Title

A Dissemination and Implementation Approach to Preventing Anxiety and Depression in Young People

Author

Rith-Najarian, Leslie

Publication Date

2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

A Dissemination and Implementation Approach to Preventing Anxiety and Depression in Young People

A dissertation submitted in partial satisfaction

of the requirements for the degree Doctor of Philosophy

in Psychology

by

Leslie Rose Rith-Najarian

2019

ABSTRACT OF THE DISSERTATION

A Dissemination and Implementation Approach

to Preventing Anxiety and Depression in Young People

by

Leslie Rith-Najarian

Doctor of Philosophy in Psychology

University of California, Los Angeles, 2019

Professor Denise Chavira, Co-Chair

Professor Bruce Chorpita, Co-Chair

Online prevention programs for anxiety and depression have great potential to alleviate the two most prevalent mental health concerns in university students. **Chapter 1** presents a systematic review of existing prevention programs for anxiety, stress, and depression in university students. Results showed that although these programs on average produce moderate effects for symptom change and have common practice elements, they are limited by non-representative samples and variable adherence rates. Informed by findings of the systematic review, an original online anxiety and depression prevention program was developed for universal delivery with university students. The intervention's implementation was tested and iteratively modified through subsequent research studies, as overviewed in Chapters 2, 3, and 4. **Chapter 2** presents two studies that collected observational data during intervention recruitment phases in order to

examine the impact of marketing strategies for the online intervention. Replicated findings showed that recruitment and branding strategies successfully engaged traditionally underserved students (e.g., male students, Asian-identifying students). **Chapter 3** presents an open trial study that collected quantitative and qualitative data to assess feasibility of the intervention's implementation and research procedures. Findings supported feasibility in terms of recruitment strategies, participant adherence/retention, program acceptability, and pre-post symptom change assessment procedures, as well as qualitative themes about participant experiences. **Chapter 4** presents a randomized controlled trial (RCT) that examined intervention effectiveness and moderators of symptom change. Results showed that the intervention condition outperformed a waitlist for improvements in depression, anxiety, and stress. Symptom improvement effects were replicated by the waitlist group, maintained through 3-month follow-up, and moderated by internal motivation for treatment. Overall, findings from all dissertation studies are discussed in terms of provided insights about how we can attract, engage, and ultimately help more young people. The **Dissertation Discussion** chapter considers overall takeaways, strengths, and future directions of the presented dissertation research.

The dissertation of Leslie Rith-Najarian is approved.

Elizabeth Gong-Guy

Michelle Craske

Robert M Bilder

Denise A Chavira, Co-Chair

Bruce Frederick Chorpita, Co-Chair

University of California, Los Angeles

2019

Table of Contents

Acknowledgements

From the conception of this dissertation to its finish, I have received a tremendous amount of support and encouragement from so many individuals.

Thank you to my dissertation co-chairs, Dr. Denise Chavira and Dr. Bruce Chorpita. Denise's practical guidance and wisdom have grounded me at times when I encountered research obstacles and her feedback helped improve my research in quality and clarity. Bruce's ability to help me see the "big picture" of projects brought more meaning to my work and the innovations he has developed provided me with valuable tools. Thank you to Dr. Liz Gong-Guy, an incredibly dedicated clinical supervisor for my dissertation studies, but also a mentor and role model in helping me develop my own professional identity. Thank you to my whole dissertation committee for the time they have dedicated and thoughtful feedback they have provided.

I am incredibly grateful to the family and friends who have provided me with a steady stream of emotional support, creative ideas, new perspectives, and inspiration. Thank you to my loving husband and life partner, Eric Kuenstner, who has balanced extra responsibilities at my busiest times while still completing his own dissertation research. Eric has listened to so many practice presentations, read so many drafts, and provided so much invaluable feedback that he probably could give his own version of my dissertation defense at this point. Thank you to my parents, Dr. Janet and Dr. Steve Rith-Najarian, for raising me to do good in this world and to ask a lot of questions. Their unwavering enthusiasm for my endeavors and inquisitive conversations have heavily fueled my passion. Thank you to my grandparents – Stephen and Christine Najarian and Dr. Donald and Joan Rith – for making higher education a bedrock of our families. Thank you to my friends, in particular Bita Mesri, Ben Smith, and Alayna Park, for their willingness to supportively listen and talk through whatever latest dissertation challenges I encountered. Thank

you to Jenn Gamarra, Resham Gellatly, Julia Hammet, and Bita Mesri for collectively conducting the suicide risk follow-up protocol which served hundreds of students.

There are still so many other incredible friends, family members, supervisors, research assistants, and peers who provided support that directly and indirectly made this dissertation possible. I am deeply grateful and have been very fortunate to have so many amazing individuals who have believed in me and helped me along my journey.

Co-authored Work in Publication or Submitted for Publication

Chapter 1 is reproduced with permission from the published article: "Rith-Najarian, L., Boustani, M., & Chorpita, B.F. (2019). A systematic review of prevention programs targeting anxiety, depression, and stress in university students. *Journal of Affective Disorders, 257*, 568-584. https://doi.org/10.1016/j.jad.2019.06.035". Dr. Boustani contributed to data collection and manuscript writing. Dr. Chorpita contributed to methodology design and writing feedback.

Chapter 2 is reproduced with permission from the published article: "Rith-Najarian, L., Sun, W., Chen, A., Chorpita, B., Chavira, D., Mougalian, S., & Gong-Guy, E. (2019). What's in a name? Branding of online mental health programming for university students. *Journal of Consulting and Clinical Psychology, 87*, 380-391. http://dx.doi.org/10.1037/ccp0000383". Ms. Sun and Ms. Chen oversaw data collection at the research site for Study 2 and contributed to the manuscript writing. Dr. Chorpita and Dr. Mougalian served as faculty sponsors and provided writing feedback. Dr. Gong-Guy served as the co-principal investigator.

Funding

This dissertation research received support in part from:

- University of California Los Angeles, through the: Life Sciences Innovation Fund Grant, Mangasar M Mangasarian Fellowship, Healthy Campus Initiative Student Grant, and Graduate Division Dissertation Year Fellowship;
- American College Health Foundation, through the Weiss Student Mental Health Award;
- JP Morgan, through the Dolores Zohrab Liebmann Fellowship; and
- American Psychological Association, through the Dissertation Research Award.

The content of this dissertation does not necessarily represent the official views of any of the funding sponsors.

Vita/Biographical Sketch

EDUCATION

M.A. awarded 2014 **University of California - Los Angeles,** Los Angeles, CA
Clinical Psychology

B.A. awarded 2012 **Harvard University,** Cambridge, MA
Honors: Magna cum laude, *Major:* Psychology, *Minor*: Germanic Studies

RELEVANT PAID POSITIONS

2018 - 19 **Clinical Psychology Intern,** Counseling and Psychological Services, UCLA
2013 - 19 **Statistical Consultant,** ghSMART & Company, Inc
2016 - 18 **Teaching Assistant,** UCLA Department of Psychology
2016 **Graduate Student Researcher,** UCLA Depression Grand Challenge
2012 - 13 **Research Coordinator,** Stress & Development Lab, Boston Children's Hospital

PEER-REVIEWED PUBLICATIONS

- **Rith-Najarian, L.,** Boustani, M., & Chorpita, B.F. (2019). A systematic review of prevention programs targeting anxiety, depression, and stress in university students. *Journal of Affective Disorder, 257*, 568–584.
- **Rith-Najarian, L.,** Sun, W., Chen, A., Chorpita, B.F., Chavira, D.A., Mougalian, S.S., & Gong-Guy, E. (2019). What's in a name?: Branding of online mental health programming for university students. *Journal of Consulting and Clinical Psychology, 87,* 380-391.
- **Rith-Najarian, L.,** Mesri, B., Park, A. L., Sun, M., Chavira, D. A., & Chorpita, B. F. (2019). Durability of cognitive behavioral therapy effects for youth and adolescents with anxiety, depression, or traumatic stress: A meta-analysis on long-term follow-ups. *Behavior Therapy, 50,* 225-240.
- Sun, M., **Rith-Najarian, L.**, Williamson, T., & Chorpita, B.F. (2019). Treatment features associated with youth cognitive behavioral therapy follow-up effects: meta-analysis. *Journal of Clinical Child and Adolescent Psychology, 48,* S269-S283.
- **Rith-Najarian, L**., Park, A., Wang, T., Etchison, A., Chavira, D., & Chorpita, B.F. (2017). Applying New Evidence Standards to Youth Cognitive Behavioral Therapies – A Review. *Behavioral Research and Therapy, 90,* 147-158.
- Ebesutani, C., Daleiden, E., Becker, K., Schmidt, L., Bernstein, A., **Rith-Najarian, L**., ... Chorpita, B. (2017). Facilitating communication of ideas and evidence to enhance mental health service quality. *Journal of Child and Family Studies, 26*, 3123-3134.
- Guan, K., Levy, M., Kim, R., Brown, T., Reding, M., **Rith-Najarian, L**., ... Chorpita, B. (2017). Managing in-session surprises: Identifying provider responses to emergent life events during evidence-based treatment implementation. *Administration and Policy in Mental Health, 44*, 164-176.
- **Rith-Najarian, L.,** Chorpita, B.F., & Daleiden, E. L. (2016). Evidence-Based Decision-Making in Youth Mental Health Prevention. *American Journal of Preventative Medicine, 51,* S132–S139.
- Higa-McMillan, C. K., Francis, S. E., **Rith-Najarian, L.,** & Chorpita, B. F. (2016). Evidence base update: 50 years of research on treatment for child and adolescent anxiety. *Journal of Clinical Child and Adolescent Psychology*, *45*, 91-113.

- Becker, K. D., Buckingham, S. L., **Rith-Najarian, L.,** & Kline, E. R. (2016). The Common Elements of treatment engagement for clinically high-risk youth and youth with first-episode psychosis. *Early Intervention in Psychiatry, 10,* 455–467.
- McLaughlin, K.A., **Rith-Najarian, L.**, Dirks, M.A., & Sheridan, M.A. (2015). Low vagal tone magnifies the association between psychosocial stress exposure and internalizing psychopathology in adolescents. *Journal of Clinical Child and Adolescent Psychology*, *44*, 314-328.
- **Rith-Najarian, L.,** McLaughlin, K.A., Sheridan, M.A., & Nock, M.K. (2014). The biopsychosocial model of stress in adolescence: Self-awareness of performance versus stress reactivity. *Stress, 17,* 193-203.

SELECTED PRESENTATIONS

- Letamendi, A., **Rith-Najarian, L.**, & Shek, Y.L. (June, 2019). *Mental Health on The Hill: An Integrated Campus Approach to Addressing Student Well-Being.* At UC Mental Health Best Practices Conference, Los Angeles, CA.
- **Rith-Najarian, L**., Gong-Guy, E., Chorpita, B., & Chavira, D. (November, 2018). *Large-Scale Prevention of Anxiety and Depression in University Students Through Online Programming*. Symposium presentation at Annual Convention of the Association for Behavioral and Cognitive Therapies. Washington, D.C. .
- Shover, C., Maust, D., **Rith-Najarian, L.,** & Darling, Z. (2018, October). *Opioid Overdose Prevention in California: A Map to Inform Public Health Policy.* Presentation at The Opioid Hack-A-Thon West. Irvine, CA.
- **Rith-Najarian, L.** (2016, September). *The Happiness Challenge: Lessons Learned and Future Directions in University Student Online Mental Health Promotion.* Presentation at Mobility and Modern Web Annual Conference. Los Angeles, CA.
- **Rith-Najarian, L.,** Park, A., Mesri, B., & Chorpita, B. (2015, November). *Meta-Analysis of Youth Cognitive Behavioral Therapy Follow-Up Effects by Treatment Practice Elements.* Poster presented at 2015 Annual Convention of the Association for Behavioral and Cognitive Therapies. Chicago, IL.

SELECTED FELLOWSHIPS, GRANTS, AWARDS, & HONORS

2017 - 18 UCLA Dissertation Year Fellowship ($20,000 + tuition)
2014 - 18 UCLA Mangasar M Mangasarian Fellowship ($10,000)
2017 American Psychological Association Dissertation Research Award ($1000)
2017 First Place, UC-Wide Grad Slam - Research Presentation Competition ($6,000)
2017 First Place, UCLA Grad Slam - Research Presentation Competition ($3,000)
2017 Advanced Graduate Student Clinical Scientist Award, UCLA Department of Psychology
2014 - 17 J.P. Morgan's D.Z. Liebmann Fellowship ($18,000)
2016 UCLA Life Science's Innovation Fund Grant ($10,000)
2014 - 15 Graduate Summer Research Mentorship ($6,000)
2015 American College Health Foundation Weiss Student Mental Health Award ($2,500)
2015 Dissemination & Implementation Science Student Presentation Award, ABCT ($75)
2014 Graduate Summer Research Mentorship ($6,000)
2012 Harvard University Phi Beta Kappa
2012 Harvard University Psychology Honors Thesis Faculty Prize ($100)
2011 Harvard College Research Program Fellowship ($2000)

Dissertation Introduction

If you build it, will they come? And will they stay? And will it help? And for whom? This dissertation seeks to answer such questions in the context of universal prevention of anxiety and depression in university students through online skills-based intervention programming.

Importance of Prevention of Anxiety, Depression, and Stress in Young Adults

The transition from adolescence to adulthood is a developmental period characterized by many changes in identity formation, sense of agency, cognitive-affective strengths, psychological wellbeing, and social wellbeing (Arnett, 2000; Conley, Kirsch, Dickson, & Bryant, 2014; Schwartz, Cote, & Arnett, 2005). Given the importance of this developmental period, it is essential to ensure that young adults are holistically well, not just academically, socially, and professionally, but also psychologically. Problematically, emerging adulthood is also a period marked with increased risk for onset of psychopathology (Schulenberg, Sameroff, & Cicchetti, 2004). For young adults entering a college environment, there are immediately declines in psychological well-being and increases in psychological distress, with such deterioration seeming to plateau later in college rather than resolve (Conley et al., 2014). Research on 12-month prevalence rates of psychological disorders in college students internationally found that nearly one in three students meet criteria for at least one disorder (Auerbach et al., 2018). High prevalence rates of psychological disorders have also been found in mental health research reviews of graduate and professional students (e.g., Evans, Bira, Gastelum, Weiss, & Vanderford, 2018; Rotenstein et al., 2016). Accordingly, it is crucial that we find ways to effectively intervene with university students in order to reduce psychopathology for as many students as possible.

