlation of $r = -0.21$ (between a small and medium effect) between the intervention/control condition and math anxiety, $r = 0.14$ (a small effect) between the intervention/control condition and math performance, and $r = -0.25$ (halfway between a small and medium effect) between math anxiety and math performance, a minimum total sample size of 208 (104 participants per group) is required. This minimum sample size is further supported as a conservative estimate according to simulation studies conducted by Fritz and Mackinnon (2007). Testing the power of the percentile bootstrap method we used with an effect size for the *a* and *b* paths of a simple mediation model halfway between a small and medium effect size, as we predict the magnitude of our effects will be, Fritz and Mackinnon estimated that a minimum sample size of 162 participants would be required to have power of 0.80 to detect an indirect effect at the $a = 0.05$ level.

Data screening. To screen for and exclude participants who demonstrated behavior consistent with insufficient effort responding, we utilized a combination of proactive and reactive approaches (Dunn et al., 2018). Proactively, we embedded directed response items (e.g. “Select high anxiety for this item.”) into the Time 1 and Time 2 AMAS surveys (Meade & Craig, 2012). As the first step of the data screening process, these items were reviewed after participants completed the Time 1 surveys and again following Time 2. Fifty-one participants were excluded from the study following Time 1 due to failing to enter the directed response and an additional 9 participants were excluded following Time 2. Next, we utilized reactive, or post-hoc, approaches. We examined participant response times for the Time 1 and Time 2 surveys separately, as overly brief response times suggest a lack of effortful engagement with the survey items. None of the remaining participant responses were under the 2s per item cutoff suggested by Huang and co-authors (2012), therefore no participants were excluded using this method. Finally, we conducted a manipulation check by analyzing participants’ text responses to the reappraisal and no-appraisal writing prompts. Participants in the treatment group who did not write about how the information in the articles could help them on the exam were excluded from the analysis. Similarly, participants in the control group who did not write about how the anatomy of birds evolved to help them fly were also excluded from the analysis. Of the remaining 276 participants, seven were excluded from the treatment group, and one participant was excluded from the control group, leaving a sample of 268 participants whose data were analyzed.

Several themes emerged from students’ text responses to the intervention. Many students wrote about learning that “stress is positive” and to not think of stress “as a negative thing.” Several students wrote that it was helpful to know that their stress response worked to provide "fuel to the brain" and get the "body prepared for the task." Other themes emerging from students' writing was that it was "reassuring" to know that their stress response was "normal" and that there was "nothing wrong with me" for experiencing the stress response. These themes tended to emerge in the writing of both high and low math anxious students (defined as one *SD* above and below the total AMAS score mean respectively). However, highly math anxious students tended to be more likely to write about learning to "change my perception of stress" and about "controlling stress," while lower math anxiety students were more likely to write explicitly about stress being "an advantage" or something to "embrace."

Data Analytic Strategy. A regression framework was used to analyze the data for questions of interest, as regression models utilizing pretest covariates reduce $SS_{residual}$ to provide more precise estimates of treatment effects and are particularly useful in testing and quantifying conditional effects of interactions between categorical predictors and continuous moderators (Darlington & Hayes, 2016). To test the hypothesis that arousal reappraisal enhances math performance, the dichotomous intervention/control condition variable and pretest math performance scores were entered into a regression model as predictors with posttest math performance scores as the outcome variable. The model was run twice, first with BMA-3 pretest scores as a predictor/covariate and BMA-3 posttest scores as the outcome variable and again with CFT pretest scores predicting CFT posttest scores.

To test the second hypothesis, that arousal reappraisal reduces math anxiety, the intervention/control condition variable and pretest AMAS factor scores (Learning Math Anxiety and Math Evaluation Anxiety) were entered in the model as predictors, with Time 2 AMAS scores as the outcome variable. To test for differential effects on LMA and MEA, the analysis was run twice, first with pretest LMA as a predictor and posttest LMA as the outcome variable and again with pretest MEA as a predictor and posttest MEA as the outcome variable.

To test the third hypothesis, that the effect of arousal reappraisal on math performance is moderated by math anxiety, we constructed moderation models using Hayes' (2017a) PROCESS macro with the intervention/control condition variable as the focal predictor (X), pretest math anxiety scores as the moderator (W), posttest math performance scores as the outcome variable (Y), and pretest math performance scores as a covariate. We tested a model with LMA as the moderator and BMA-3 scores as the outcome variable, then tested the model again with CFT scores as the outcome variable. We completed the same process with MEA as the moderator, constructing four models in total.

To test the fourth hypothesis, that reductions in math anxiety will mediate the positive effect of arousal reappraisal on Time 2 math performance, we used Hayes's PROCESS macro (Hayes, 2017a) to construct mediation models with the intervention/control condition variable as the focal predictor (X), posttest math anxiety scores as the mediator (M), and posttest math performance scores as the outcome variable (Y). As an inferential test of mediation, for each model, we calculated bootstrap (10,000 samples) 95% confidence intervals for the indirect effect of the intervention/control condition on Time 2 math performance scores through Time 2 math anxiety. Time 1 math performance scores were included in the models as a covariate. As with the moderator analyses, the mediator model was run four times, with combinations of Time 2 LMA or MEA scores as the mediator (M), and Time 2 BMA-3 or CFT scores as the outcome variable (Y).

Results

Preliminary Analyses

Table 2.1 presents descriptive statistics and basic correlations for the math anxiety and math performance variables by treatment group. Time 1 LMA and MEA scores and internal consistency were comparable to other recent studies with undergraduates (Jamieson et al., 2020; Schillinger et al., 2018). Time 1 BMA-3 scores were also in line with those reported by past studies with college samples, though the internal consistency was poor and lower than past studies (Douglas & LeFevre, 2018; Steiner & Ashcraft, 2012). Estimates of the internal consistency of the BMA scores for the intervention group in this study indicated that the scores for that group were comprised of about equal parts of true score and measurement error, making it extremely hard to draw valid statistical conclusions from scores obtained using that measure. This low internal consistency of the BMA-3 scores in our study may be in part due to the escalating difficulty of BMA-3 items. When correcting for time allowed for the task, students in our sample scored 6-7 points lower on average on the CFT at Time 1 than the Canadian undergraduate students assessed in previous studies (Douglas & LeFevre, 2018; Sowinsky et al., 2014), which may be due to differences between the Canadian and U.S. educational systems. CFT internal consistency was good and slightly lower than in past studies.

To evaluate the success of randomization procedures, chi-square tests of independence were conducted to ensure the intervention and control groups did not significantly differ on gender $\chi^2(2, N = 268) = 1.11$, $p = 0.57$ or race $\chi^2(6, N = 268) = 8.68$, $p = 0.19$. Independent sample t-tests were conducted to evaluate pre-treatment intervention/control differences on LMA $t(266) = 1.05$, $p = 0.29$, MEA $t(266) = -0.99$, $p = 0.32$, BMA-3 $t(265) = 0.45$, $p = 0.66$, and CFT scores $t(266) = -0.97$, $p = 0.33$. The null hypothesis that the control and intervention groups differed on pre-treatment or demographic variables was rejected for each of these tests, supporting the efficacy of randomization procedures.

Regression diagnostics were conducted to evaluate the structure of the data, determine if the fitted regression model adequately represented the data, identify influential cases, and detect violations of the typical regression assumptions of homoscedasticity, normality, and linearity. We used procedures described by Darlington and Hayes (2016) to assess for potential heteroscedasticity by examining the relationship between the square of the *t*-residual (the *t*-residual is referred to as the “studentized deleted residual” in SPSS software) and the regressors in each model we constructed. Under the standard assumptions of regression, the variance of the *t*-residuals should be the same regardless of the values of the regressors. Therefore, any association found between the squared *t*-residual and a regressor or set of regressors is evidence against the assumption of homoscedasticity. This test involves transforming each squared *t*-residual into normalized scores known as Van der Waerden scores. The transformation is accomplished by first forcing the distribution of the squared *t*-residual to be approximately normal by replacing the values of the squared *t*-residual with their ranked position in the distribution (e.g. the smallest squared *t*-residual will have a value of 1, the second smallest will have a value of 2, etc.). These ranks are then divided by $1 + N$ and converted to Z-scores, in our case by using the IDF.NORMAL function in SPSS, to produce the Van der Waerden scores, which are approximately normally distributed. The Van der Waerden scores are then regressed on all the regressors in the model being tested. If the multiple correlation, R, is statistically significant, there is evidence that the Van der Waerden scores are not independent of the regressors and that the assumption of homoscedasticity is not tenable. We completed these procedures for each model we constructed, finding violations of the assumption of homoscedasticity for both the LMA ($R = .31$, $p < .001$) and MEA ($R = .20$, $p = .004$) models tested as part of Hypothesis 2 as well as for the moderation model regressing Time 2 BMA-3 scores on the dichotomous condition variable, Time 1 MEA scores, and their interaction term ($R = .19$, $p = .04$) tested as part of Hypothesis 3.