Depression and anxiety are especially prevalent among university students and cause significant impairment. Currently, major depression disorder and generalized anxiety disorder have the highest 12-month prevalence among college students internationally (18.5% and 16.7%, respectively), and are comparably much more prevalent than any other diagnosis (Auerbach et al., 2018). Academically, depression has been found to predict lower GPA and drop out, especially for students with comorbid anxiety (Eisenberg, Golberstein, & Hunt, 2009). Physically, depression symptoms have been shown to precede subsequent poor physical health and associated work impairment in young adulthood (Keenan-Miller, Hammen, & Brennan, 2007). In terms of resulting significant distress, students screening positive for depression or anxiety have shown more frequent suicidal behaviors (Keyes et al., 2012). Anxiety and mood disorders are even consequences of each other, as anxiety symptoms have been found to predict later onset of depressive symptom, and vice versa (Jacobson & Newman, 2017). Moreover, anxiety and depression symptoms are highly comorbid, and so there is increasing support for targeting these internalizing symptoms together (e.g., Barlow, Allen, & Choate, 2004; Craske, 2012). Considering the high prevalence rates of depression and anxiety individually and their high comorbidity, prevention programs that target both types of symptoms have potential to help the highest number of young people with such problems.

In addition, stress is strongly associated with anxiety and depression symptoms (Mahmoud, Staten, Hall, & Lennie, 2012; Regehr, Glancy, & Pitts, 2013), and so reduction in maladaptive stress responses is also an important component for prevention of internalizing symptoms in university students. Many students report mental health concerns associated with unique stressors related to the university environment (Eisenberg, Gollust, Golberstein, & Hefner, 2007; Hyun, Quinn, Madon, & Lustig, 2006). For example, one study found that

students' most commonly cited sources of psychological stress were academic performance, pressure to succeed, and post-graduation plans (Beiter et al., 2015). Higher stress symptoms have been associated with higher anxiety and depression symptoms in undergraduate and graduate students (e.g., Norton, 2007; Saravanan & Wilks, 2014). Therefore, intervention programs targeting anxiety and depression should also focus on promotion of general stress coping and accordingly measure change in subjective stress.

Given the nature of anxiety and depression disorders, prevention programming for university students has some unique benefits relative to treatment services. College students with earlier onset of anxiety disorder or mood disorder are much more likely to experience greater severity and chronicity of illness (Pedrelli, Nyer, Yeung, Zulauf, & Wilens, 2015), so intervening early is important. For depression, risk of future major depressive episode increases progressively with each successive recurrence of symptoms (Solomon et al., 2000). Therefore, if initial episodes can be prevented altogether for university students, then there can be lower risk of depression symptoms long-term. Indeed, a research review estimated that evidence-based prevention can prevent 22% of new depression cases each year (Cuijpers, Straten, Smit, Mihalopoulos, & Beekman, 2008). Regarding anxiety, individuals with an anxiety disorder are more likely than not to remain symptomatic during young adulthood, often experiencing homotypic and heterotypic continuity such that they maintain symptoms of the original anxiety disorder as well as develop symptoms of other anxiety disorders (Ranøyen et al., 2018). Therefore, prevention of a single anxiety disorder theoretically lowers one's risk of later developing multiple anxiety disorders. Fortunately, a meta-analysis of anxiety prevention programs estimated a 45% reduction in anxiety incident rate due to intervention (Moreno-Peral et al., 2017). Taken together, prevention of anxiety and depression disorders has the potential to

reduce incidence of internalizing psychopathology in the short-term and long-term, before treatment services become necessary.

Benefits and Difficulties of Implementing Prevention Programming for Students

There are benefits of offering prevention programs to university students specific to their age as well as their environment. It has been suggested that implementing prevention programs during developmental stages when individuals are still in educational systems can promote durable change because young people's cognitive and behavioral patterns are not yet rigidly engrained (Werner-Seidler, Perry, Calear, Newby, & Christensen, 2017). Correspondingly, we might consider offering prevention for non-student young adults as well as university students, given that the elevated psychological disorder prevalence at this age is regardless of post-secondary enrollment (Blanco et al., 2008). However, implementation of prevention programming for students has some unique advantageous. First, campus environments lend themselves to population-level intervention, such as universal prevention, because educational institutions are often already in the practice of offering formal health services (e.g., student health clinics) as well as various community-level campaigns (Reavley & Jorm, 2010). Moreover, prevention efforts can more holistically intervene with students, given that university campuses are at the intersection of many aspects of their life, including academic activities, health services, housing, social activities, and more (Reavley & Jorm, 2010). Finally, there is ease in accessing a student population when offering intervention programming given that the population is geographically centralized and clearly defined by student enrollment status. Accordingly, there is currently much attention given to the development of prevention programming for university students.

Over recent decades, researchers and practitioners have made substantial progress in addressing anxiety and depression in young people, producing hundreds of prevention programs for university students (e.g., Conley, Durlak, & Kirsch, 2015), and expanding formal mental health services to meet increasing demand (Eisenberg, Hunt, Speer, & Zivin, 2011). However, there are many barriers – stigma concerns, accessibility issues, etc. – that still prevent young adults from seeking mental health services (Eisenberg et al., 2011; Gulliver, Griffiths, & Christensen, 2010). Thus, despite overall increased service use by college students, for years research has unfortunately found that most students with an anxiety or mood disorder do not use mental health services (Blanco et al., 2008; Hunt & Eisenberg, 2010). Internationally, as few as 16% of students meeting criteria for a psychological disorder were found to use mental health services (Auerbach et al., 2016). The resulting mental health service need-use gap is an important issue to address.

Intervention via technology has been recommended as one solution to address these low rates of service seeking (Farrer et al., 2013; Pedrelli et al., 2015). Accordingly, dozens of online mental health programs have been developed for university student populations and have demonstrated effectiveness in reducing anxiety, depression, and stress (Conley et al., 2015; Davies, Morriss, & Glazebrook, 2014; Farrer et al., 2013). However, such programs are often beset with low intervention completion and adherence, whether it be no longer accessing the online program midway through delivery or only spending a few minutes passively viewing the intervention content without practicing the learned material (Karyotaki et al., 2015; Kelders, Kok, Ossebaard, & Van Gemert-Pijnen, 2012). Though there is a lack of direct evidence explaining why engagement is low for electronic mental health programs, theories offer that programs are not user-friendly or have their own barriers, such a privacy concerns (Torous,

Nicholas, Larsen, Firth, & Christensen, 2018). The low adherence and engagement rates of online interventions have been considered centrally related to their implementation in real world settings outside of research studies (Fleming et al., 2018; Torous et al., 2018). Thus, despite vast research support for online interventions' effectiveness, their limitations decrease chances of successful implementation and dissemination on college campuses. Therefore, there remains a great need to examine online prevention programs that address existing implementation limitations.

Designing for Dissemination and Implementation

Given the increasing prioritization of reducing the science-practice gap for evidence-based interventions, dissemination and implementation (D&I) research is an expanding field. According to the National Institutes of Health program announcement on D&I research funding (National Institutes of Health, 2013):

- Dissemination research is defined as "studies identifying mechanisms and approaches to package and convey the evidence-based information necessary to improve public health and clinical care services."
- Implementation research is defined as "the scientific study of methods to promote the integration of research findings and evidence-based interventions into healthcare policy and practice."

As such, traditional D&I research does not involve intervention development per se, but rather the testing of interventions after they have been through efficacy and/or effectiveness research stages. More recently, D&I research leaders in public health have been advocating for "Designing for Dissemination (D4D)" (e.g., Brownson, Colditz, & Proctor, 2017). The process of D4D involves the development of interventions (among other products of research) such that

they match the unique needs and characteristics of their target audience in order to increase their chances of successful real-world implementation and dissemination (Rabin & Brownson, 2017). Some examples of D4D actions that can be taken during intervention development research are: identify gaps in existing research; identify barriers to implementation; involve stakeholders as early as possible; and collect data relevant to implementation feasibility during development (e.g., adherence data, maintenance of effects post-intervention; Owen et al., 2017). Thus, rather than conducting D&I research after an intervention is established as "evidence-based," a D4D approach incorporates D&I principles during the intervention designing and effectiveness testing.

The D4D approach can be applied to research on online prevention programming for depression, anxiety, and stress in university students. Research studies should iteratively inform and test the development and implementation of such an online intervention. For example, given that many relevant online interventions already exist, systematic review could both identify common evidence-based elements while identifying intervention research limitations at the meta-analytic level. Thereafter, once an initial intervention platform has been correspondingly developed, research studies could investigate student response (e.g., enrollment, adherence, feedback) to intervention implementation, and then a revised version of the intervention could be tested for effectiveness. Emerging evidence suggests there are benefits to tailoring implementation strategies to address context-specific barriers (Powell et al., 2017). Therefore, student help-seeking barriers and campus implementation barriers will be important considerations. Taking such an approach may facilitate the development of prevention programming that is appealing, feasible, and effective for those young adults on campuses who could benefit from intervention, and do so in a time-efficient way.

Dissertation Aims

Through four studies, this dissertation research aims to inform the development of an online universal prevention program for anxiety and depression in university students and examine its implementation and effectiveness. This dissertation seeks to develop an intervention that: (a) includes evidence-based practices, (b) reaches students in need, (c) is feasible for implementation, and (d) produces better symptom outcomes than receiving no intervention. The study in **Chapter 1,** entitled *A Systematic Review of Prevention Programs Targeting Depression, Anxiety, and Stress in University Students*, reviews 68 relevant prevention programs with the aims of: (1) identifying strengths and limitations, (2) examining effect sizes, and (3) distilling common practice elements. The studies in **Chapter 2**, entitled *What's in a Name?: Branding of Online Mental Health Programming for University Students*, examine intervention enrollment data at two different campuses to determine if a lowered-barrier recruitment approach: (1) attracts underserved students in need, (2) yields demographically representative samples, and (3) differentially attracts students by alternative brand name. The study in **Chapter 3**, *Open Trial Feasibility Study of an Online Universal Prevention Program for Anxiety and Depression in University Students*, uses an open trial design to examine the feasibility of the developed online intervention program in terms of: (1) recruitment strategy yields, (2) retention and adherence rates, (3) program acceptability, (4) assessment procedure sensitivity, and (5) preliminary safety check. Finally, the study in **Chapter 4**, entitled *Randomized Controlled Trial of an Online Program for the Universal Prevention of Anxiety and Depression in University Students*, uses an RCT design to examine the effectiveness of the online prevention program when compared to a waitlist group to determine if the online intervention produces pre-post effects that are: (1) significantly better compared with the waitlist, (2) replicated by the waitlist once receiving

delayed intervention access, (3) maintained through follow-up, and (4) moderated by baseline characteristics of individuals. Evaluating the findings of these four studies together will elucidate the ways in which online anxiety and depression prevention programming might – and perhaps might not – help reduce the service need-use gap for university students. Such evaluation is considered in the **Dissertation Discussion** chapter, as is evaluation of the strengths and remaining directions of this dissertation's research approach.

Chapter 1: A Systematic Review of Prevention Programs Targeting Depression, Anxiety, and Stress in University Students

Abstract

Background: Given the prevalence of anxiety, depression, and stress among university students, it is important to assess the effectiveness of prevention programs for these problems. Beyond examining effect sizes, applying a common elements approach can enhance our understanding of which practice elements are most frequently included in symptom-reducing programs. **Method:** This review examined effective (i.e., outcome-producing) prevention programs targeting depression, anxiety, and/or stress in university students. Programs could be delivered in a group-based, online/computer-delivered, or self-administered format at the universal, selective, or indicated prevention level. **Results:** The resulting sample of 62 articles covered 68 prevention programs for college, graduate, or professional students across 15 countries. Average effect sizes for programs were moderate (overall $g = 0.65$), regardless of delivery format or prevention level. The most common practice elements (overall and for programs producing large effects) were: psychoeducation (72%), relaxation (69%), and cognitive monitoring/restructuring (47%). Many programs were limited by: (a) symptom target-outcome mismatches, (b) disproportionately female samples, and (c) inconsistently reported adherence data. **Conclusion:** The outcome-producing prevention programs in our sample had common practice elements and produced moderate reduction in symptoms overall. Future research of depression, anxiety, and stress prevention programs for university students can investigate practice elements' unique and combined impact on outcomes, further explore under-tested practice elements, and use findings to inform intervention design.

Introduction

University students are at high risk for depression and anxiety symptoms (American College Health Association, 2018; Zivin, Eisenberg, Gollust, & Golberstein, 2009) and are exposed to multiple stressors unique to this developmental period (Beiter et al., 2015; Drake, Sladek, & Doane, 2016), highlighting the need for interventions that target these mental health issues. Although psychological disorder prevalence in young adults is elevated regardless of post-secondary enrollment (Blanco et al., 2008), campus-based prevention programming has some natural advantages for large-scale implementation, given geographical proximity of students to each other, and available campus resources. Prevention programs for university students have the potential to reduce experiences of depression, anxiety, and stress, along with associated consequences, such as university drop-out, decreased academic performance, social functioning, and suicidal behavior (American College Health Association, 2018; Keyes et al., 2012; NAMI, 2012; Salzer, 2012).

Reviews of Prevention Programming for University Students

There have been multiple systematic reviews of prevention programs for depression, anxiety, and/or stress in university students, across which findings have been inconsistent. For example, Conley, Durlak, and Kirsch (2015) reviewed universal mental health prevention programs for university students that targeted depression, anxiety, and stress, and found small effects for skill-training programs, as long as they included supervised skills practice. Another review focused on technology-based prevention programs for depression, anxiety, and stress in university students, and found moderate effects for these programs compared with inactive comparision groups (e.g., waitlist), but not when compared with another active group (e.g., attentional control; Davies, Morriss, & Glazebrook, 2014). Similarly, a review of anxiety and

depression prevention programs for adolescents and young adults found symptom reduction in only 60% of programs (Christensen, Pallister, Smale, Hickie, & Calear, 2010). Moreover, a review on depression prevention programs for college students concluded that evidence of effective outcomes in the literature is currently insufficient, due to a lack of standardized interventions that have been implemented and evaluated across multiple student contexts (Buchanan, 2012). Given the mixed findings in reviews examining varying effectiveness by prevention program differences, it may be important to step back and further explore commonalities across effective programs. To this point, although these reviews have synthesized information about demographics, outcome effect sizes, and comparisons by broad treatment type (e.g., cognitive behavioral therapy versus relaxation training; skills-training versus psychoeducation), none have examined specific intervention practice elements common across prevention programs.