Regarding normality, Shapiro-Wilk tests (Shapiro & Wilk, 1965) indicated a lack of support for the assumption that data for the Time 1 and Time 2 MEA, LMA, CFT, and BMA-3 variables were sampled from a population that is normally distributed on those variables (ps < .001). Plots of Pearson residuals versus fitted values and versus the values of individual regressors displayed evidence of systemic features and, in the case of the moderation models, isolated points, highlighting possible violations of the assumption of normality (Fox & Weisberg, 2019). Additionally, Studentized residuals, Cooks distances, DFFTS, and DFBETAS calculated for each model revealed the presence of influential cases that significantly affected the regression surface and, in turn, decisions about which models best fit the data. These cases were scrutinized and errors in data recording or calculation were ruled out.

Due to the violations of linear regression assumptions detailed above, we adjusted our analytic approach in two major ways. To account for heteroscedasticity and the presence of influential data points, standard errors were calculated using the HC4 heteroscedasticity-consistent standard error estimator. In simulations involving high leverage data points, HC4 has exhibited less bias in inferential tests than other heteroscedasticity-consistent standard error estimators (Cribari-Neto, 2004). The HC4 estimator is available through the PROCESS (Hayes, 2017a) and RLM macros (Darlington & Hayes, 2016) for SPSS. We also constructed 95 percentile bias-corrected bootstrap confidence intervals, utilizing 10,000 bootstrap samples, for all the regression coefficients in the models tested, which are the confidence intervals we report in the text. Methodological studies (Kelley, 2005; Russell & Dean, 2000; Wood, 2005) support the use of nonparametric bootstrap techniques when analyzing nonnormal data, or when traditional linear regression assumptions are untenable. Additionally, bootstrapping has more power to detect interaction effects than alternatives such as logarithmic transformation, without the interpretive complications that arise when data is transformed.

To test the assumption of linearity, we created quadratic functions by adding the square terms of predictors to our regression models and assessed for significant ($p < 0.05$) change in model R^2 (Cohen et al., 2003; Darlington & Hayes, 2016). For example, to assess for nonlinearity in the moderation model regressing Time 2 BMA-3 scores on the dichotomous condition variable, Time 1 MEA scores, and their interaction term, tested as part of Hypothesis 3, we added the square term of Time 1 MEA scores as well as the interaction term of the squared Time 1 MEA scores and the dichotomous condition variable (Hayes, 2017b). The squared term of the dichotomous condition (X) variable was not included in these models because there is no mathematical difference between it and its squared term. For all the nonlinear models tested, none of the squared terms or interaction terms including squared predictors were significant predictors of the dependent variable and adding the squared terms did not result in significant changes in model R^2 (ps > .05). These findings support the assumption that the relationship between the predictors and outcome variables is a linear one.

There was little missing data in our dataset. Time 1 and Time 2 BMA scores were both missing a single case out of 268 (0.4%) and the gender variable had 4 missing cases (1.6%). All other variables of interest contained complete data. Little's missing completely at random test $\chi^2(14, N = 268) = 36.20$, $p = .001$ supported the hypothesis that data were not missing completely at random. Missing cases were handled via listwise deletion. Accordingly, analyses incorporating BMA scores included 266 cases, while all other analyses included the full 268 cases.

Regression Analyses

Hypothesis 1

In Hypothesis 1, we asserted that the intervention/control condition would significantly predict posttest math performance scores, when controlling for pre-test scores. For both the BMA-3 and CFT models, the intervention/control condition variable was not a significant predictor of the posttest math performance scores, ps > .26 (see Table 2.2), and both sets of confidence intervals straddled zero, indicating a null treatment effect and failure to reject the null hypothesis. For both models, pre-test math performance scores significantly predicted post-test performance.

Hypothesis 2

In Hypothesis 2, we predicted that the intervention/control condition would significantly predict Time 2 math anxiety scores and expected intervention recipients to have significantly less math anxiety than members of the control group. As shown in Table 2.2, the condition variable did not significantly predict Time 2 LMA or MEA scores, ps > .26, and confidence intervals for the coefficient of the condition variable in both models contained zero, indicating a lack of support for Hypothesis 2. In both LMA and MEA models, Time 1 math anxiety scores predicted math anxiety scores at Time 2.

Hypothesis 3

In Hypothesis 3, we postulated that Time 1 math anxiety would moderate the effect of the intervention/control condition on Time 2 math performance scores. We expected the intervention to have a greater positive conditional effect on math performance for participants with high Time 1 levels of math anxiety compared to those with lower Time 1 levels of math anxiety. As shown in Table 2.3, of the four models tested, one model showed evidence of an effect of the intervention/control condition variable on Time 2 math performance scores conditioned on the level of Time 1 math anxiety. Specifically, the intervention/control condition variable interacted with Time 1 MEA to significantly predict Time 2 CFT scores ($p < .05$, 95% CI [-1.63, -0.08]). While, in that model, the intervention/control condition variable ($p = .04$, 95% CI [0.18, 6.32]) and Time 1 CFT scores ($p < .01$, 95% CI [0.69, 0.90]) significantly predicted Time 2 CFT scores as well, the significant interaction effect takes interpretive precedence. We explore that interaction further in the following section. Regarding the other moderation models tested (see Table 2.3), there was no evidence of the intervention/control condition variable or the condition x math anxiety interaction term predicting Time 2 math performance scores. Time 1 MEA scores significantly predicted Time 2 BMA scores ($p = .04$, 95% CI [-0.37, -0.01]), as did the Time 1 BMA score covariate ($p < .01$, 95% CI [0.57, 0.77]). In the two models testing Time 1 LMA scores as the moderator variable, the Time 1 math performance score covariates were the only significant predictors (ps < .01) of the Time 2 math performance score outcome variables.

Probing the conditional effect of the intervention and MEA on CFT scores. We probed the significant intervention/control condition x Time 1 MEA interaction, shown graphically in Figure 2.1, by utilizing the pick-a-point approach (Rogosa, 1980). Specifically, we plotted CFT score estimates for individuals in the intervention and control conditions at low, medium, and high levels of the moderator. Low, medium, and high levels of Time 1 MEA were defined as the 16th, 50th, and 84th percentiles of the distribution (Hayes, 2017a). Examining group differences, we see that for those with low levels of Time 1 MEA, predicted Time 2 CFT scores were 1.43 points higher (Cohen's $d = 0.32$) for participants who received the intervention than for those in the control group. For those with medium levels of Time 1 MEA, predicted Time 2 CFT scores were 0.22 points higher for participants in the treatment group (Cohen's $d = 0.05$) than for control group participants. For individuals with high levels of Time 1 MEA, predicted

Time 2 CFT scores were 0.62 points *lower* for treatment group participants (Cohen's d = -0.14) than for control group participants, though this negative effect on highly math anxious participants was not statistically distinguishable from a null effect.

To further explore the interaction, we employed the Johnson-Neyman technique (Bauer & Curran, 2005), to identify the value or values of the moderator at which the conditional effect of the intervention on Time 2 CFT scores shifts between statistical significance and non-significance at the $\alpha = .05$ level. These values mark the boundaries of the region of significance for the effect. Visualized in Figure 2.2, the conditional effect decreases linearly as Time 1 MEA increases, becoming statistically indistinguishable from a null effect as Time 1 MEA increases past 2.55 and the lower limit of the 95% confidence interval for the effect crosses zero. In our sample, over 22% of the participants had Time 1 MEA scores falling in the region of significance, or below 2.55, indicating they would likely have experienced a statistically significant increase in Time 2 CFT scores, had they received the intervention.

Hypothesis 4

In Hypothesis 4, we stated that the intervention/control condition variable would have a significant effect on Time 2 math performance scores through Time 2 math anxiety. We predicted that individuals who received the intervention would show reductions in Time 2 math anxiety which would, in turn, predict increases in math performance scores. Results for the mediation models tested are displayed in Table 2.4. None of the four models tested showed significant indirect effects of the condition variable on math performance through math anxiety, as the 95% confidence intervals for the indirect effects all included zero. This is unsurprising given the failure to find significant differences between the intervention and control groups on Time 2 math performance or Time 2 math anxiety in the analyses for Hypotheses 1 and 2.

Discussion

This study set out to answer several questions. Our primary question was whether the arousal reappraisal intervention developed and tested by Jamieson et al. (2016) would continue to be effective in an online setting. Counter to our hypothesis, we found that when delivered online, the arousal reappraisal intervention did not significantly improve participants' scores on online math performance tasks. These findings diverged from those of several previous studies (e.g. Jamieson et al., 2010; Jamieson et al., 2016; John-Henderson et al., 2015), but aligned with the findings of Hangen et al. (2018). Hangen and co-authors found null effects for arousal reappraisal on math performance, but in their exploration of moderators, found that men who received the intervention performed better than men who did not, while female participants experienced no effect. This points to the importance of continued research on moderators of the effects of arousal reappraisal, a topic we return to in the next paragraph. In addition to math performance, we were also interested in whether the arousal reappraisal intervention would continue to be effective at reducing math anxiety prior to an evaluative math task when delivered online. Once again, counter to the hypothesized outcome and to Jamieson et al. (2016), the online arousal reappraisal intervention did not significantly reduce participants' math anxiety.

Beyond the above questions, we also wanted to answer the call put forth by Jamieson and co-authors (2016) for further research on whether math anxiety moderates the effectiveness of arousal reappraisal interventions. Here we found evidence that, as Jamieson et al. speculated, the online arousal re-appraisal intervention was effective in improving math performance scores in students with low to moderate levels of Math Evaluative Anxiety by 11-23%, with effects decreasing as MEA increased. However, there was no evidence for its effectiveness with more math anxious students. More specifically, the higher a student's level of MEA, the less effective the intervention appeared. These findings were opposite our predictions that highly math anxious students who received the intervention would have the greatest math performance gains. Work on uncovering the mechanisms behind arousal reappraisal is still in its early stages. Nevertheless, past theory suggests that when effective, arousal reappraisal forestalls negative cognitive cycles regarding physiological stress responses and enhances recipients' perceptions of their available resources, promoting challenge rather than threat states (Jamieson et al., 2016). With our most math anxious students, this intended effect did not occur, leaving us to wonder, why? It is possible that many students with high levels of math anxiety have difficulty quickly attending to and reappraising their physiological arousal before returning their attention to the math task at hand. Unlike less anxious students, arousal reappraisal interventions could cause highly math anxious students to focus *more* attention for longer periods of time on their physiological stress response and related cognitions than they would normally. As you may recall, according to the disruption account of math anxiety (Ramirez et al., 2018), attending to negative cognitions and physiological sensations for an extended period may act as a secondary task that pulls working memory resources from math performance tasks. Math tasks requiring strategies rather than simple memory retrieval, such as the ones our participants completed, place high demands on working memory and performance tends to degrade when working memory resources are allocated elsewhere (Ashcraft & Krause, 2007).

Additionally, students with long-standing experience of high levels of math anxiety and associations between physiological stress responses and academic struggles may view the arousal reappraisal intervention as invalidating or providing evidence for beliefs that their inability to harness their stress response means that there is something lacking or wrong with them. This could reinforce negative appraisals about their ability to meet math performance challenges, decreasing their perceptions of the resources they have available to meet the demands of the task and making a threat response more likely than a challenge response. These highly math anxious students may require more practice to change their mental habits and successfully re-appraise their stress responses than less anxious students, something that was not accounted for by our one “dose” intervention.

Also, for those students with low to moderate levels of MEA, math performance improved only in the CFT task, a timed math task designed to place heavy demands on working memory. The intervention had no effect on scores in the untimed BMA-3 task, regardless of students’ level of MEA. One interpretation is that this indicates further support for working memory as a mechanism through which arousal reappraisal interventions exercise their positive effects on math performance. However, we encourage caution when interpreting this finding, as the poor internal consistency of the BMA-3 measure during both time points of the study indicated that the observed scores for the intervention group contained an unacceptable amount of measurement error, resulting in reduced power to uncover any intervention effects.

Limitations

While this study begins to address significant gaps in the literature around math anxiety and arousal reappraisal interventions, there are several limitations that are important to consider. In their 2002 book, a successor to classics on experimental and quasi-experimental design by Campbell and Stanley (1963) and Cook and Campbell (1979), Shadish and co-authors discuss threats to validity relevant to the current study. We wanted to extend previous studies by testing the intervention in an online environment. Accordingly, there was a lack of control in our study over the experimental setting, as participants completed the study in whatever environment they found themselves in when logging in to the survey. Therefore, extraneous variance may have been introduced in the form of distractions, interruptions, technical difficulties, or other factors that may have been better controlled in a laboratory setting. This extraneous variance could have led to increased error that detracts from the validity of our statistical conclusions, though these conditions also help replicate the environment that many students are in when they engage in online learning. These concerns may be lessened somewhat by Klein et al.'s (2018) massive Many Labs 2 replication project, which found there was little variation between effect sizes when online and in-lab samples were given the same intervention.

Another limitation is the threat to construct validity presented by the mono-operation and mono-method bias inherent to our use of a single self-report measure to operationalize the construct of math anxiety. It is possible, due to this method bias, that we may be omitting important aspects of math anxiety, such as physiological responses or processes occurring beneath the level of verbal awareness, from our analyses. However, as Cipora et al. (2019) discuss in their review of math anxiety measures, physiological and neurological methods are not yet able to differentiate math anxiety from other stress reactions. There are also several threats to external validity that limit the generalizability of our findings. Our sample is one such threat. It was drawn from a single university and contained significantly more Black or African American and Asian or Asian American students and significantly less White, Non-Hispanic and Hispanic or Latin-x students proportionally than a representative sample of U.S. undergraduate students (U.S. Department of Education, 2021). Our sample also contained about 10% more women and 10% fewer men than is representative of U.S. undergraduates. Additionally, our sample consisted of undergraduate students who chose this study out of several other study options to gain course credit or extra credit. While the study description and instructions that students saw were kept as general as possible, students in our sample may have more interest in math or reasoning problems than a typical student. Due to the lack of research on moderators of arousal reappraisal, it is difficult to say how these characteristics of our sample may have influenced our results, but they remain important to note when considering how this intervention may be used or tested on other participants.

A second threat to external validity involves the randomized pre-test post-test design used in our study. While the randomized pre-test post-test design is excellent at controlling for alternative hypotheses that threaten internal validity, the inclusion of a pre-test introduces a threat to external validity in the form of a potential interaction between taking the pre-test and engaging in the intervention. Past research has shown that pretests can dampen the effect of interventions in randomized pre-test post-test designs (Campbell & Stanley, 1963). Though the threat of the pre-test x intervention effect in our study may be minimal due to the commonality of testing in our students' lives, researchers may want to consider utilizing a Soloman (1949) Four-Group Design to control for pre-test effects.

Additional limitations concern whether our findings will generalize outside of the experimental setting. While we attempted to create evaluative threats at Time 2 through a stress inducing prompt that have been used in other studies, we did not measure whether this prompt had its intended effect. It is possible that the effects of the online arousal reappraisal intervention on students will change in higher stakes evaluative situations when students may perceive a greater threat or have greater levels of motivation, such as during course exams or standardized tests. Another threat to external validity involves the fact that these data were collected during the Fall 2020 semester, amidst the global COVID-19 pandemic and an unprecedented shift from in-class instruction to online or hybrid educational formats (College Crisis Initiative, 2021). In the Fall 2020 version of their annual national survey of 32,000 students, the Healthy Minds

Network (2021) found that over half of students were reporting significant levels of depression or anxiety, the highest prevalence in the history of the survey. Two-thirds of student respondents reported feeling lonely or isolated and 83 percent reported that their mental health had negatively affected their academic performance in the past month. Son et al.'s (2020) qualitative research with university students in Texas conducted during the Spring 2020 semester tells a similar story, with 71% of the students describing increases in depression and anxiety, 86% reporting greater social isolation, 89% discussing difficulties with concentrating on academic work, and 38% of students reporting that their greatest academic challenge was the transition to online classes. While preassessment math anxiety scores in our sample were comparable with past studies (Andrews & Brown, 2015; Hopko et al., 2003), we assume that the COVID-19 pandemic affected our participants in multiple ways that are impossible to account for or replicate, including a rapid and widespread shift to online learning at the university from which the sample was recruited.

Future Research

Given the limitations discussed above, future projects may attempt to replicate the study in more naturalistic online environments with real-world threats, for example, integrating the study into online university math courses and using course exams as the pre-test and post-test measures. End of semester course grades could also be analyzed to investigate moderator effects over greater timespans. Further exploration into mechanisms of arousal re-appraisal would also be welcome additions to the literature. Studies with a more robust measurement strategy including physiological measures and measures of potential mediators such as affect, demand and resource appraisal, course engagement, and approach and avoidance motivation could help illuminate the paths through which these interventions operate and help identify why arousal reappraisal may backfire with some participants. Many potential moderators such as gender, race, perceived math ability or self-concept, math motivation, physiological interoceptivity, and implicit theories about stress and intelligence also remain unexplored or underexplored. In their discussion of math anxiety, Ramirez et al. (2018) point to the importance of students' interpretations of their lived experiences with math, which are influenced by their social environments as well as their internal narratives. Continued work on mediators and moderators can further elucidate the roles that external and internal factors play in math anxiety and the relationship between those factors and the effectiveness of interventions such as arousal reappraisal, leading to more tailored or combinatorial approaches to treatment and prevention. One such approach, as yet untested, has been suggested by Jamieson et al. (2018), who propose packaging interventions that cultivate a stress-is-enhancing general mindset with arousal reappraisal interventions to promote more comprehensive change in students' interpretations of stress.