Why Identify Common Practice Elements of Effective Intervention Programs

Research on practice elements – also sometimes called treatment components, common elements, treatment elements, or common components – has been growing in the field of clinical psychology (e.g., Chorpita, Daleiden, & Weisz, 2005; Murray et al., 2014; Weisz, Ugueto, Herren, Afienko, & Rutt, 2011). Practice elements are the discrete clinical strategies and techniques that are common across a specified set of interventions. Practice elements can be instructional strategies (e.g., psychoeducation about symptoms a client is experiencing) or skills-based strategies that promote improvement in symptoms (e.g., introducing cognitive restructuring principles to a client and having them practice the skill). Identifying and understanding practice elements can help us understand what is common across treatments, rather than only emphasize the differences between treatments (Arch & Craske, 2008; Rotheram-

Borus, Swendeman, & Chorpita, 2012). Furthermore, it has been argued that identifying common practice elements, in conjunction with theory and logic frameworks for organizing elements, can inform protocol design or adaptation within a given context or population to increase compatibility of treatments (see Chorpita & Daleiden, 2014 for examples of successful component-based approaches to design) and prevention programs (Rith-Najarian, Daleiden, & Chorpita, 2016). Furthermore, in the absence of an existing evidence-based treatment for a specific population, relying on known practice elements can provide at least an evidence-informed treatment approach, rather than forcing selection of an existing treatment that is inappropriate to the context or making clinical decisions completely from scratch (Southam-Gerow et al., 2014). Finally, research that identifies common practice elements can also highlight which practice elements are relatively under-researched and could benefit from further testing. The field of mental health prevention could reap benefits by applying the scientific approaches used by practice element research to better synthesize what we know about the common elements of effective prevention programming.

How to Identify Practice Elements of Interventions

One strategy that allows for the systematic identification and aggregation of intervention elements is called the distillation and matching model (DMM; Chorpita et al., 2005). When applying the DMM, studies of intervention programs (for some pre-defined population, targets, contexts, etc.) are first reviewed in order to code their treatment protocols, which can be considered composites of practice elements. Based on protocol content descriptions, information about underlying pragmatic clinical approaches (e.g., insight-building, exposure) can be extracted and reduced into a set of atheoretical practice element codes – the "distillation." The "matching" part involves analyzing whether unique profiles or clusters of elements emerge

within distinct contexts in the literature (e.g., treatments targeting anxiety versus those treating depression; treatments delivered in a group versus one-on-one; treatments tested with children versus adolescents). Even without conducting the full analytic procedures for the "matching" component, "distillation" allows for identification of common elements from treatment protocols to determine which individual practice elements are most frequently included across effective programs of certain characteristics (e.g., universal-level prevention for anxiety).

The distillation methodology has already led to the identification of common elements for: universal adolescent prevention programs (Boustani et al., 2015); home-based prevention of child maltreatment (Kaye, Faber, Davenport, & Perkins, 2018), psychosocial child treatments for disruptive behavior (Garland, Hawley, Brookman-Frazee, & Hurlburt, 2008); engagement practices in children's mental health services (Becker, Boustani, Gellatly, & Chorpita, 2018; Lindsey et al., 2014); and engagement practices for first-episode psychosis (Becker, Buckingham, Rith-Najarian, & Kline, 2016). Distillation of common elements of prevention programs targeting university student depression, anxiety, and stress will enhance our understanding of such programs, tendering the benefits for the clinical decision-making (e.g., inform stakeholders and clinicians in contexts or with certain populations for which there is no existing evidence-based intervention) and research planning (e.g., identify under-studied practice elements needing further research). Broadly, a distillation approach will allow synthesis of practice element commonalities across effective prevention programming within the university student prevention literature for the first time.

The Current Study

This review aimed to systemically examine programs targeting depression, anxiety, and/or stress in university students at the universal, selective, and indicated prevention level,

including programs that are group-based, online/computer-delivered, and self-administered. A preliminary aim for this review was to provide an up-to-date summary of characteristics of such prevention program studies (e.g., delivery format types, targets, sample demographics, adherence/competition data) and their aggregated effect sizes. Then, the primary aim was to identify the practice elements that are common to these prevention programs, using the distillation approach. In addition to examining overall frequency of practice elements across programs, we compared element frequencies among: 1) universal, selective, and indicated prevention programs; 2) group-based, online/computer-delivered, and self-administered programs; and 3) programs producing small, moderate, and large effect sizes. This research aim was exploratory given its novelty within the university student prevention literature, and therefore there were no a priori hypotheses about which elements would be most frequent nor about any differences in distribution of frequency across categories. Overall, our review strategy expands on those of previous similar reviews by examining not just prevention program types, research study/sample characteristics, and outcome effect sizes, but also practice elements at the aggregate level.

Methods

Article Selection and Sample

Article identification. We first identified reviews conducted on related topics and compiled articles from their reference lists: Bamber and Kraenzle Schneider (2016), $n = 57$; Buchanan (2012), $n = 16$; Christensen and colleagues (2010), $n = 36$; Conley, Durlak, and Kirsch (2015), $n = 91$; Davies, Morriss, and Glazebrook (2014), $n = 36$ (list of included and excluded articles); Farrer and colleagues (2013), $n = 27$; Larun and colleagues (2006), $n = 16$; Regehr, Glancy, and Pitts (2013), $n = 32$. Next, we identified additional articles on relevant

interventions through an online search of the Cochrane Library. The search, last updated in April 2018, used the following simultaneous constraints (including word variations) to identify 242 records that: (a) were in the *trials* category, (b) had a title containing the word *prevention*, *intervention*, *program*, *training*, *semester*, or *stress reduction*; (c) had a title containing the word *depression*, *anxiety*, *stress*, or *transdiagnostic*; (d) had the word *student* appearing in the title, abstract, or keywords; and (e) had a title not containing the words *child*, *adolescent*, or *school*. Duplicate articles (either the same article, $n = 118$, or articles on the same study or the same prevention group, $n = 4$, e.g., Braithwaite & Fincham, 2007, 2009) were removed, leaving 243 articles from other reviews and 188 articles identified in the Cochrane Library to be prescreened. The flowchart of article inclusion/exclusion is presented in Figure 1.1.

Inclusion criteria. Included studies were required to have: (a) investigated a primary intervention; (b) included post-secondary students (age-equivalent non-student samples were not eligible); (c) been designed to primarily target anxiety (i.e., general anxiety, worry, performance/social, perfectionism, panic), depression (i.e., major depression, mood, affect), or stress (i.e., psychological stress, physiological stress, or distress); (d) used random assignment to groups; (e) used a total study sample size of 30 students or more, and (f) been published in a peer-reviewed journal (e.g., dissertations that were not later accepted for publication were not eligible, articles presented at conferences were not eligible). For articles identified, inclusion criteria were determined based on the extracted article information provided in review articles or by reading the full article text of the original article, when information from the reviews was insufficient. For articles identified in the Cochrane Library, inclusion criteria were determined by reading the abstracts, or again by reading the full article text, when the abstract was insufficiently informative. Based on these inclusion criteria, 140 articles were retained for review.

Exclusion criteria. We then reviewed the remaining 140 full text articles to apply our exclusion criteria. First, we excluded interventions that were delivered in one-on-one sessions, as we deemed them characteristically different and less scalable for prevention-level programming at universities. Second, we excluded interventions that were tested on clinical samples in which the majority or all students met criteria for an anxiety or mood disorder in DSM-5 or surpassed a clinical cut-off score (subclinical cut-off scores were permissible). We deemed such interventions as more representative of a tertiary intervention, or "treatment level." We did not, however, exclude studies that included some diagnostic-level participants, as long as the majority (i.e., > 50%) were subclinical or low symptomatic.

For the final exclusion criterion, we reviewed the results sections of the remaining 112 articles to confirm significant results for a "winning" intervention group on an eligible outcome measure of interest (n = 50 excluded). To qualify as a "winning" intervention group, we required at least one positive and significant (i.e., $p < .05$) between-group result for an established outcome measure of: depression, anxiety, stress, or positive/negative affect. This "winning" criterion has been used by other common element distillation reviews (Boustani et al., 2015; Chorpita et al., 2005) and was applied because we wanted our review results to represent findings for "successful" prevention programs, and comparisons between effective and ineffective programs was deemed beyond the scope of this paper. The first author coded results for all articles, and a subset of 58 articles were double coded by the second author to confirm reliability of exclusion decision ($\kappa = 0.78$). There must have been either a between-group result observed for the intervention group at post-assessment (not mid-assessment or follow-up), or a significant group by time interaction from pre- to post-assessment. If a study reported a significant between-group differences on any variables at baseline, then the respective covariate

must have been included in the analysis. In contrast, if results were only significant when including an unjustified covariate in analyses, then significant results were not counted for that respective outcome. Lastly, studies that demonstrated significant between-group results only on a study-created measure or only on a clinically assessed diagnostic status (i.e., not using a standardized structured interview) were excluded, because we considered such measures to be unstandardized (2 articles excluded). A final set of 62 articles were retained for coding, all of which were written in English.

Coding Procedure

Coding system creation and training. A coding manual was developed, informed by the PracticeWise Clinical Coding System (PracticeWise, 2012), which has been the basis for multiple other published distillation reviews (Becker et al., 2016; Boustani et al., 2015; Chorpita & Daleiden, 2009; Higa-McMillan, Francis, Rith-Najarian, & Chorpita, 2016; Lindsey et al., 2014). The PracticeWise codebook has been adapted for prevention program review before (Boustani et al., 2015), and in a similar manner a specific manual was developed for this study (see details below) to better suit coding of prevention programs for university students. In addition to coding information reported in the study article, data were supplemented with publicly available supplemental materials or cited information. Five article coders (two college graduates, three advanced undergraduates, all psychology majors) initially met to be trained on the coding manual. Each coder was assigned five practice articles (not from the final review sample), after which the principal investigator provided individualized feedback on their reliability. Every article was coded independently by a pair of coders (27 articles by one pair, 21 articles by a second pair, and 14 articles by a third pair). Coding validation meetings were held weekly with each pair to resolve discrepancies, with the first author serving as a third rater for

unresolved discrepancies.

Inter-rater reliability was calculated with Cohen's kappa statistic. Reliability statistics were calculated based on coder's data collection prior to resolving discrepancies. We used previously established guidelines to characterize magnitude of inter-rater reliability as: excellent for $\kappa \geq 0.75$, good for $\kappa \geq 0.60$, moderate for $\kappa \geq 0.40$, fair for $\kappa \geq 0.20$, and poor for $\kappa < 0.20$ (Banerjee, 1999; Landis & Koch, 1977). Mean inter-rater reliability was 0.69 with a range from 0.41 to 1.00. All inter-rater reliability results are presented along with the full code definitions in Table 1.1.

Prevention level type. Gordon's (1983) prevention classification system was used to define three prevention levels: universal, selective, and indicated. Categorization was coded based on participant recruitment and inclusion criteria information reported in the methods and often confirmed by information in article introduction (e.g., narrative about importance of prevention with a given target population; description of study aims). For example, typically the only inclusion criteria for a "universal" level program would be regarding age to consent and current enrollment as a student within the recruitment context (e.g., school of nursing).

Delivery format type. Programs were categorized into three delivery format types: group-based, online/computer-deliver, or self-administered. Categorization was coded based on the intervention description reported in the methods of each study article.

Intervention target. Target refers to the symptom(s) or problem area(s) that the prevention program was designed to address: anxiety, depression, and stress. For the purpose of this review, prevention program target was coded if it was common to two or more of: (a) the selection criteria for the study sample (unless universal); (b) the study-reported target for the intervention, and (c) the measure outcomes reported in the results. Although many studies might

include measures of anxiety, depression, and/or stress, this fact alone does not mean that those outcomes were intended targets. For example, if a study included students with subclinical anxiety, provided anxiety and depression prevention programming, and examined anxiety, depression, and stress outcomes, then the targets would be both depression and anxiety. Programs could have multiple targets. For this reason, we could not categorize each program into one target category, but rather coded for the presence of each target. Although we did not include programs that primarily targeted other problem areas solely (e.g., externalizing symptoms, substance use, eating disorders), such outcomes could have been additional primary targets or secondary targets.

Study sample characteristics. A write-in field was used to collect the geographic and demographic information (e.g., gender, student type) about each study's sample, which was directly extracted from sample descriptions and descriptive data reported in the study article results.

Intervention adherence/completion data. A write-in field was used to collect any adherence/completion data about the study's intervention. Data could be reported in terms of adherence, engagement, program use rate, compliance, program drop-out, etc.

Practice elements. The PracticeWise Clinical Coding System is design to evaluate studies and protocols relevant to treatment of children and adolescent mental health concerns, and therefore, not all codes were relevant for our target population. From the 73 original PracticeWise element codes, we excluded 15 parent-related codes (e.g., "parental monitoring"), and 39 codes with low prevalence in the current program sample (i.e., present in 5% or less of program sample; e.g., "stimulus/antecedent control," "behavioral contracting, "response prevention"). Based on our prescreening of articles, we added two new practice element codes:

"peer engagement" and "stress management/coping skills." We also created new specific types of "problem solving" to represent problem solving practices specific to "sleep hygiene" and "time management", and of specific types of "psychoeducation" to differentiate between "psychoeducation about symptoms" and "psychoeducation about services/resources," a distinction which has been made in other distillation manuals (Becker et al., 2018). The final coding system included 24 elements, as described in Table 1.1. The inter-rater reliability for practice elements was excellent for 5 elements, good for 15 elements, and moderate for 4 elements (see Table 1.1). The principal investigator coded all articles during a final expert review round to confirm presence or absence of elements with moderate inter-rater reliability.

Effect Sizes

Between-group effect sizes were calculated based on the means, standard deviations, and sample size of the intervention group compared with the comparison group at post-treatment. Effect sizes were based on outcome data from the measures with a qualifying significant between-group result (defined above). Cohen's *d* with Hedges' correction for small samples (Hedges, 1981) was calculated using the formula below.

$$\textit{Hedges' } g = \frac{M_1 - M_2}{\sqrt{\frac{(n_1 - 1){SD_1}^2 + (n_2 - 1){SD_2}^2}{n_1 + n_2 - 2}}} \times \left(1 - \frac{3}{4(n_2 - 1) - 1}\right)$$

One effect size was assigned to each intervention group. For interventions that had multiple effect sizes calculated, the largest one was selected, excluding effect sizes based on outcome data with significant differences at baseline favoring the intervention group, as such data could artificially inflate the effect size. Each intervention group was then categorized into three groups of effect size magnitude: small (g = 0.20 – 0.49), moderate (g = 0.50 – 0.79), and

large (g = 0.80+). We used Comprehensive Meta-Analysis software version 3 (Borenstein, Hedges, Higgins, & Rothstein, 2013) to calculate the average effect sizes with a random effects model.