Conclusion

To summarize, the online arousal reappraisal we tested was only effective in boosting math performance in a timed, working-memory intensive task, for the 22% percent of students with moderate or lower math anxiety. There was no evidence that online arousal reappraisal reduced math anxiety, raising additional questions regarding mechanisms for effectiveness. With its finding that the relationship between arousal reappraisals and math performance is moderated by math anxiety, this research contributes needed complexity to the literature about brief social-psychological interventions. It is our hope that mediators and moderators of these interventions continue to be investigated to best understand how, when, and for whom brief social-psychological interventions are effective and for whom they are contra-indicated, as more information regarding these questions is necessary before widespread dissemination and adoption is attempted.

References

Allen, E. I., & Seaman, J. (2017). *Digital learning compass: Distance education enrollment report*. Retrieved from https://onlinelearningsurvey.com/reports/digtiallearningcompassenrollment2017.pdf

Andrews, A., & Brown, J. (2015). The effects of math anxiety. *Education*, *135*(3), 362-370.

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General, 130*(2), 224–237. doi:10.1037/0096-3445.130.2.224

Ashcraft, M. H., & Krause, J. A. (2007). Working memory, math performance, and math anxiety. *Psychonomic Bulletin & Review, 14*(2), 243–248. doi:10.3758/bf03194059

Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate Behavioral Research, 40*(3), 373–400. doi:10.1207/s15327906mbr4003_5

Bonham, B. S., & Boylan, H. R. (2011). Developmental mathematics: Challenges, promising practices, and recent initiatives. *Journal of Developmental Education, 34*(3), 2.

Bourassa, A. (2014). *Numerical sequence recognition: Is familiarity or ordinality the primary factor in performance?* (Master's thesis, Carleton University, Ottawa, Canada). Retrieved from https://curve.carleton.ca/282e5db5-4f36-4e5f-93fa-7a47bbaf5640

Burke, L. (2021, April 22). *Charting a new normal*. Inside Higher Ed. https://www.insidehighered.com/news/2021/04/22/community-colleges-keeping-more-classes-online

Campbell, D. T., Stanley, J. C., & Gage, N. L. (1963). *Experimental and quasi-experimental designs for research.* Houghton Mifflin.

Cook, T.D. and Campbell, D.T. (1979) *Quasi-experimentation: Design and analysis issues for field settings*. Houghton Mifflin.

Cipora, K., Szczygieł, M., Willmes, K., & Nuerk, H.-C. (2015). Math anxiety assessment with the Abbreviated Math Anxiety Scale: Applicability and usefulness: Insights from the polish adaptation. *Frontiers in Psychology, 6*. doi:10.3389/fpsyg.2015.01833

Chatterjee, S., & Yilmaz, M. (1992). A review of regression diagnostics for behavioral research. *Applied Psychological Measurement, 16*(3), 209–227. doi:10.1177/014662169201600301

Chipman, S. F., Krantz, D. H., & Silver, R. (1992). Mathematics anxiety and science careers among able college women. *Psychological science, 3*(5), 292-296.

College Crisis Initiative. (2021). *C2i Dashboard*. https://collegecrisis.shinyapps.io/dashboard/

Cribari-Neto, F. (2004). Asymptotic inference under heteroskedasticity of unknown form. *Computational Statistics & Data Analysis, 45*(2), 215–233. doi:10.1016/s0167-9473(02)00366-3

Darlington, R. B., & Hayes, A. F. (2016). *Regression analysis and linear models: Concepts, applications, and implementation*. Guilford Publications.

Douglas, H. P., & LeFevre, J.A. (2018). Exploring the influence of basic cognitive skills on the relation between math performance and math anxiety. *Journal of Numerical Cognition, 3*(3), 642–666. doi:10.5964/jnc.v3i3.113

Dunn, A. M., Heggestad, E. D., Shanock, L. R., & Theilgard, N. (2016). Intra-individual response variability as an indicator of insufficient effort responding: Comparison to other indicators and relationships with individual differences. *Journal of Business and Psychology, 33*(1), 105–121. doi:10.1007/s10869-016-9479-0

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. doi:10.3758/bf03193146

Follette, K., McCarthy, D., Dokter, E., Buxner, S., … Prather, E. (2015). The Quantitative Reasoning for College Science (QuaRCS) Assessment, 1: Development and validation. *Numeracy, 8*(2). doi:10.5038/1936-4660.8.2.2

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression*. Sage publications.

Frazier, P. A., Tix, A. P., & Barron, K. E. (2004). Testing moderator and mediator effects in counseling psychology research. *Journal of Counseling Psychology, 51*(1), 115–134. doi:10.1037/0022-0167.51.1.115

Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science, 18*(3), 233–239. doi:10.1111/j.1467-9280.2007.01882.x

Hangen, E. J., Elliot, A. J., & Jamieson, J. P. (2018). Stress reappraisal during a mathematics competition: testing effects on cardiovascular approach-oriented states and exploring the moderating role of gender. *Anxiety, Stress, & Coping, 32*(1), 95–108. doi:10.1080/10615806.2018.1530049

Hayes, A. F. (2017a) *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Publications.

Hayes, A. F. (2017b) *Hacking PROCESS for estimation and probing of linear moderation of quadratic effects and quadratic moderation of linear effects*. Unpublished white paper. Retrieved from https://www.afhayes.com/public/quadratichack.pdf

Healthy Minds Network. (2021) *The Healthy Minds Study Fall 2020 Data Report*. https://healthymindsnetwork.org/wp-content/uploads/2021/02/HMS-Fall-2020-National-Data-Report.pdf

Hopko, D. R., Mahadevan, R., Bare, R. L., & Hunt, M. K. (2003). The Abbreviated Math Anxiety Scale (AMAS). *Assessment, 10*(2), 178–182. doi:10.1177/1073191103010002008

Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. (2012). Detecting and deterring insufficient effort responding to surveys. *Journal of Business and Psychology, 27(*1), 99–114. doi:10.1007/s10869-011-9231-8

Jaccard, J., Turrisi, R., & Wan, C. K. (1990). *Interaction effects in multiple regression.* Newbury Park, CA: Sage.

Jamieson, J. P., Crum, A. J., Goyer, J. P., Marotta, M. E., & Akinola, M. (2018). Optimizing stress responses with reappraisal and mindset interventions: an integrated model. *Anxiety, Stress, & Coping, 31*(3), 245–261. doi:10.1080/10615806.2018.1442615

Jamieson, J. P., Mendes, W. B., Blackstock, E., & Schmader, T. (2010). Turning the knots in your stomach into bows: Reappraising arousal improves performance on the GRE. *Journal of Experimental Social Psychology, 46*(1), 208–212. doi:10.1016/j.jesp.2009.08.015

Jamieson, J. P., Mendes, W. B., & Nock, M. K. (2013). Improving acute stress responses. *Current Directions in Psychological Science, 22*(1), 51–56. doi:10.1177/0963721412461500

Jamieson, J. P., Peters, B. J., Greenwood, E. J., & Altose, A. J. (2016). Reappraising stress arousal improves performance and reduces evaluation anxiety in classroom exam situations. *Social Psychological and Personality Science, 7*(6), 579–587. doi:10.1177/1948550616644656

John-Henderson, N. A., Rheinschmidt, M. L., & Mendoza-Denton, R. (2015). Cytokine responses and math performance: The role of stereotype threat and anxiety reappraisals. *Journal of Experimental Social Psychology, 56*, 203–206. doi:10.1016/j.jesp.2014.10.002

Kazelskis, R., Reeves, C., Kersh, M. E., Bailey, G., Cole, K., Larmon, M., … Holliday, D. C. (2000). Mathematics anxiety and test anxiety: Separate constructs? *The Journal of Experimental Education, 68*(2), 137–146. doi:10.1080/00220970009598499

Kelley, K. (2005). The effects of nonnormal distributions on confidence intervals around the standardized mean difference: Bootstrap and parametric confidence intervals. *Educational and Psychological Measurement, 65*(1), 51–69. doi:10.1177/0013164404264850

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams Jr, R. B., Alper, S., ... & Sowden, W. (2018). Many Labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science, 1*(4), 443-490. doi:10.1177/2515245918810225

Ma, X. (1999). A meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *Journal for Research in Mathematics Education, 30*(5), 520. doi:10.2307/749772

Ma, X., & Willms, J. D. (1999). Dropping out of advanced mathematics: How much do students and schools contribute to the problem? *Educational Evaluation and Policy Analysis, 21*(4), 365–383. doi:10.3102/01623737021004365

Maloney, E. A., Schaeffer, M. W., & Beilock, S. L. (2013). Mathematics anxiety and stereotype threat: shared mechanisms, negative consequences and promising interventions. *Research in Mathematics Education, 15*(2), 115–128. doi:10.1080/14794802.2013.797744

Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women’s intellectual performance. *Journal of Experimental Social Psychology, 42*(2), 236–243. doi:10.1016/j.jesp.2005.04.010

McKenzie L. (2021, April 27). *Students want online learning options post-pandemic*. Inside Higher Ed. https://www.insidehighered.com/news/2021/04/27/survey-reveals-positive-outlook-online-instruction-post-pandemic

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods, 17*(3), 437–455. doi:10.1037/a0028085

Noel-Levitz, Inc. (2006). *Student success in developmental math: Strategies to overcome barriers to retention*. Iowa City, IA: Noel-Levitz. Retrieved from https:// www.noellevitz. com/NR/rdonlyres/B4148B72-C135- 4AD4-A04C-2F66821C872C/0/ENABLEMATH_ paper_0706indd.pdf

Primi, C., Busdraghi, C., Tomasetto, C., Morsanyi, K., & Chiesi, F. (2014). Measuring math anxiety in Italian college and high school students: Validity, reliability and gender invariance of the Abbreviated Math Anxiety Scale (AMAS). *Learning and Individual Differences, 34*, 51–56. doi:10.1016/j.lindif.2014.05.012

Ramirez, G., Shaw, S. T., & Maloney, E. A. (2018). Math anxiety: Past research, promising interventions, and a new interpretation framework. *Educational Psychologist, 53*(3), 145–164. doi:10.1080/00461520.2018.1447384

Rogosa, D. (1980). Comparing nonparallel regression lines. *Psychological Bulletin*, *88*(2), 307. doi:10.1037/0033-2909.88.2.307

Rounds, J. B., & Hendel, D. D. (1980). Measurement and dimensionality of mathematics anxiety. *Journal of Counseling Psychology*, *27*(2), 138–149. doi:10.1037/0022-0167.27.2.138

Russell, C. J., & Dean, M. A. (2000). To log or not to log: Bootstrap as an alternative to the parametric estimation of moderation effects in the presence of skewed dependent variables. *Organizational Research Methods, 3*(2), 166–185. doi:10.1177/109442810032002

Schillinger, F. L., Vogel, S. E., Diedrich, J., & Grabner, R. H. (2018). Math anxiety, intelligence, and performance in mathematics: Insights from the German adaptation of the Abbreviated Math Anxiety Scale (AMAS-G). *Learning and Individual Differences, 61*, 109–119. doi:10.1016/j.lindif.2017.11.014

Schoemann, A. M., Boulton, A. J., & Short, S. D. (2017). Determining power and sample size for simple and complex mediation models. *Social Psychological and Personality Science, 8,* 379-386.

Shadish, W., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3/4), 591. doi:10.2307/2333709

Shi, Z., & Liu, P. (2016). Worrying thoughts limit working memory capacity in math anxiety. *PLOS ONE, 11*(10), e0165644. doi:10.1371/journal.pone.0165644

Solomon, R. L. (1949). An extension of control group design. *Psychological bulletin*, *46*(2), 137.

Son, C., Hegde, S., Smith, A., Wang, X., & Sasangohar, F. (2020). Effects of COVID-19 on college students' mental health in the United States: Interview survey study. *Journal of Medical Internet Research, 22*(9), e21279. doi:10.2196/21279

Sowinski, C., Dunbar, K., & LeFevre, J. (2014). *Calculation Fluency Test* (Unpublished technical report). Math Lab, Carleton University, Ottawa, Canada.

Steiner, E. T., & Ashcraft, M. H. (2012). Three brief assessments of math achievement. *Behavior Research Methods, 44*(4), 1101–1107. doi:10.3758/s13428-011-0185-6

Trusty, J., & Niles, S. G. (2003). High-school math courses and completion of the bachelor's degree. *Professional School Counseling*, 99-107.

U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), (2021), Retrieved from https://nces.ed.gov/ipeds/use-the-data

Wilkinson, G. S. (1993). *Manual for the Wide Range Achievement Test* (3rd ed.). Wilmington, DE: Wide Range.

Wood, M. (2005). Bootstrapped confidence intervals as an approach to statistical inference. *Organizational Research Methods*, *8*(4), 454–470. doi:10.1177/1094428105280059

APPENDICES

Appendix A: Participant Measures

Demographics

1. What is your sex?
    a. Male
    b. Female
2. What is your gender?
    a. Male
    b. Female
    c. Transgender
    d. Other, please specify
    e. Decline to answer
3. What is your age? (write the number)
4. Please indicate your **Hispanic Origin**:
    a. Hispanic or Latino
    b. Not Hispanic or Latino
5. Please indicate your race, the specific group(s) that you identify with the most (you can select more than one):
    a. White
    b. Black or African American
    c. American Indian or Alaska Native
    d. Asian
    e. Native Hawaiian or Other Pacific Islander
    f. Other, please specify
6. What is your country of birth?
    a. United States
    b. Other
7. What is your current marital status?
    a. Single
    b. Married/Partnered
    c. Separated
    d. Divorced
    e. Widowed
    f. Other
8. What is your academic major?
9. What interest area is your major in?
    a. Business
    b. Education
    c. Health Professions
    d. Humanities & Arts
    e. Law
    f. Policy/Social Science
    g. STEM

Appendix B: Participant Measures

Abbreviated Math Anxiety Scale

Test Format: Each of the measure's 9 items are responded to using a 5-point Likert-type scale, ranging from 1 (low anxiety) to 5 (high anxiety), with the total score representing a summation of the nine items. Items 1, 3, 6, 7, and 9 are summed to score the Learning Math Anxiety factor. Items 2, 4, 5, and 8 are summed to score the Math Evaluation Anxiety factor.

1. Having to use the tables in the back of a math book.
2. Thinking about an upcoming math test 1 day before.
3. Watching a teacher work an algebraic equation on the blackboard.
4. Taking an examination in a math course
5. Being given a homework assignment of many difficult problems that is due at the next class meeting.
6. Listening to a lecture in math class.
7. Listening to another student explain a math formula.
8. Being given a “pop” quiz in math class.
9. Starting a new chapter in a math book.

Appendix C: Participant Measures
Calculation Fluency Test Pre-test Items

Addition Part 1 (30 seconds)

28 +13	51 +10	42 +53	71 +11	95 +52	74 +38	14 +19	99 +63	57 +83	17 +39
19 +27	40 +44	67 +38	98 +59	42 +13	17 +19	90 +82	45 +91	55 +58	83 +42
98 +31	34 +22	20 +54	63 +92	40 +59	26 +89	18 +39	27 +36	44 +80	88 +77

Subtraction Part 1 (30 seconds)

89 -60	52 -48	60 -39	51 -28	85 -23	18 -11	49 -37	83 -57	42 -23	68 -47
52 -19	91 -23	60 -31	42 -31	94 -45	98 -64	50 -33	53 -19	61 -45	41 -27
39 -23	61 -37	29 -19	31 -14	54 -12	92 -65	60 -43	43 -27	70 -31	94 -24

Multiplication Part 3 (30 seconds)

73	41	69	29	16	63	60	52	85	36
x 8	x 5	x 3	x 9	x 8	x 8	x 4	x 4	x 6	x 7
52	98	41	19	15	49	71	30	48	81
x 9	x 3	x 8	x 6	x 4	x 2	x 9	x 8	x 7	x 5
45	32	79	37	19	52	17	47	39	78
x 9	x 6	x 2	x 8	x 9	x 6	x 5	x 2	x 3	x 7

Appendix D: Participant Measures
Calculation Fluency Test Post-test Items

Addition Part 1 (30 Seconds)

47	23	41	47	59	23	87	31	38	34
+17	+48	+53	+85	+16	+18	+58	+53	+49	+78
41	86	58	25	86	29	74	34	15	83
+38	+93	+34	+77	+55	+22	+31	+19	+26	+19
37	13	38	51	78	89	34	56	23	47
+98	+87	+67	+65	+45	+32	+65	+45	+43	+39

Subtraction Part 2 (30 Seconds)

48	42	95	81	40	51	42	97	93	74
-19	-31	-65	-62	-31	-27	-18	-18	-45	-23
65	36	80	51	82	91	68	75	64	42
-39	-22	-46	-27	-31	-64	-59	-34	-24	-37
99	61	91	82	25	73	58	57	59	31
-45	-27	-60	-47	-19	-45	-32	-17	-42	-27

Multiplication Part 3 (30 seconds)

14	38	50	80	61	52	97	72	16	49
x 3	x 8	x 2	x 9	x 7	x 7	x 4	x 7	x 5	x 6
52	71	96	47	83	16	44	50	62	39
x 7	x 2	x 8	x 6	x 3	x 8	x 3	x 8	x 6	x 5
13	68	75	67	45	94	52	83	61	54
x 8	x 4	x 5	x 8	x 9	x 7	x 6	x 7	x 6	x 9