We assessed impact of publication bias on our effect size results by: (a) examining the funnel plot (Torgerson, 2006) of standard error (y-axis) and respective effect sizes (x-axis); running Egger's weighted regression test (Egger, Smith, Schneider, & Minder, 1997), and (c) computing a fail-safe N (Rosenthal, 1979).

Data Analysis Plan

The majority of analyses were descriptive, reporting results by frequency counts or percentages. Chi-squared analyses were used to examine (a) if certain delivery format types were more likely at different prevention levels and (b) if each practice element was significantly more/less frequent in certain programs. To examine differences in mean number of practice elements, one-way ANOVAs were used with Bonferroni post hoc tests.

Results

Program Sample

The final article sample contained 62 articles, with 68 unique winning prevention program groups to code, because 6 articles tested multiple prevention programs that produced significant effects. Articles were published between 1978 and 2018. References for review sample are listed in Appendix 1A. Program characteristics from articles included in our review are outlined in Appendix 1B.

Program Prevention Levels and Delivery Format Types

Programs were at the prevention levels of: universal (n = 25), selective (n = 20), and indicated (n = 23). Program delivery formats were: group (n = 37), online/ computer-delivered

(n = 16), and self-administered (n = 15). Of the universal prevention programs, 72% were group-based, 16% were self-administered, and 12% were online/computer-delivered. Universal prevention programs were significantly more likely to be group-based (χ^2 = 4.93, df = 1, p = .03) than non-universal-level programs. Of the selective prevention programs, 55% were group-based, 35% were self-administered, and 10% were online/computer-delivered. Although the proportion of self-administered groups for selective prevention programs was higher than for non-selective-level programs, the disproportion was not statistically significant (χ^2 = 2.76, df = 1, p = .10). Of the indicated prevention programs, 48% were online/computer-delivered, 35% were group-based, and 17% were self-delivered. Indicated prevention programs were significantly more likely to be online/computer-delivered (χ^2 = 11.40, df = 1, p = .001) than non-indicated-level programs.

Intervention Targets

The programs were designed to target: stress only (n = 22); depression, anxiety, and stress (n = 12); depression and anxiety (n = 11); anxiety only (n = 10); anxiety and stress (n = 6); depression only (n = 5); and depression and stress (n = 2). However, the target type(s) of an intervention was not necessarily the outcome type(s) with significant results (i.e., an intervention described as targeting stress produced a significant result on a depression measure). In fact, 23.5% of groups produced significant results on symptom measures that were not their identified target(s) – 17.6% had significant results on non-target measures in conjunction with target-congruent measures and 5.9% had significant results only on non-target measures.

Study Sample Characteristics

A range of student types were covered by the interventions in this review, including undergraduate students (e.g., first year students, four-year college students, student varsity

athletes, psychology majors, community college students), graduate students, and professional students (e.g., nursing students, medical students, law students, dentistry students). Studies took place across 15 countries. Excluding studies with single gender samples (n = 11), the majority of samples were disproportionately female, with 51% having a sample that was two-thirds or more female and 31% having a sample that was at least three-quarters or more female. See Appendix 1B for details.

Adherence and Completion Data

Intervention adherence and/or completion data were reported for 38 (56%) of the interventions: 18 (49%) group-delivered, 11 (69%) online/computer-based, and 10 (67%) self-administered. Studies that only reported research dropout rates were not counted. Measures of adherence/completion varied widely (e.g., minutes online, homework competition, intervention initiation, skill practice frequency, reading of material, group attendance). Some only provided the percentage of participants who "received" or "completed" the intervention without further definition. Context was an important factor to note, as some of the highest full completion rates reported were for single session attendance (Cukrowicz & Joiner, 2007; Yusoff & Esa, 2015) or for studies with greater incentives (e.g., $300+ compensation for complete intervention participation, Rose et al., 2013; completion of all sessions required to receive course credit, Anshel, 1996). Because of the variability in adherence measures used, it was not possible to calculate average rates or present ranges. See Appendix 1B for details.

Effect Sizes

The overall average between-group effect size (n = 64) was g = 0.65 [0.57-0.73]. There were three articles (covering four groups) that had insufficient data to calculate an effect size. By delivery format, the effect sizes were: 0.69 [0.58 - 0.81] for group (n = 36); 0.65 [0.50-0.81] for

self-administered (n = 14); and 0.52 [0.41-0.63] for online/computer-delivered (n = 14). By prevention level, effect sizes were: 0.69 [0.55-0.83] for universal (n = 24); 0.73 [0.59-0.87] for selective (n = 17); and 0.53 [0.44-0.63] for indicated (n = 23). Of the 64 groups, large effect sizes were produced by 29% of groups; by 30% of selective, 28% of universal prevention, and 26% of indicated programs; and by 35% of group-delivered, 27% of self-administered, and 19% of online/computer-based programs. See Figure 1.2 for a forest plot of all effect sizes.

Our analyses of impact of publication bias on effect sizes suggested that our data was biased, as the funnel plot was asymmetric and Egger's weighted regression test was significant (t = 2.06, p = .04). However, the Fail-safe N indicated that 8,508 groups with an effect size of 0 would need to be added to this review sample in order for the average intervention effect size to become non-significantly different than zero.

Practice Element Frequencies

Figures 1.3 – 1.5 display the frequencies of practice elements and proportional differences in their presence by prevention level type, delivery format, and effect size magnitude. The most common two practice elements overall were psychoeducation and relaxation, regardless of prevention level, delivery type, or effect size magnitude.

Upon reviewing patterns that emerged from the practice element frequency results, we created post hoc groupings of similar skills-based practice elements. Skills-based practice elements that were physiological in nature were: relaxation (69%), physical exercise (25%), and bio/neurofeedback (7%). Skills-based practice elements that were metacognitive in nature were: cognitive monitoring/restructuring (47%), mindfulness (31%), insight-building (26%), and self-monitoring (13%). Skills-based practice elements that were behavioral in nature were: problem solving (21%), communication skills (13%), exposure (13%), activity scheduling (12%), time

management (9%), sleep hygiene (9%), and assertiveness training (6%). Other stress management or coping skills not otherwise specified were present in 31% of programs. Examining programs that produced large effect sizes, skills-based practice elements that were physiological in nature were even more frequently occurring: relaxation (84%), physical exercise (37%), and bio/neurofeedback (16%).

Programs had an average of 4.82 (*SD* = 3.04) practice elements (excluding accessibility promotion from totals, given that it was only coded for group-delivered program). There was a significant difference by delivery format type for number of practice elements ($F = 10.15$, $p < .001$). Bonferonni post hoc tests revealed that self-administered prevention programs had significantly fewer practice elements ($M = 2.20$, $SD = 1.26$) compared with group-delivered programs ($M = 5.22$, $SD = 3.00$, $p = .002$) and online/computer-delivered programs ($M = 6.38$, $SD = 2.90$, $p < .001$). There was also a significant difference by prevention level for number of practice elements ($F = 7.49$, $p = .001$). Bonferonni post hoc tests revealed that indicated prevention programs had significantly more practice elements ($M = 6.39$, $SD = 3.47$) than selective prevention programs ($M = 3.10$, $SD = 2.07$, $p = .001$). Universal prevention programs did not significantly differ in terms of practice elements ($M = 4.76$, $SD = 2.55$) as this average fell between those of indicated and selective programs. Comparing programs that produced small, moderate, or large effect sizes, there were no significant differences in number of practice elements ($F = 0.68$, $p = .6$).

Discussion

This review systemically examined 68 prevention programs targeting depression, anxiety, and/or stress in university students. Our results provide an updated summary of prevention program study characteristics and effect sizes for programs delivered via in-person groups,

online/computer platforms, or self-administered materials at the universal, selective, or indicated prevention level. We have also expanded on past reviews by identifying practice elements common across outcome-producing prevention programs. The findings of this review highlight certain strengths and areas for improvement in prevention programming for university students, as well as provide initial evidence into overlapping program content of effective outcome-producing prevention programs (i.e., programs that produced at least one significant between-group result on an eligible target outcome measure).

Strengths of Reviewed Programs

Some encouraging findings relate to the diversity of outcome-producing prevention programs available to university students as well as the effect sizes produced by these programs. Programs that produced significant symptom reduction were represented in each delivery format type, at each prevention level, and in various combinations of delivery format and prevention level. The overall effect size was moderate (g = 0.65), as were the effect sizes for programs of each delivery format (group, g = 0.69; self-administered, g = 0.65; online/computer-delivered, g = 0.52) and at each prevention level (universal, g = 0.69; selective, g = 0.73; indicated, g = 0.53). The results did not indicate that effect sizes significantly differed by delivery format or prevention level, as confidence intervals for each effect size overlapped. Although the average effect size was moderate, these results would almost definitely have been smaller had we not (a) excluded those prevention programs without significant results and/or (b) selected the largest effect size for each respective program. However, 29% of programs produced a large effect size, which is promising given that these programs were delivered at the prevention level, which typically produce relatively smaller effect sizes than treatment level interventions (e.g., Weisz, Sandler, Durlak, & Anton, 2005). Large effect sizes were produced by some portion of programs

within each delivery format and at all prevention levels, suggesting there is a range of strong prevention program options for depression, anxiety, and/or stress symptoms available to university students.

Practice Elements

The practice elements findings contributed meta-analytic insights that extend beyond results of other similar reviews. We identified elements that are common to these outcome-producing prevention programs, and examined differences in element frequencies by prevention level, delivery format, and effect size magnitude.

Certain practice elements emerged as common across all programs, and some emerged as relatively more frequent among programs of specific prevention levels, delivery formats, and effect size magnitude categories. The practice elements of psychoeducation of symptoms, relaxation, and cognitive monitoring/restructuring skills were common to more than one third of programs. Interestingly, skills-based practice elements of physiological nature were most common, whereas skills-based practice elements of behavioral nature were relatively less common overall. This finding might be specific to prevention programming for university students, as a review of common elements within prevention programs for high school students found cognitive and behavioral skills to be more common for depression and anxiety programming (Boustani et al., 2015).

Because prevention level and delivery format type of programs were confounded, it is difficult to draw conclusions about differences in element frequencies when comparing across these categories. For example, selective prevention programs and self-administered programs had significantly lower frequency of psychoeducation of symptoms. However, given that online/tech-delivered were more likely to be indicated prevention, and group-delivered programs

were more likely to be universal prevention, the higher rates of psychoeduation of symptoms in these types of programs are confounded. Thus, it is unclear whether prevention level or delivery format may be driving such differences in element frequencies. Still, findings suggest that self-administered programs had fewer practice elements per program overall, as well as less coverage of diversity of practice elements. This finding is not surprising, given that the nature of these self-administered prevention programs often had a narrow intervention focus, for example providing students only with relaxation recordings (e.g., Grassi, Gaggioli, & Riva, 2009), or only with a biofeedback machine (e.g., Ratanasiripong, Ratanasiripong, & Kathalae, 2012), and the self-administration emphasized enacting skills rather than instructing about skills. Ultimately, these results outline the current state of practice elements as they generalize across the reviewed prevention programs.

Although the commonness of certain elements does not implicate their effectiveness, we may have more confidence in those practice elements that were frequent among effect-producing prevention programs, as were reviewed in this study. Moreover, the comparison of element frequencies between programs that produced large effects to those that produced moderate or small effects shows compelling associations that describe the content of those programs found to be more effective. Physiological skills (e.g., relaxation) were already more common than behavioral skills (e.g., communication skills) among all programs, and they were even more common for programs with large effect sizes. For programs with large effect sizes, all three physiological oriented skills (i.e., relaxation, physical exercise, biofeedback) were relatively more frequent (by 9-14% compared with groups with moderate effect sizes; by 16-24% compared with groups with small effect sizes), but the disproportions were statistically insignificant. Differences in metacognitive skills were inconsistent: cognitive

monitoring/restructuring was more common to programs with large effect sizes (8% and 25% more than groups with moderate and small effect sizes, respectively), but mindfulness and insight building were less common. These findings call for further investigation into the unique role of physiological skills in prevention programs for university students.

We have not interpreted results to imply specific suggestions for how common practice elements should or could inform intervention design. Analysis of common practice elements within a given literature is one way to identify candidate practices to inform design or adaption of programs for a given context, but not the only way nor necessarily the best way, depending on intervention design goals (Chorpita & Daleiden, 2018). Identifying candidate practice elements is one small part of program design, which also involves consideration of program interface, combination of elements, sequence of elements, dose of each element, conditionality logic, etc. Still, at a minimum, identification of common elements gives an "ingredient list" to consider, where you might not otherwise have one, even though you will still need a "recipe" and the actual resources to select and combine ingredients into the desired product.

Recommendations for Improvement

Some findings pointed to potential areas for improvement. Of note, these areas for improvement are identified within a purely outcome-producing prevention program sample, which is a good reminder that establishing between-group results is not the only important consideration for being an "effective" program.

In many cases, there were target-outcome mismatches such that the identified target(s) for each prevention program would not align with the symptom measures that showed results. For example, a study on an online stress management intervention produced no significant post-intervention between-group results on any stress measures, but did so on an anxiety measure

(Chiauzzi, Brevard, Thum, Decembrele, & Lord, 2008). In contrast, a study that hypothesized its acceptance-based behavioral group therapy would decrease anxiety, depression, and stress symptoms only produced significant changes on the depression subscale of an outcome measure (Danitz & Orsillo, 2014). These findings highlight the transdiagnostic ambiguity among interventions targeting depression, anxiety, and/or stress. These labels also may occasionally be used interchangeably despite distinct symptoms of these three targets. It will be important for future prevention research to disentangle what predicts improvement on one of these outcomes versus another.

In addition, several samples had imbalanced gender ratios. Although there was diversity represented in terms of student type and geographic location, female students were over-represented in the testing of prevention programs. Our results are congruent with other similar meta-analyses that found that most studies have a majority female student sample (Conley et al., 2015; Davies et al., 2014; Louise Farrer et al., 2013; Regehr et al., 2013). These findings are consistent with research that finds that female students are more likely to engage in help-seeking and mental health service use (Eisenberg, Downs, Golberstein, & Zivin, 2009; Eisenberg, Hunt, Speer, & Zivin, 2011). However, the difference in proportion of males and females reporting symptoms of depression, anxiety, and stress is not as extreme – though females are still higher – meaning that the degree of gender imbalance in the reviewed samples is not necessarily an accurate reflection of need for services by males (American College Health Association, 2018). If future research investigates strategies to attract more male students in need, then prevention programs may be more equipped to serve students in need, regardless of gender.