Appendix E: Participant Measures
Brief Math Assessment-3 Pre-test items

$\begin{array}{r} 37 \\ \underline{-16} \end{array}$	$\begin{array}{r} 67 \\ \underline{+25} \end{array}$	$\begin{array}{r} 6 \\ \underline{\times 3} \end{array}$	$\frac{8}{4} =$ ____

$7\frac{1}{2} + 1\frac{1}{2} =$	$\begin{array}{r} 2\frac{1}{3} \\ 5\frac{1}{6} \\ \underline{+\ 1\frac{1}{9}} \end{array}$	$\begin{array}{r} 5\frac{2}{5} \\ \underline{-3\frac{3}{4}} \end{array}$
Write as a common fraction in lowest terms: 0.75 = _______	$3x - y = 22$ $2x - y = 13$ $x =$ _____ $y =$ _____	Reduce: $\frac{2x + x}{x^2 - 1} \bullet \frac{x^2 - x}{x^2}$

Appendix F: Participant Measures
Brief Math Assessment-3 Post-test items

1.	2.	3.	4.
$\begin{array}{r} 42 \\ -21 \\ \hline \end{array}$	$\begin{array}{r} 56 \\ +17 \\ \hline \end{array}$	$\begin{array}{r} 8 \\ \times 5 \\ \hline \end{array}$	$\frac{9}{3} = \_\_\_\_\_$

5.	6.	7.
$3\frac{1}{2} \ + \ 2\frac{1}{2} \ = \ \_\_\_\_\_$	$\begin{array}{r} 4\frac{1}{4} \\ 3\frac{1}{8} \\ +\ 2\frac{1}{2} \\ \hline \end{array}$	$\begin{array}{r} 8\frac{1}{4} \\ -5\frac{2}{3} \\ \hline \end{array}$

8.	9.	10.
Write as a common fraction in lowest terms: .025 = _____	$5j - w = 18$ $4j - w = 14$ $j =$ _____ $w =$ _____	Reduce: $\frac{P^2 + P}{P^2} \bullet \frac{2P - 2}{P^2 - 1}$ Answer: ____

Appendix G: Participant Measures
Intervention

Thank you for being a part of this study!

As you might expect, taking a math assessment can be a very stressful experience. Before starting your assessment, we are going complete a short reading exercise designed to help you perform well. In the following pages, you will be presented with summaries of scientific studies.

After reading each study, we ask that you please write a short summary in your own words describing how the information presented can help you on today's exam.

This excerpt is adapted from Jamieson & Mendes' 2010 study that appeared in the *Journal of Experimental Psychology*

In stressful situations, people experience changes in their body. They might experience these as “unsettled feelings” or “butterflies in their stomach,” and conclude that they are nervous. However, bodily changes that happen during stress can be good. For instance, scientists have found that feelings of “butterflies” indicate that the body is gathering resources to meet situational demands. In other words, the body needs energy to perform and stress helps deliver this energy to your brain.

Stress can be “good” or “bad,” and depends on our perceptions and beliefs. For example, imagine you are a skier staring down a steep slope with no other way off the mountain than going down this dangerous trail. Regardless of whether you like skiing, this situation is stressful. Expert skiers experience the stress as “excitement” because they believe they can handle the difficult trail, whereas novices experience the stress as “fear” because the difficult trail exceeds their skill level. Thus, the skier’s response (excitement vs. fear) depends on how they perceive stress.

Research from students with anxiety indicates that stress does not hurt performance, but can actually help because our brain releases chemicals that help us think quickly. So, during the test today, try to view your own stress as a coping tool.

In your own words please briefly describe how this information can help you perform well on your assessment today:

____

____

____

____

This excerpt is adapted from Nock et al.'s (2011) study that appeared in the *Journal of Clinical Psychology*

Stress is a normal reaction that helps you face the challenges in your life. It is not harmful. In fact, if we did not have stress reactions we could not survive. If stress is helpful, then why do most people see it as a negative experience?

Research indicates that negative reactions to stressful situations like taking an exam are the result of how we think about stress (also known as 'cognitive appraisals'). When the "*fight or flight"* system activates, our brain searches for possible sources of harm. However, in modern society there is often no physical threat. When no explanation can be found, the brain invents explanations such as, "There must be something wrong with me." Nothing could be further from the truth. Stress is adaptive and good.

During stressful situations remember that your body's responses are beneficial. Increased heart rate, sweating, and heavy breathing are all signs that your body is delivering oxygen (fuel for thinking) to your brain.

In your own words please briefly describe how this information can help you perform well on your assessment today:

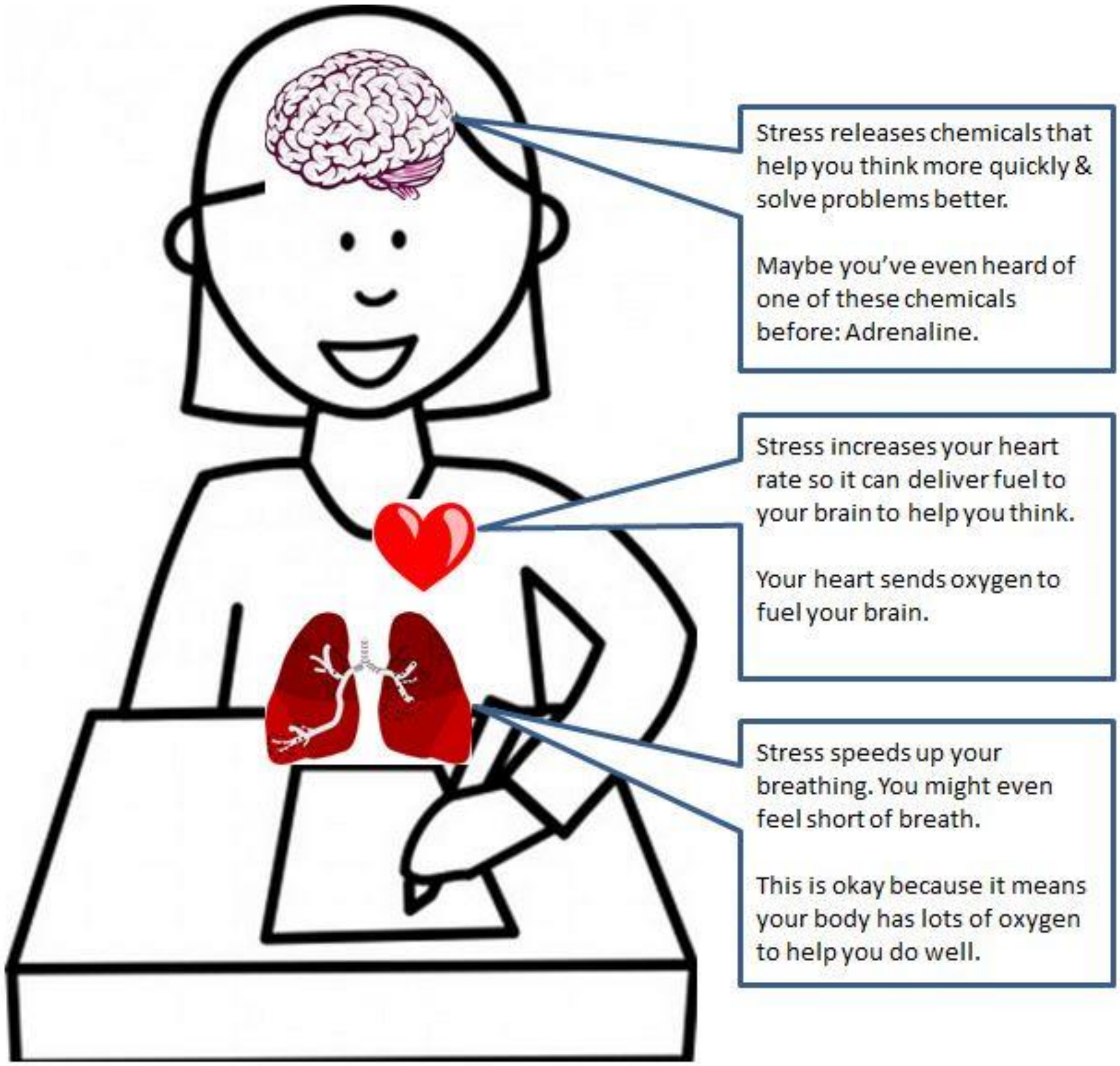
____

____

____

____

The following is an illustrative diagram that shows the biological changes that happen when we experience stress. Please take a minute to note where the changes occur and how these help us do well.



In your own words please briefly describe how this information can help you perform well on your assessment today:

____

____

____

Great job! You have finished the exercise. We would now like for you to answer some brief questionnaires before starting your assessment.

Remember during the assessment today, we ask that you try to remind yourself that your body's responses to the stressful testing situation will help you to perform well.

Good luck!