Finally, there were inconsistencies in adherence/completion data reporting. Only about half of the programs had reported adherence data, and there was great inconsistency in how

adherence or completion rates were reported. Thus, it is unclear if the variability in adherence rates is due to how adherence was measured versus how engaged participants were with the intervention. The field could benefit from a deeper understanding of the unique adherence challenges to prevention programming, and how adherence relates to successful prevention program outcomes. Intervention research would also benefit from more standardized reporting of program adherence and completion, while also considering the impact of research-related variables (e.g., participant payment) on program completion. This important data needs to be tracked and transparently reported, in order to further our understanding of the effectiveness of prevention programming. Peer-reviewed journal publishers and reviewers can play an important role in encouraging the reporting of adherence and completion data.

Limitations

The findings of this review should be considered within the context of a few limitations. First, some practice elements were coded with moderate inter-rater reliability. Second, it is impossible to disentangle the source of intervention effects (i.e., which elements, if any, are driving change) from a common elements distillation approach, and the most common elements are not necessarily the most effective elements (Chorpita, Becker, & Daleiden, 2007). It was not possible to run a meta-regression with all elements as predictors of effect size, as the analysis would be underpowered for including so many predictors. Given that interventions are understood as the sum of their practice elements, it was also deemed inappropriate to run a meta-regression with only one practice element at a time as a predictor. Moreover, because the coding of element presence was binary, the data did not capture how extensively an element was covered (e.g., two programs could both have "relaxation" coded, but one program might have included this practice element in a single lesson, whereas the other program might have taught

this skill exclusively). A limitation to the calculation of our effect sizes was that some studies used inactive comparisons while others used active groups, and an average aggregated effect size does not weight these studies differently. A similar review found effect size to be non-significantly different from zero when a prevention program was compared with an active comparison group (Davies et al., 2014), so the effect sizes for each program in our review may have been confounded by comparison group type. Relatedly, because we did not include non-"winning" prevention programs, we could not compare differences in practice elements between programs with or without effect sizes below the small magnitude, i.e., $g < 0.2$, which will be an important extension for future research. Finally, prevention program elements were identified based on coding of articles rather than manuals. We decided to code for practice elements based on content presented within research articles because not all of the included programs used a manual, and so we did not want to unintentionally credit more elements to those programs with manuals. However, a study found that there are some inconsistencies when relying on manuals versus articles when coding for common elements (Knudsen, Boustani, Chu, Wesley, & Chorpita, 2018), thus it is possible that the current findings under-represent frequencies of certain elements.

Future Directions

There are multiple avenues for next steps in this field of university student prevention programming. First, there are other potentially informative parameters by which prevention programs could be categorized in order to compare differences in practice elements. For example, program adherence data could be used to define programs with low, moderate, or high adherence in order to investigate which practice elements are common to programs with higher adherence. Second, in order to understand the relative effectiveness and impact on outcomes of

certain elements, a next step could be to code how each experimental group fared relative to a comparison group on each measured outcome, to determine whether positive outcomes on certain measures of interest are more likely to occur when certain intervention elements are present (e.g., Becker et al., 2018). Third, research that experimentally tests interventions with single practice elements or different combinations of common practice elements could help disentangle which elements are responsible for driving larger effect sizes. After all, some elements that are relatively less common are not necessarily less effective. Fourth, a common element distillation approach could be expanded to reviewing university student prevention programs that target other symptoms (e.g., disordered eating, substance use, traumatic stress). The findings of the current review may be unique to programs targeting depression, anxiety, and stress. For example, substance use rates have been similar for male and female students (American College Health Association, 2018), and therefore sample gender ratios for tested substance abuse prevention programming might not be as disproportional. Fifth, findings from the current study could inform further testing of under-tested practice elements for certain program types. For example, psychoeducation of symptoms and cognitive techniques were common overall, but significantly less often included in selective prevention programs. Finally, in contexts in which new prevention programs must be designed, or existing programs adapted, this review provides a promising and circumscribed set of practice element options to consider for program content selection, rather than selecting intervention content at random or based on a slice of relevant the research literature.

Conclusion

Decades of research and intervention development have produced dozens of outcome-producing prevention programs for university students. Further progress in the prevention of

depression, anxiety, and stress symptoms relies not just on the creation of new interventions but also on the evaluation of existing programs in order to learn from the current state of the evidence. There is already an abundance of prevention programs that have been tested. Therefore, there may be diminishing returns in a sole focus on creating new programs, rather than furthering our understanding of what works and applying what we already know (e.g., Chorpita et al., 2011). With such a wealth of research to review, we are well positioned to further explore specific intervention strategies that are a part of effective prevention programs. The reviewed programs were available in a variety of formats and all produced significant effects – moderate effects on average – despite some limitations (e.g., gender imbalanced samples, variable adherence rates). With a better understanding of what elements are common to effective prevention programs, we can begin to use the knowledge created by past research inform the future of mental health promotion for university students.

Table 1.1

Code Definitions and Corresponding Cohen's Kappa Values

Code	Description	Kappa
Prevention Program Level		.62
Universal	offered to any student in the population, focused on promoting behaviors that generally reduce negative outcomes. An example of a universal prevention program would be a psychoeducational program for all students to learn how to identify depression symptoms in themselves and others.	
Selective	offered to subsets of the population – identified by age, sex, or other distinguishable characteristics – that are expected to uniquely experience benefits of prevention programming. The subgroup of population is selected for a reason related to intervention purpose. An example of a selective prevention program would be a program that targets depression through communication skills and therefore is only delivered to students currently in romantic relationships.	
Indicated	offered to individuals with a sub-clinical level of symptoms related to the intended intervention target. Often a study testing an indicated prevention program will use the inclusion criteria related to a cut-off score. An example of an indicated prevention program would be a program that is tested on any students with anxiety scores in the "at risk" range.	
Delivery Format		.76
Group-based	delivered in-person to multiple participants at the same time	
Online/computer-delivered	self-delivered by working through content one a website or a computer program, remotely or in a designated location	
Self-administered	delivered via written materials (exercise instructions, handouts, books, manuals) or audio/video materials that participants used on their own time. Some programs included one initial individual session for the purposes of explaining to the participant how to conduct the intervention on their own	
Target		
Anxiety	general anxiety, worry, performance anxiety, social anxiety, perfectionism, and/or panic	.81

Depression	major depression, depressed mood, negative affect	.86
Stress	psychological stress, physiological stress, or distress	.86
Practice Elements		
Activity Scheduling	Skills to facilitate involvement in enjoyable, rewarding, or enriching activities and experiences	.65
Accessibility Promotion	Strategies to increase convenience and accessibility for participating in program	.52[a]
Assertiveness Training	Exercises or techniques designed to promote the student's ability to assert his or her needs appropriately with others	.65
Biofeedback/ Neurofeedback	Strategies to provide information about one's own physiological activity, often done with specialized equipment	1.00
Cognitive Monitoring/ Restructuring	Techniques designed to monitor and alter unhelpful thoughts	.72
Communication Skills	Training in how to communicate more effectively with others to increase positive functioning, increase consistency, or minimize stress.	.89
Exposure	Techniques or exercises that involve direct or imagined experience to practice approaching feared situations	.67
Goal Setting	The explicit selection of a goal for the purpose of working towards it throughout the program	.65
Insight Building	Activities specifically designed to help achieve greater self-understanding about one's emotions, reactions, triggers, or preferences	.66
Maintenance/ Relapse Prevention	Exercises and training designed to consolidate skills already developed and to anticipate future challenges	.41
Mindfulness	Exercises designed to facilitate present-focused, non-evaluative observation of experiences as they occur, with a strong emphasis on being "in the moment"	.89
Modeling**	Demonstration of a desired behavior, typically presented by a program leader, program materials, or peers	.65
Motivational Enhancement	Exercises designed to increase readiness to participate in the programs, such as cost-benefit analysis, persuasion, or Socratic questioning	.42

Relaxation	Techniques or exercises designed to induce physiological calming, including muscle relaxation, breathing exercises, and meditation	.72
Peer engagement*	Encouragement to interact with or learn from peers who experience the same target problem	.65
Performance Feedback	Providing information about one's own or another's performance, based on assessment or observation	.68
Problem Solving	Training in the use of techniques, discussions, or activities designed to bring about solutions to targeted problems	.67
Psychoeducation: services/resource awareness*	Providing information about available services and resources available to participants, relevant to their needs	.63
Psychoeducation: symptoms	The formal review of information about a problem area, the development of a problem, and/or its relation to a proposed intervention	.62
Self-Monitoring	The repeated measurement of one's own symptoms, feelings, behaviors	.61
Sleep hygiene*	Specific problem solving skills for sleep difficulties.	.88
Stress management/ Coping skills*	Strategies to help build skills to cope and deal with stressors, not otherwise specified	.46
Time management*	Specific problem solving skills for time management issues.	.71

Notes.
* Elements not in the PracticeWise Clinical Coding manual at the time of this review
** Elements with adjusted definitions to better match prevention programs
[a] Only coded for in-person group-delivered interventions. The nature of online/computer-delivered or self-administered interventions is to be accessible but not all studies of these interventions necessarily described accessibility-promoting strategies explicitly within their protocol description section. Therefore, it was deemed that coding the presence or absence of this practice element would be ambiguous for online/computer-delivered or self-administered interventions.

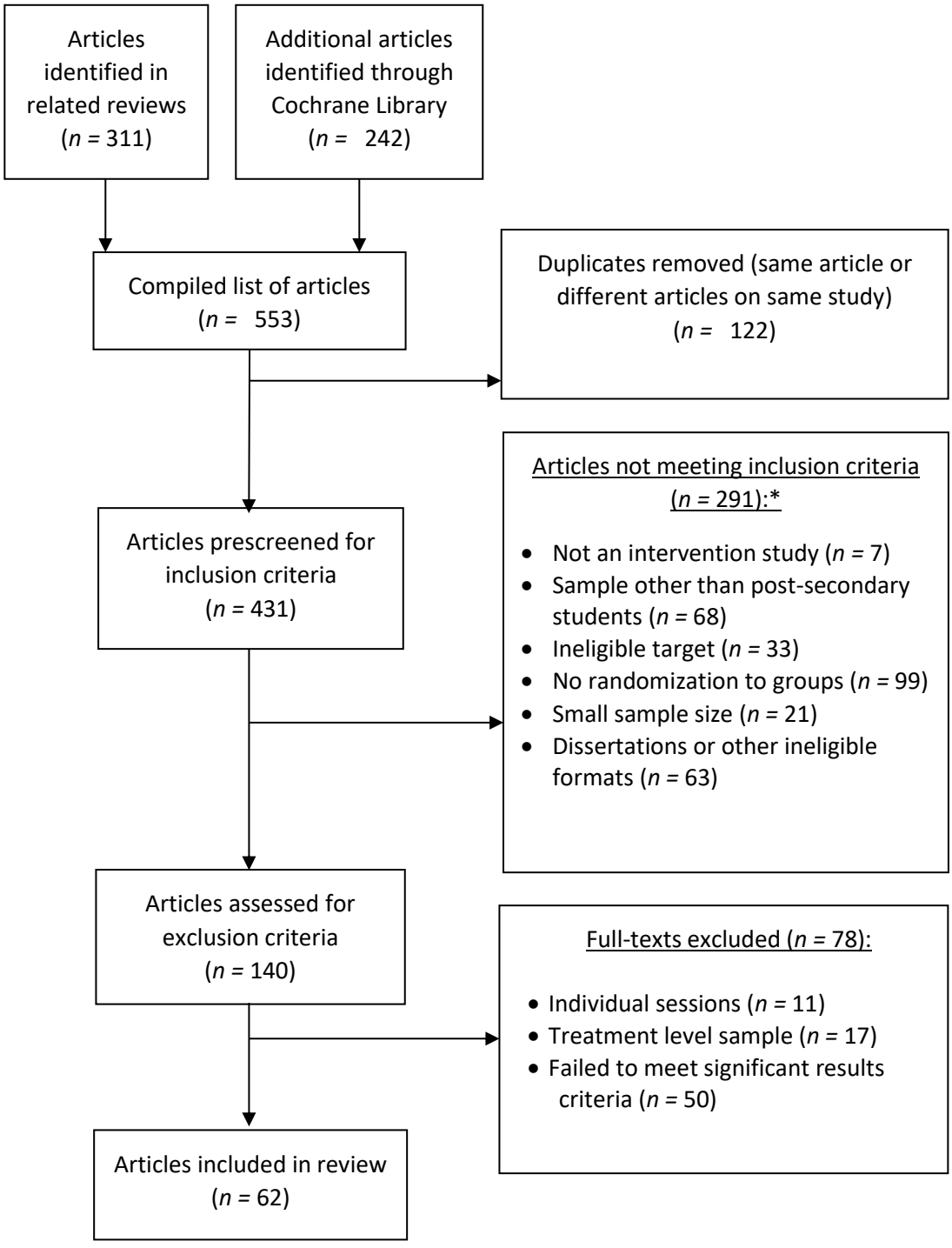


Figure 1.1

Flowchart of article identification, inclusion, and exclusion.

Note: Articles may have failed to meet multiple inclusion criteria, but we report their removal only once, according to this order of inclusion criteria.