Appendix H: Participant Measures
Control

Thank you for being a part of this study!

As you might expect, taking a math test can be a very stressful experience. Before starting your exam, we are going complete a short reading exercise designed to help you perform well. In the following pages, you will be presented with summaries of scientific studies.

After reading each study, we ask that you please write a short summary in your own words describing how the information presented can help you on today's exam.

This excerpt is adapted from Martin & Fahrig's 2011 article that appeared in *Ecology*

Almost every part of a bird's anatomy has evolved in some part to enhance flight. Birds must be lightweight to fly, so have evolved very lightweight hollow bones. The structure of their bones resembles honeycomb, making them very strong but also very light. For example, frigate birds have a wingspan of over two meters, but the skeleton weighs about 4 ounces. Birds also have fewer organs (e.g. only one ovary) and no teeth. Birds use a digestive organ called a gizzard to grind up food.

Birds maintain higher body temperatures than mammals, about 104 degrees Fahrenheit. This enables cells in their muscles to work around 2.2 times faster, and allows muscles to relax more rapidly. The higher body temperature is enabled through the insulating properties of feathers, and in some species, a layer of fat.

Birds have highly efficient respiratory and circulatory systems which keep their tissues well supplied with oxygen and nutrients, supporting a high metabolic rate. Bird lungs are full of elastic air sacs that help to dissipate heat and reduce the density of their bodies. The eyesight of birds is said to be the best of all vertebrates. Excellent eyesight and coordination helps birds to fly safely.

In your own words, using the above information, please briefly describe how the anatomy of birds has evolved to help them fly:

____

____

____

____

This excerpt is adapted from Campbell and Reese (2002) article that appeared in *Nature*

Despite the popular expression “bird-brained” being used to describe somebody who lacks intelligence, birds’ brains are proportionately larger than those of reptiles and amphibians (their closest living relatives), and research has shown that birds are capable of very complex behavior. Indeed, some birds migrate over 12,500 miles per year, without any of the satellite navigation devices we use to find our way.

The most obvious adaptations for flight are bird’s wings. To flap their wings and provide power for flight, birds contract their large pectoral (breast) muscles which are anchored to a keel on their sternum (breastbone). Many birds, such as birds of prey use air currents to soar and glide, whereas birds such as hummingbirds must flap continuously to hover while feeding. Some birds have evolved into flightless birds, in the absence of natural predators or in the case of penguins, to enable them to swim. In all birds capable of flight, it is the shape and arrangement of feathers which enables them to create lift with their wings.

In your own words, using the above information, please briefly describe how the anatomy of birds has evolved to help them fly:

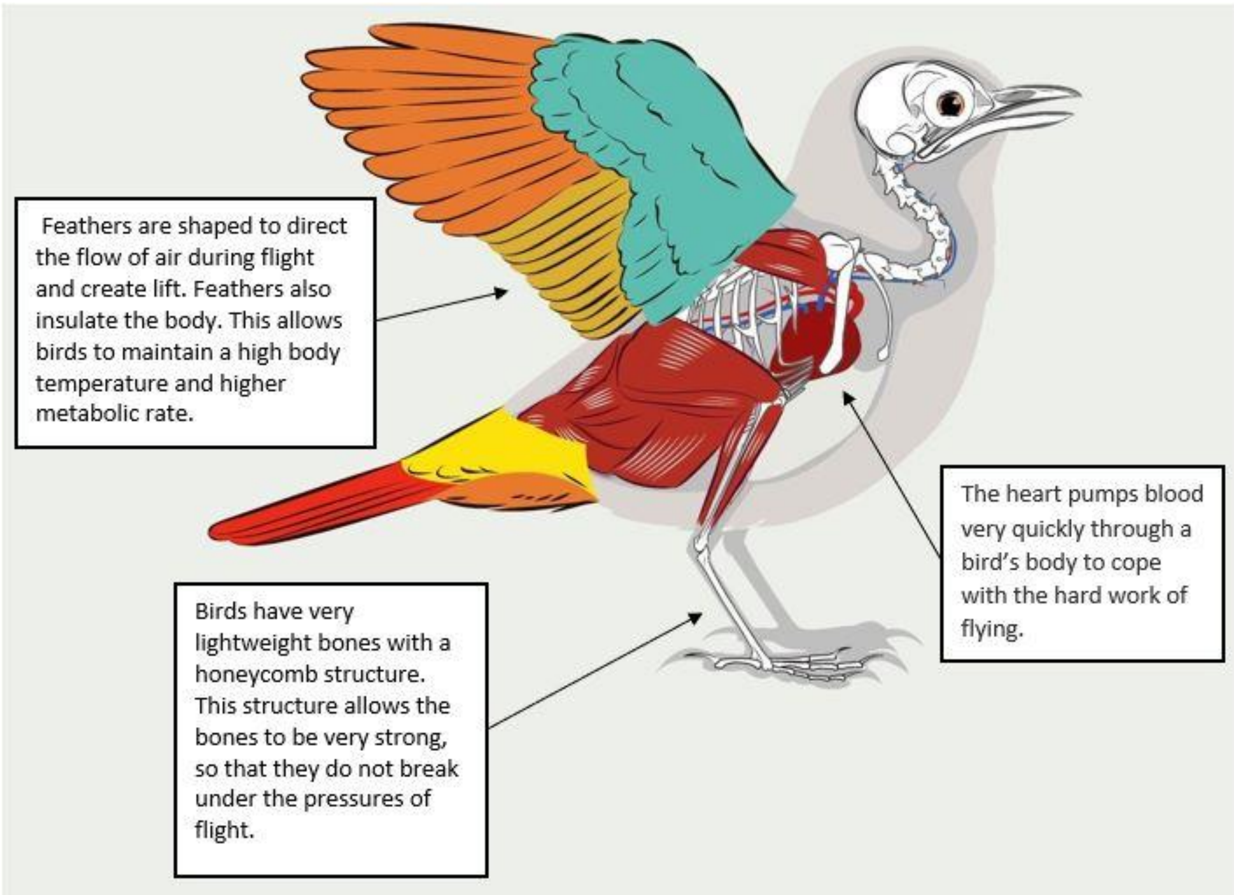
____

____

____

____

The following is an illustrative diagram that shows the biological adaptations that contribute to birds' ability to fly. Please take a minute to note where the changes occur and how these contribute to flight.



In your own words, using the above information, please briefly describe how the anatomy of birds has evolved to help them fly:

____

____

____

____

Great job! You have finished the exercise. We would now like for you to answer some brief questionnaires before starting your test.

Good luck!

Table 2.1

Descriptive Statistics and Bivariate Correlations between Math Anxiety and Math Performance Across Condition

Measure	LMA T1	LMA T2	MEA T1	MEA T2	BMA-3 T1	BMA-3 T2	CFT T1	CFT T2
LMA T1	-	0.76**	0.65**	0.56**	-0.22*	-0.22**	-0.24**	-0.18*
LMA T2	0.78**	-	0.65**	0.65**	-0.20*	-0.20*	-0.28**	-0.27**
MEA T1	0.46**	0.43**	-	0.87**	-0.17	-0.12	-0.17*	-0.16
MEA T2	0.42**	0.50**	0.80**	-	0.01	-0.10	-0.21*	-0.18*
BMA-3 T1	-0.08	-0.12	-0.07	-0.01	-	0.66**	0.27**	0.26**
BMA-3 T2	-0.17	-0.26**	0.00	0.05	0.59**	-	0.37**	0.43**
CFT T1	-0.01	-0.12	0.00	-0.04	0.26**	0.14	-	0.76**
CFT T2	-0.10	-0.14	-0.23**	-0.15	0.24**	0.23**	0.66**	-
Control[a]								
M	1.89	1.98	3.45	3.45	6.58	7.17	9.27	10.45
SD	0.82	0.89	1.13	1.10	1.59	1.86	4.24	4.55
a	0.83	0.89	0.90	0.91	0.58	0.69	0.82	0.83
Intervention[b]								
M	1.78	1.83	3.58	3.61	6.50	7.33	9.75	11.17
SD	0.85	0.74	1.06	1.02	1.52	1.57	3.84	4.50
a	0.85	0.81	0.86	0.86	0.49	0.51	0.79	0.84

Note. T1 = Time 1; T2 = Time 2; LMA = Learning Math Anxiety; MEA = Math Evaluation Anxiety; BMA-3 = Brief Math Assessment-3; CFT = Calculation Fluency Test.

[a] Control only (above the diagonal) N = 137.

[b] Re-appraisal only (below the diagonal) N = 131.

*Correlation is significant at the 0.05 level.