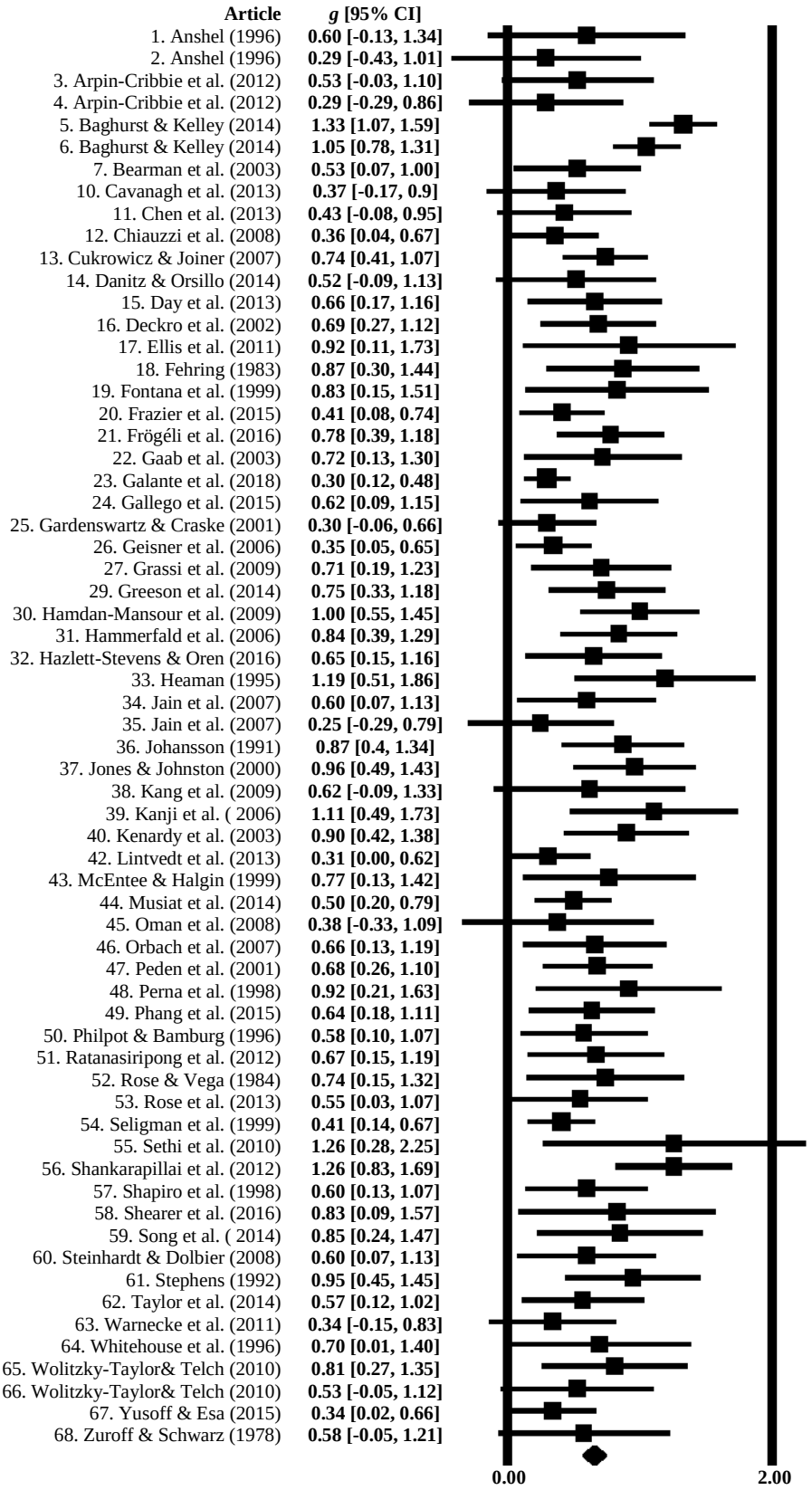


Figure 1.2
Forest Plot of Program Effect Sizes (Hedges' *g*)

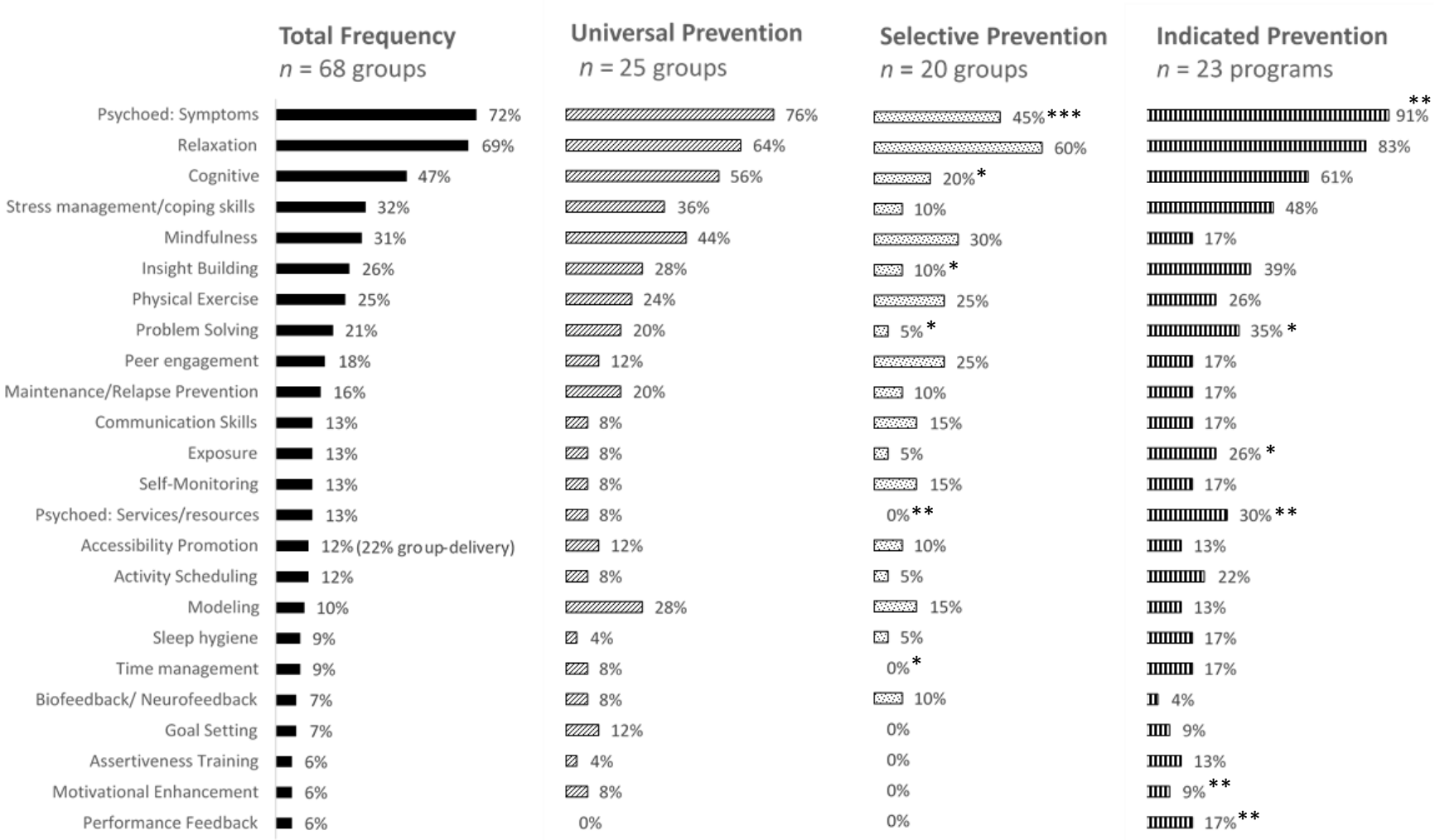


Figure 1.3
Frequency of elements common to all groups and to groups of different prevention levels

Note. Chi-square tests (or likelihood ratios for analyses based on estimated cell counts < 5): * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

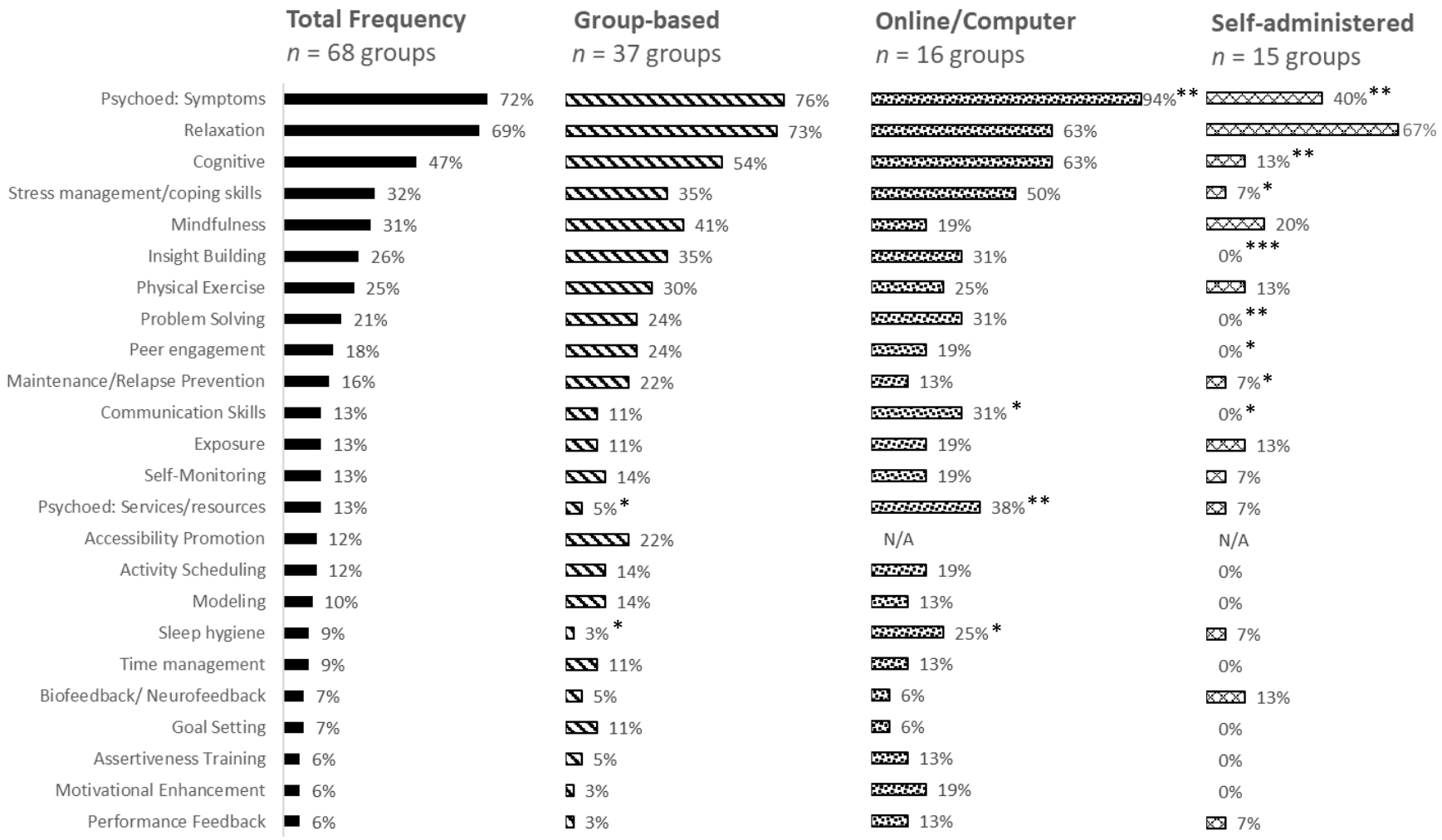


Figure 1.4
Frequency of elements common to all groups and to groups of different delivery types

Note. Chi-square tests (or likelihood ratios for analyses based on estimated cell counts < 5): * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

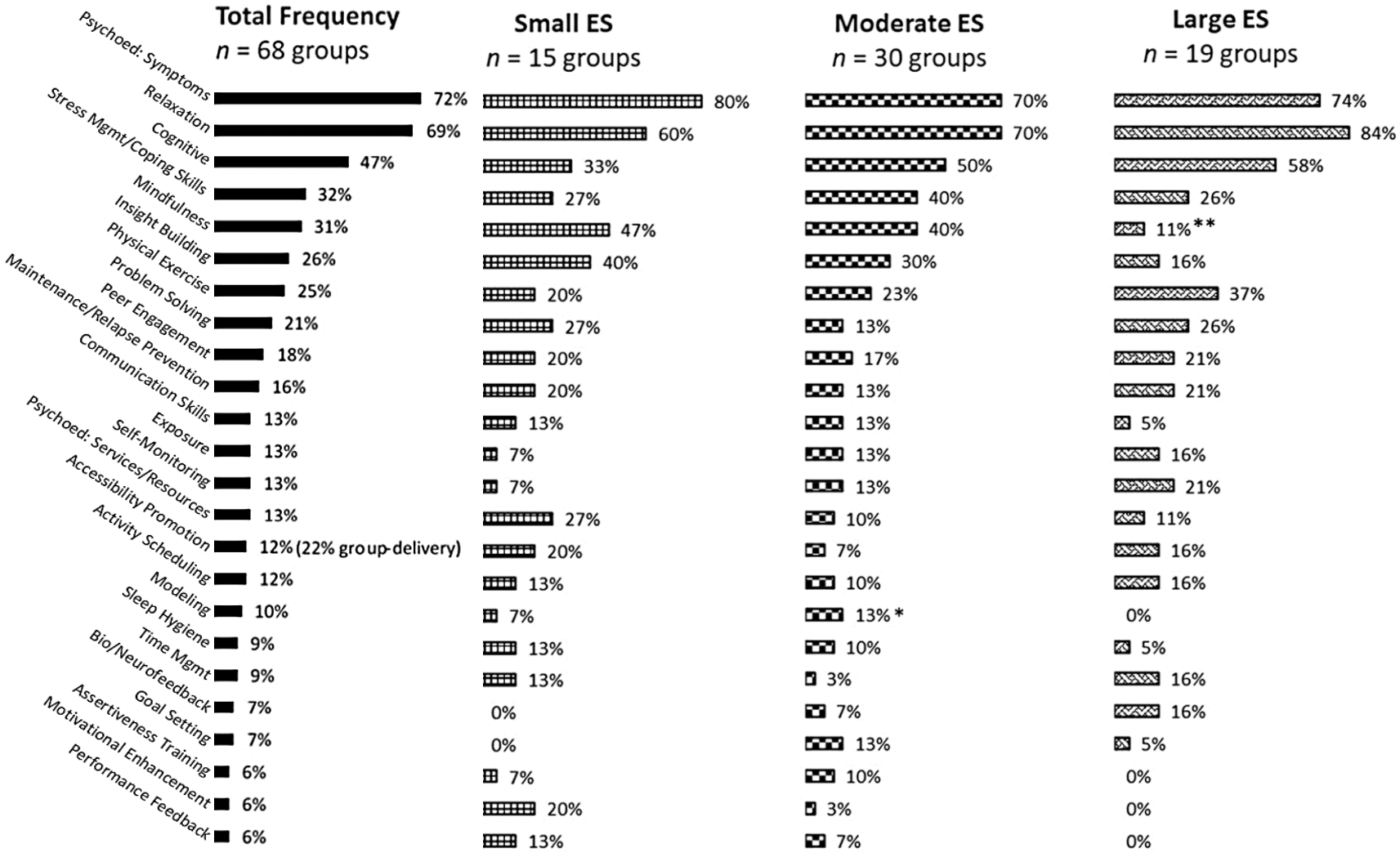


Figure 1.5

Frequencies of practice elements and proportional differences by effect size magnitude.
Note. Chi-square tests (or likelihood ratios for analyses based on estimated cell counts < 5): * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

Chapter 2: What's in a Name?: Branding of Online Mental Health Programming for University Students

Abstract

Objective: University students experience many help-seeking barriers and thus not all students who could benefit from mental health services enroll in them. This study aimed to examine student enrollment in response to strategic marketing of an online prevention program for anxiety and depression. **Method:** Data were collected from students at two universities during recruitment phases for the online program. The program was branded as "The Happiness Challenge" or "ReBoot Camp" through parallel sets of recruitment materials using language intended to address help-seeking barrier concerns (e.g., stigma, inaccessibility). The yielded samples were examined for unaddressed psychological need rates, demographic composition, and differential enrollment by student subgroups into either program brand. **Results:** Replicated results between Study 1 (n = 651 students, 71.2% undergraduate, 80.3% female, 27.9% white non-Hispanic) and Study 2 (n = 718 students, 60.6% undergraduate, 73.4% female, 53.2% white non-Hispanic) showed that: more than a third of students qualified as having "unmet need" for services; enrollment was disproportionately self-identified female and Asian students; Asian students were less likely to report prior service use and more likely to be categorized as having "unmet need"; "ReBoot Camp" was disproportionately selected by male students. **Conclusion:** Findings suggest that recruitment effectively reached students with unaddressed mental health need, including high enrollment by Asian students, who historically seek services less often. Additionally, important gender differences emerged in preferences for program name. These findings could inform how to market services in university settings in order to reach more students, including those from underserved subgroups.