**Correlation is significant at the 0.01 level.

Table 2.2

Model Coefficients, Standard Errors, and 95% Bootstrap Confidence Intervals

Predictor	*B*	*SE*	Bootstrap 95% CI
Cond → BMA T2 (R^2 = 0.39)**			
Condition	0.18	0.16	-0.13 to 0.49
BMA T1	0.67**	0.05	0.57 to 0.78
Cond → CFT T2 (R^2 = 0.51)**			
Condition	0.29	0.39	-0.44 to 1.07
CFT T1	0.79**	0.06	0.68 to 0.88
Cond → LMA T2 (R^2 = 0.40)**			
Condition	-0.07	0.07	-0.20 to 0.05
LMA T1	0.75**	0.05	0.67 to 0.84
Cond → MEA T2 (R^2 = 0.40)**			
Condition	0.05	0.07	-0.09 to 0.19
MEA T1	0.81**	0.03	0.75 to 0.87

Note. T1 = Time 1; T2 = Time 2; LMA = Learning Math Anxiety; MEA = Math Evaluation Anxiety; BMA-3 = Brief Math Assessment-3; CFT = Calculation Fluency Test; CI = confidence interval. Unstandardized regression coefficients are reported.

*Indicates significance at the 0.05 level.

**Indicates significance at the 0.01 level.

Table 2.3

Moderation Model Coefficients, Standard Errors, and 95% Bootstrap Confidence Intervals

Predictor	*B*	*SE*	Bootstrap 95% CI
Cond X LMA T1 → BMA T2 (R^2 = 0.40)**			
Condition	0.21	0.40	-0.56 to 1.00
LMA T1	-0.18	0.14	-0.43 to 0.08
Condition X LMA T1	-0.03	0.21	-0.42 to 0.37
BMA T1	0.66**	0.05	0.55 to 0.75
Cond X MEA T1 → BMA T2 (R^2 = 0.40)**			
Condition	-0.72	0.50	-1.68 to 0.28
MEA T1	-0.19*	0.09	-0.37 to -0.01
Condition X MEA T1	0.26	0.14	-0.25 to 0.54
BMA T1	0.67**	0.05	0.57 to 0.77
Cond X LMA T1 → CFT T2 (R^2 = 0.52)**			
Condition	1.11	0.99	-0.64 to 3.05
LMA T1	-0.04	0.34	-0.65 to 0.62
Cond X LMA T1	-0.44	0.54	-1.46 to 0.47
CFT T1	0.80**	0.06	0.69 to 0.90
Cond X MEA T1 → CFT T2 (R^2 = 0.54)**			
Condition	3.25*	1.56	0.50 to 6.42
MEA T1	-0.15	0.23	-0.57 to 0.29
Cond X MEA T1	-0.81*	0.41	-1.63 to -0.08
CFT T1	0.79**	0.05	0.69 to 0.89

Note. T1 = Time 1; T2 = Time 2; LMA = Learning Math Anxiety; MEA = Math Evaluation Anxiety; BMA-3 = Brief Math Assessment-3; CFT = Calculation Fluency Test; CI = confidence interval. Unstandardized regression coefficients are reported.

*Indicates significance at the 0.05 level.

**Indicates significance at the 0.01 level.

Table 2.4

Mediation Model Coefficients, Standard Errors, and 95% Bootstrap Confidence Intervals

	(X) Cond (M) LMA T2 (Y) BMA T2		(X) Cond (M) MEA T2 (Y) BMA T2		(X) Cond (M) LMA T2 (Y) CFT T2		(X) Cond (M) MEA T2 (Y) CFT T2	
Predictor	*B*(*SE*)	95% CI	*B*(*SE*)	95% CI	*B*(*SE*)	95% CI	*B*(*SE*)	95% CI
Math Anxiety	**R^2 = 0.60****		**R^2 = 0.70****		**R^2 = 0.60****		**R^2 = 0.71****	
Condition	-.08 (.07)	-.22 to .04	.04 (.07)	-.11 to .18	-.06 (.06)	-.19 to .06	.05 (.07)	-.08 to .20
LMA T1	-.33 (.17)*	-.06 to .01	-	-	.74 (.05)**	.65 to .82	-	-
MEA T1	-	-	.81 (.03)**	.74 to .87	-	-	.81 (.03)**	.74 to .87
BMA T1	.65 (.05)**	.66 to .84	.02 (.02)**	-.02 to .06	-	-	-	-
CFT T1	-	-	-	-	-.02 (.01)**	-.04 to -.01	-.01 (.01)	-.03 to .00
Performance	**R^2 = 0.41****		**R^2 = 0.39****		**R^2 = 0.52****		**R^2 = 0.53****	
Condition	.13 (.16)	-.17 to .44	.18 (.16)	-.12 to .50	.29 (.39)	-.45 to 1.07	.38 (.39)	-.35 to 1.14
LMA T2	-.33 (.17)*	-.64 to .00	-	-	-.30 (.35)	-.95 to .39	-	-
MEA T2	-	-	-.05 (.15)	-.35 to .22	-	-	.52 (.39)	-.18 to 1.33
BMA T1	.65 (.05)**	.54 to .75	.67 (.05)**	.57 to .77	-	-	-	-
CFT T1	-	-	-	-	.78 (.06)**	.67 to .89	.79 (.05)	.69 to .89
LMA T1	.05 (.16)	-.26 to .36	-	-	-.05 (.36)	-.72 to .61	-	-
MEA T1	-	-	-.03 (.14)	-.29 to .25	-	-	-.95 (.44)	-1.83 to -.17
Indirect Effect	.03 (.03)	-.01 to .10	.00 (.01)	-.03 to .02	.02 (.04)	-.03 to .11	.03 (.05)	-.05 to .17

Note. T1 = Time 1; T2 = Time 2; LMA = Learning Math Anxiety; MEA = Math Evaluation Anxiety; BMA-3 = Brief Math Assessment-3; CFT = Calculation Fluency Test; CI = confidence interval. Unstandardized regression coefficients are reported.

*Indicates significance at the 0.05 level.

**Indicates significance at the 0.01 level.

Figure 2.1

Time 2 CFT Scores by Condition Predicted by Time 1 Math Evaluation Anxiety

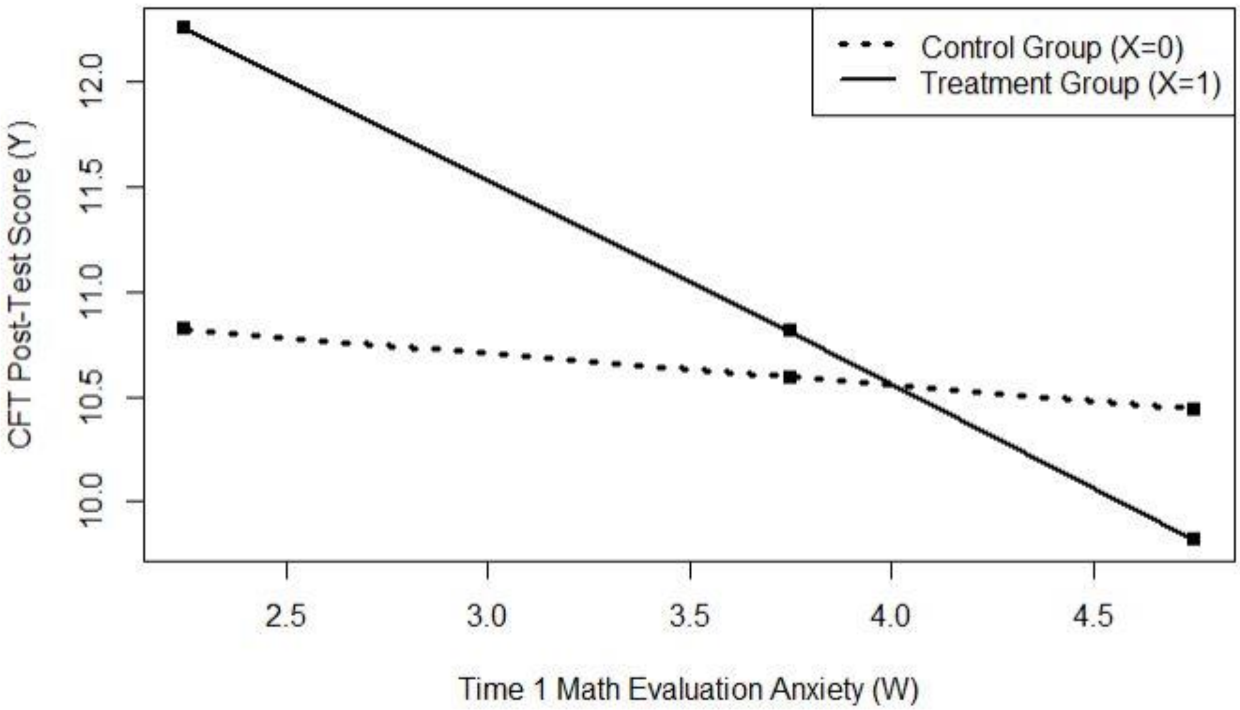


Figure 2.2

Johnson-Neyman Plot of Conditional Effect of Intervention on Time 2 CFT Scores