Introduction

Lifetime prevalence rates of psychological disorder are highest during the ages that students are attending university, yet fewer than half of students with anxiety or mood disorders seek treatment, leaving many with unmet need (Blanco et al., 2008; Hunt & Eisenberg, 2010; Zivin et al., 2009). There are a number of help-seeking barriers that university students experience – lacking awareness of services (i.e., what services are available or how to access them); stigma about mental illness; concerns with confidentiality and trust; inability to self-assess symptoms; accessibility issues (e.g., not having enough time, inconvenient location, cost); and preference for self-reliance (Eisenberg, Golberstein, & Gollust, 2007; Gulliver, Griffiths, & Christensen, 2010; Yorgason, Linville, & Zitzman, 2008). To reduce this service need-use gap, exploration of commercial marketing for student mental health services has been recommended (Yorgason, Linville, & Zitzman, 2008).

However, marketing of psychosocial interventions has gone relatively unexplored. In contrast, public health efforts have widely adopted marketing campaigns since the 1990s that target young people's health behavior (e.g., diet, drug use, skin cancer prevention, etc.; Andreasen, 2004), and the pharmaceutical industry has spent more than 4 billion annually on consumer marketing since the mid-2000s (Kornfield, Donohue, Berndt, & Alexander, 2013; Mackey, Cuomo, & Liang, 2015). A direct-to-consumer marketing approach has been recommended to increase the uptake of psychological interventions, because information about available evidence-based services is otherwise solely provided to treatment providers (Becker, 2015). There is already evidence that direct-to-consumer marketing increases helping-seeking for pharmaceutical interventions (Gallo, Comer, & Barlow, 2013) and initial evidence that psychological interventions may do so as well (Gallo, Comer, Barlow, Clarke, & Antony, 2015).

Beyond simply increasing awareness of services, marketing also needs to increase student interest by presenting services in an appealing way. Effective social marketing must convey to potential users why the potential benefits of use are greater than the perceived costs, or barriers (Andreasen, 2004). Most help-seeking barrier concerns can be addressed through intervention descriptions (e.g., listing a low financial cost; highlighting accessible timing and location). Because mental health stigma can prevent students from even seeking information on health services (Lannin, Vogel, Brenner, Abraham, & Heath, 2015), resource offerings may need to be described without explicit labeling of "mental health problems." One study found that college students were more likely to engage in self-help if they read an advertisement describing the life of a person with depression (without naming the disorder), rather than an advertisement that simply listed depression symptoms (Chang, 2008). Thus, in order to reach students typically deterred by help-seeking barriers, we need to understand how to effectively advertise campus services through strategic descriptions and destigmatizing language.

Some student subgroups are disproportionately less likely to seek help from mental health services, including male, Asian, black, and Hispanic students (Eisenberg et al., 2011; Rickwood, Deane, Wilson, & Ciarrochi, 2005). Male students have also been underrepresented in research studies examining university student prevention programming (**Chapter 1**). In particular, stigma barriers are higher for male and ethnic minority students (Clement et al., 2015), as is the preference for self-reliance a barrier for young men (Ellis et al., 2012). Male and ethnic minority students still experience mental health concerns though (Eisenberg, Hunt, & Speer, 2013; Mahmoud, Staten, Hall, & Lennie, 2012); thus, their lower service use rate does not reflect a lack of need. To ultimately reach more students, research should examine how service marketing attracts students universally or differentially.

Services that are advertised in a way that reduces barrier concerns will likely attract more students overall, but we can also use branding to specifically target under-served student groups. Student interest in counseling center services has been found to differ not only in how services are described in brochures (Rochlen, Blazina, & Raghunathan, 2002), but also in response to the name of a counseling center (Brown & Chambers, 1986). Marketing research on other consumer products has amassed insight into gender preferences, which has implications for psychosocial service branding to attract male students. For example, technology products are perceived as more masculine (Lieven, Grohmann, Herrmann, Landwehr, & van Tilburg, 2014). Even subtle name differences can change gender associations, with brand names with more stop-consonants (e.g., "k", "t") rather than fricative-consonants (e.g., "h", "s") perceived as more masculine (Klink, 2000). We are not aware of existing research investigating differential student enrollment in response to intervention name branding, and we believe it is a worthwhile help-seeking engagement strategy to explore.

The Current Research

This study aims to examine the enrollment of students into an online anxiety and depression universal prevention program in response to strategic recruitment efforts. We take a marketing approach by advertising the prevention program in a way that addresses common help-seeking barriers, such as accessibility concerns, preference for self-reliance, and stigma. The program was advertised under two distinct names, "The Happiness Challenge" and "ReBoot Camp." These two names were selected based on previous pilot data and review of general marketing research literature. When the program had been first piloted two years previously to this study, it was called "The Happiness Challenge" and the enrolled sample was 90% female (Rith-Najarian, 2015). Because males are historically less likely to use mental health services,

attention was given to attracting this subpopulation. “ReBoot Camp” was added because this name purposefully sounds more technology-oriented, and it contains more stop-consonants (ReBoot Camp) than fricative-consonants (The Happiness Challenge).

A universal online prevention program was deemed a particularly appropriate intervention type for our research aim. Offering programming online – rather than in person – allows for service use in private, anywhere/anytime, and through self-help. Indeed, young people commonly cite accessibility, anonymity, and lower stigma as reasons for their use of online services (Kauer, Mangan, & Sanci, 2014). The accessibility factor is especially important, as online counseling service requests are usually made outside standard business hours (Richards, 2009). Regarding a universal-level intervention, research findings suggest that universal programs elicit less stigma than indicated prevention programs (Rapee et al., 2006). Prevention programming fortunately also lends itself to larger-scale delivery, and prevention programs for anxiety and depression in particular can be far-reaching given these symptoms are the most common mental health concerns among university students (ACHA, 2016; Eisenberg, Hunt, & Speer, 2013). Students in distress have indicated greater intentions to use online universal prevention programming over formal service options (Ryan, Shochet, & Stallman, 2010), and one third of young people already use the internet for finding information about mental health problems (Burns, Davenport, Durkin, Luscombe, & Hickie, 2010). Therefore, offering an online prevention program can leverage this common help-seeking behavior among young people. Marketing strategies are particularly pertinent for prevention programs, into which individuals self-select, as opposed to treatment-level interventions in which treatment selection is often made jointly with a therapist or doctor.

Overall, we sought to investigate if our branding and marketing language approaches

would reach students who were (a) in need of services, (b) representative of the student body population, and (c) differentially attracted by branding. Specific research questions were:

- How many students in our sample have elevated symptoms? How many and which students in our sample have not used mental health services? These research questions had no a priori hypotheses. Instead, we planned to descriptively examine the proportion of enrolled students who self-reported anxiety/depression symptoms while also reporting no prior use of mental health services in order to determine if our marketing approach successfully reached students with unaddressed psychological need.
- Does the sample of university students enrolling in a prevention program demographically represent the respective population? Because our recruitment approach was designed to address barriers that differentially deter certain students, we expected our sample to produce a generally representative sample in terms of demographics.
- Does "The Happiness Challenge" versus "ReBoot Camp" attract students disproportionally from certain demographic groups? It was expected that ReBoot Camp would attract relatively more males. Other comparisons by demographic variables were planned as exploratory analyses, and no other a priori hypotheses were made.

To address these research aims, data was collected at two universities during the online program's enrollment phase. Given our interest in the effects of prevention program marketing and branding in a real campus environment, using controlled conditions was deemed impractical. For example, students cannot be randomized to only encounter one set of materials or the other around the campus. We instead used a naturalistic data collection design that was replicated at two different U.S. universities, one a public university on the West Coast and the other a private university on the East Coast. Using different student cohorts was intended to increase confidence

that any significant results were generalizable beyond a single student sample. Ultimately, answering our research questions could provide insight into the importance of how mental health services are presented to their intended target population.

Study 1 Methods

The data for this study were collected in spring 2016 at a large public university in California during the recruitment phase of an open trial for the online prevention program for anxiety and depression. The 8-week online program delivered content via emails and the website. Each weekly module teaches skills such as behavioral activation, physical exercise, cognitive restructuring, and mindfulness. Students can select which skills (and how many) they want to practice each week and can access relevant resources. For example, if the current weekly module is "mindfulness," a student would read about the benefits of practicing mindfulness, read a list of meditations and mindfulness skills options with instructions, select the strategies they want to test, and use relevant worksheets and audio recordings. To facilitate engagement, the program incentivizes skills practice through prize drawings and sends practice reminder emails throughout the week. The study was approved by the UCLA Institutional Review Board (Protocol ID #15-001724).

Participants

Inclusion criteria were: 1) age 18 or older; 2) currently enrolled as an undergraduate, graduate, or professional student; and 3) enrolled in the online program. Participation in the research study was voluntary, and students could enroll in the prevention program regardless of research involvement. Research participation was incentivized with entry into a gift card prize drawing. Of the 1,655 students that enrolled in our prevention program online, 671 students provided online consent at the beginning of the research survey, of which 651 were eligible to

participate in our research study. The 20 ineligible students were excluded because they either not meet eligibility criteria (n = 9), knew someone on research team (n = 7), or straight-lined survey responses (i.e., a respondent indicated the same rating for all survey questions, even when questions were assessing opposite experiences; n = 4). Full demographic details of the resulting sample are presented in Table 2.1.

Procedure

The recruitment materials were disseminated through a mass email from the registrar, student group email-lists, campus wellness newsletters, flyers in dormitories and other campus location, and social media posts. All recruitment materials prominently addressed help-seeking concerns commonly experienced by male and female students: online accessibility and timing convenience ("use it when you want, where you want"), cost ("It's free!"), and preference for self-reliance ("set your own goals and work on them week-by-week"). The program was described as a "habit change program" for "all students" in order to reduce help-seeking barriers related to stigma about mental illness. We offered the program to all students (rather than indicating the program was targeting only certain students), given the evidence that universal programs are less stigmatizing (Rapee et al., 2006). The benefits of the program were described in terms of improving behaviors rather than symptoms. For example, "procrastination" and "avoiding uncomfortable situations" were mentioned, as procrastination and avoidance are related to both anxiety and depression (Flett, Blankstein, & Martin, 1995; Ottenbreit & Dobson, 2004). It was thought describing behaviors in this way would remove the barrier of self-assessing symptoms, and would also be less stigmatizing than saying "anxiety" and "depression" symptoms. In fact, studies have found that people have less stigmatizing attitudes towards "burnout" than "depression," even if they are describing the same construct (Bahlmann,

Angermeyer, & Schomerus, 2013; Bianchi, Verkuilen, Brisson, Schonfeld, & Laurent, 2016). See Figure 2.1 for demonstration of how these elements were included in materials. See Appendix 2A for full content of example materials.

The program was always advertised under both names "The Happiness Challenge" and "ReBoot Camp." Each version had identical content, but the tone of written content and the visual appearance for the recruitment materials and online program interface were slightly different to match the brand names (See Figure S1 in Appendix 2A). For example, The Happiness Challenge description mentioned "focusing on your happiness" whereas ReBoot Camp was described as functioning like a "drill sergeant" (See Figure S2 in Appendix 2A). Otherwise, program descriptions were identical. Recruitment materials were mirrored with the two names and their respective logos and materials for both programs were posted in similar places; see Figures S1-S5 in Appendix 2A for examples. Interested students could visit a central sign-up website that provided more information about the programming, and there they could enter their account set-up info for the program of their choice.

Participants provided online consent by completing the pre-program survey for the open trial. In the survey, participants responded to symptoms measures, reported campus service use, and indicated which of two programs they selected, which we verified by cross-checking with our website's database. Students also provided their student ID number to allow us to link their research responses to the demographic information in their student record. The list of student IDs in our research study was provided to the campus research office, and they returned the various student associated demographic variables. Population-level demographic information for the entire student body was also provided by the campus research office.

Variables

Demographics. Student record data provided demographic information on gender, age, student status (undergraduate versus graduate), ethnicity, and undergraduate/graduate academic area (e.g., physical sciences, humanities). The population-level data for the campus was also based on student records, therefore ensuring the same definition of demographic variables in our sample as for the population. Ethnicity data from student records was based on self-reported identification with ethnic categories defined by the Common Data Set (CDS Advisory Board, 2018). For gender, although the student record system allowed students to select non-binary identification options for internal uses, it required students to select "Male or "Female" for aggregate reporting purposes. Therefore, the gender data accessible for this study was binary.

Clinically elevated symptoms. Anxiety was measured with the Spielberger State Trait Anxiety Inventory – State (STAI-S) scale, and depression was measured with the Patient Health Questionnaire-2 (PHQ-2). Clinically elevated anxiety was defined by the recommended cut-off score of 40 (Knight, Waal-Manning, & Spears, 1983) and a positive depression screen was defined by a score of 2+ on either item 1 or 2 of the PHQ-2 (Löwe, Kroenke, & Gröfe, 2005).

Use of counseling center. Using a self-report checklist of campus services and resources, students could endorse one of three options for each resource: "I don't know what it is," "I know it, but haven't used it," or "I've used it". Student were categorized as a service-user if they endorsed using the university counseling center for individual services, group therapy, or mental health skills workshops.

Data Analysis

We used two-tailed z-tests to test for differences in demographic proportions in our student sample versus the campus student body population. Within our sample, we used chi-squared analyses to test differences in proportions by demographics (e.g., gender, ethnicity) on

respective outcomes (e.g., prior service use, program name selection). Chi-squared analyses of gender by ethnicity were non-significant (*p*-values ranging .24 – .87), which justified conducting subsequent chi-squared analyses using gender or ethnicity as independent variables in their own separate analyses.

Study 1 Results

Participants. Our sample (*n* = 651) was comprised of 71.2% undergraduate students (28.8% graduate and professional student) and 80.3% females. There was complete baseline survey data for 534 students, and the remaining 117 students had either (a) partial survey data or (b) demographic data only. See Table 2.1 for full sample demographic details.

We also compared the initially enrolled sample to the program completer sample in order to check if the demographic make-up of the enrolled sample differed substantially from the completer sample. A program completer was defined as a student who had logged program activity through the last program week and had missed no more than two program modules. Chi-squared analyses showed the demographic make-up (i.e., gender proportions, ethnic proportions, clinically elevated symptom status, service use) of the starting sample did not significantly differ from the program completer sample (*p*-values ranging .13 – .87).

Symptoms and mental health service use. Of participating students, 60.7% had clinically significant symptoms: 3.5% had clinically elevated depression alone (PHQ-2), 39.6% had clinically significant anxiety alone, and 17.6% had both. Of students in our sample, 69.4% had never used the campus counseling center. Of the students that had elevated symptoms, 63.3% had never used the counseling center. Thus, 38.6% of the full Study 1 sample was categorized as having "unmet need." Running chi-squared tests by each ethnic group individually, we found that Asian students were significantly less likely to use the counseling

center, X^2 (1, 534) = 9.25, p = .002. No other chi-squared tests were significant by gender (p = .32) or ethnic group (p-values ranging .09 – .80). We further examined the subsample of clinically elevated, non-service using students (N = 219) compared with the other students in the sample. Using chi-squared analyses, we found that these clinically elevated, non-service using students were no more or less likely to be female, Black, Hispanic, Multiracial, or White (p-values .12 – .86), but there were proportionally more Asian students in this group, X^2 (1, 534) = 3.96, p = .047.

Representativeness of sample. We ran z-tests to compare the proportions of ethnicity, gender, academic department affiliation, and undergraduate/graduate student status of our student sample to those of the entire student body. Differences between the demographics of our sample and the full student body included significantly higher proportions of: female students (z = 15.71, p < .001) and Asian students (z = 2.26, p = .02).[1] There were no other differences in the yielded student sample by other ethnic categories or undergraduate/graduate student status. A comparison of all sample demographics versus population (i.e., campus study body) demographics is presented in Table 2.1.

Name selection. The Happiness Challenge was selected by 390 participants (59.9%) and ReBoot Camp was selected by 261 participants (40.1%) in our sample. Compared with students who selected The Happiness Challenge, students who selected ReBoot Camp were more likely to be male, X^2 (1, 548) = 7.22, p = .007. There were no significant differences by ethnicity (p-values ranging .18 – .91).

Before examining differences in program name selection by academic major, we ran a

[1] After reviewing Study 2 results, we also post hoc examined the difference in proportions of Asian students in our sample versus in the student body for undergraduates only, and again there were disproportionately more Asian undergraduates in our sample (4.8% more, z = 2.09, p = .04).

chi-squared test to check for disproportionate representation of gender across academic areas. There were gender differences by academic area, X^2 (5, 546) = 30.37, $p < .001$, and therefore differences in name selection by academic area were examined separately for males and females. For males, academic major predicted program enrollment, with those in the physical sciences disproportionately selecting ReBoot Camp, X^2 (1, 107) = 5.53, $p = .02$. For females, those in the life sciences disproportionately selected The Happiness Challenge X^2 (1, 439) = 6.24, $p = .01$. There were no other significant differences by academic department for males (*p*-values ranging .16 – .88) or females (*p*-values ranging .16 – .80).

We ran post hoc chi-squared analyses to check clinical equivalency between participants selecting each brand. There were no differences by elevated symptom status ($p = .28$) nor by service-user status ($p = .75$).

Study 2 Methods

The data for this study were collected in fall 2017 at a private university in Connecticut. Data collection for this study occur during an enrollment phase for the same online program. Otherwise, the data collection procedure was nearly identical to that of Study 1. The study was approved by the Yale Institutional Review Board (Protocol ID #2000020985).

Participants

Inclusion criteria were the same as for Study 1. De-identified data from the prevention program platform were supplemented with data from surveys completed for research voluntarily. Research participation was incentivized with either entry into a gift card prize drawing, or credit for a class research requirement. Of the 718 students who created online platform accounts, 265 consented to participate in the research component, of which 260 provided usable data. The omission of survey data for five students was due to unreliable survey responses (i.e., either

straight-lining responses or highly discrepant scores on validation item pairs, such as providing the same rating to items with reverse scoring). Full demographic details of the resulting sample are presented in Table 2.2.

Procedure

The recruitment materials were disseminated through a mass email to all students through a comprehensive undergraduate student group email-list, dorm newsletters, campus-wide graduate and professional newsletters, and fliers in residencies and other campus buildings. Given that recruitment at the campus in Study 1 involved emailing all enrolled students, we ensured that all undergraduate, graduate, and professional students at the second university site also received program recruitment emails. All undergraduate students were emailed directly through a student group email-list twice and all graduate students were emailed first through a graduate student newsletter and then through comprehensive graduate student group email-list. The second university site adapted direct copies of the original marketing materials with campus-specific information. See Appendix 2A. This direct adaptation was crucial for the purposes of replicating the first study findings, as altering marketing materials could produce different patterns of student yield, and the changes would be confounded with campus and student population differences. Accordingly, all recruitment materials again prominently addressed the same help-seeking concerns, and both "The Happiness Challenge" and "ReBoot Camp" versions were advertised. Parallel sets of fliers for either "The Happiness Challenge" or "ReBoot Camp" were distributed in the same campus building to reduce possible demographic biases due to building usage.

All recruitment materials directed users to a web page that explained the two versions of the same program and provided instructions with a unique code. The web page directed users to

a separate platform site, where users could then sign-up for the program of choice using the unique code. Some participants then provided online consent for the additional research survey component, which included the same questionnaires as in Study 1. One difference was in how demographic data was collected. In Study 1, the university site had a mechanism to provide researchers with student record data, while the university site for Study 2 had no such mechanism. De-identified data from the online platform database was used for all demographic and brand name analyses. Population-level demographic data for the student body were based on publicly available campus data for the respective academic year during which data collection took place. Research survey data was used for any analyses involving symptom measures and use of counseling services.

Variables

Demographics. Gender, student status, and ethnicity were collected from students' account information on the intervention platform, which was based on self-reported demographic identification during the account set-up process. Because the platform does not collect data on age or international status, these demographic variables were not included for Study 2. Population-level ethnicity data was thus extracted with international students collapsed within self-reported ethnicity categories. Self-reported gender options allowed students to select "Male", "Female," or "Other." The publicly available population-level demographic data for the campus was also based on self-reported identification, with ethnic categories based on the Common Data Set (CDS Advisory Board, 2018) and gender categories based on a binary definition.

Clinically elevated symptoms. Again, clinically elevated anxiety was defined by the recommended cut-off score of 40 (Knight, Waal-Manning, & Spears, 1983) and a positive

depression screen was defined by a score of 2+ on either item 1 or 2 of the PHQ-2 (Löwe, Kroenke, & Gröfe, 2005). One limitation of Study 1 was reliance on PHQ-2 alone for depression symptom assessment. We had used the PHQ-2 with trigger logic to prompt the full PHQ-9 for students with elevated PHQ-2 scores in order to shorten the online survey, thus reducing survey burnout and increasing completion rates. However, the PHQ-2 has shown variable sensitivity (Gilbody, Richards, Brealey, & Hewitt, 2007), which could affect results. Therefore, Study 2 included the full PHQ-9 for all students in order to compare if symptom elevation rates change by use of PHQ-2 versus PHQ-9. Survey burnout was of less concern for Study 2 because we were able to streamline other components of the data collection. A positive depression screen using the full PHQ-9 was defined by a score of 10+ (Kroenke, Spitzer, & Williams, 2001).

Use of mental health services. One limitation of Study 1 was that the use of the campus counseling center does not represent service use of any mental health services, as some students may use services off-campus. Therefore, we asked students about which services they "have used or are using" through a self-reported checklist, including (a) campus counseling center services and (b) any mental health services, on- or off-campus (e.g., local mental health centers, off-campus therapy/psychiatry). A student was categorized as a service-user if endorsed any listed mental health services.

Data Analysis

Analyses were planned again as two-tailed z-tests and chi-squared analyses. Chi-squared analyses of gender by ethnicity were again non-significant (*p*-values ranging .10 – .81), allowing for the separate testing of them as independent variables in Study 2 as well.

Study 2 Results

To determine replication of findings, the same main analyses as were run in Study 1 were

run using this sample, with only some additional or omitted specific analyses based on difference in data collection procedure.

Participants. The sample (n = 718) was comprised of 60.6% undergraduate students (39.4% graduate and professional students) and 73.4% females (74.5% female of binary-identifying students). See Table 2.2 for full sample demographic details.

Symptoms and mental health service use. Of the 260 students who completed the full research survey, 69.4% had clinically significant symptoms: 0.4% had clinically elevated depression alone (PHQ-2), 52.0% had clinically significant anxiety alone, and 16.9% had both. When using the PHQ-9 instead to define clinically elevated status, exactly 69.4% still had clinically elevated symptoms, with 0.4% having clinically elevated depression alone (PHQ-9), 39.9% having clinically significant anxiety alone, and 29.0% having both. Therefore, the same percentage of students demonstrated elevated symptoms by either definition, but more students fell into the "both elevated anxiety and depression symptoms" category with the latter definition.

Of the research survey sample, 72.3% had never used the campus counseling center and 70.0% had never used any on- or off-campus mental health services. Of the students with elevated symptoms, 66.8% had never used the campus counseling center and 63.4% had never used any mental health services. Thus, 44.0% of the full Study 2 sample was categorized as having "unmet" need. Female students were significantly more likely to use the counseling center, X^2 (1, 256) = 6.41 p = .01 or any mental health services, X^2 (1, 256) = 8.35, p = .004. Running chi-squared tests by each ethnic group individually, we found that Asian students were significantly less likely to use the counseling center, X^2 (1, 260) = 9.03, p = .003 or any mental health services X^2 (1, 260) = 11.78, p = .001. No other chi-squared tests were significant by ethnic group (p-values .053 – .87).

We next compared the subsample of clinically elevated, non-service-using students (N = 109) with the other students in the sample. Running chi-squared tests by each ethnic group individually, we found that Asian students were significantly more likely to be in this unmet need group: counseling center non-use definition: X^2 (1, 250) = 6.19, p = .01; any mental health service non-use definition: X^2 (1, 250) = 8.75, p = .003. White students were less likely to be in this unmet need group when defining non-use by any mental health services X^2 (1, 250) = 5.40, p = .02. No other chi-squared tests of clinically elevated, non-service-using students were significant by gender or ethnic group (p-values .12 – .70).

Representativeness of sample. We ran z-tests to compare the proportions of ethnicity, gender, academic department affiliation, and undergraduate/graduate student status of our student sample to those of the entire student body. Differences between the demographics of our sample and the full student body included significantly higher proportions of: female students (z = 12.93, p < .001) and undergraduate students (z = 8.75, p < .001). Given that ethnic composition in the student body did vary by undergraduate/graduate status at the population level (p-values all < .001 for each ethnic group), and that our sample is significantly under-representative of graduate students, the ethnicity results are likely confounded. Therefore, we examined proportional differences by ethnicity for undergraduate students only, as they constituted the majority of the Study 2 sample. Examining ethnicity proportions of the sample versus the population for undergraduates only, we found significantly higher proportion of Asian students in our sample (z = 2.83, p = .005) but no differences by other ethnic groups (p-values .14 – .98). A comparison of all sample demographics versus population (i.e., campus study body) demographics is presented in Table 2.2.

Name selection. The Happiness Challenge was selected by 430 participants (59.9%) and

ReBoot Camp was selected by 288 participants (40.1%) in this sample. Students who selected ReBoot Camp were more likely to be male, X^2 (1, 706) = 8.59, p = .003, compared with students who selected The Happiness Challenge. Examining ethnic groups individually, Black students were more likely to select ReBoot Camp, X^2 (1, 718) = 3.83, p = .05, compared with non-Black students. There were no other significant differences by ethnicity (p-values ranging .28 – .92).

Before examining differences in program name selection by academic major, we ran a chi-squared test to check for disproportionate representation of gender across academic areas. Although there were not gender differences by academic area (p = .18), to parallel Study 1 we examined differences in name selection by academic area separately for males and for females. For males, academic department predicted program enrollment, with those in the physical sciences and engineering disproportionately selecting ReBoot Camp, X^2 (1, 56) = 4.41, p = .04. There were no other significant differences by academic department for males (p-values ranging .09 – .59) or females (p-values ranging .51 – .96).

We ran post hoc chi-squared analyses to check clinical equivalency between participants selecting each brand. There were no differences by elevated symptom status (p-values ranging .88 – .90), counseling center use (p = .62), nor by broad mental health service use (p = .80). We also examined research survey participation equivalency and found that students selecting either brand were not significantly more likely to have participated in the research survey (p = .14).

Assessment of Results Replication

To compare results between Study 1 and Study 2 we followed a replication analysis procedure proposed by Asendorpf and colleagues (2013) to assess: (a) which (if not all) results are in the same direction; (b) if the significant findings for both data sets are found for the same

respective variables; and (c) if the effect sizes for each test statistic from Study 2 are within the confidence interval for that test statistic from Study 1.

First, we assessed which results were in the same direction. For a summary of the results comparison, see Table S1 in Appendix 2B. For both schools: more than a third of participants had elevated symptoms while reporting no prior service use (33.5%; 44.0%); Asian participants were less likely to report prior use of a counseling center; Asian participants were over-represented in the clinically elevated, non-service using students; the majority of participants were undergraduates (71.2%; 60.6%); the majority of participants were female (80.3%; 74.4%) and were overrepresented relative to the respective campus population; Asian undergraduate students were overrepresented relative to the respective campus population; the majority of students selected "The Happiness Challenge," (59.9%; 59.9%); male students were more likely to select "ReBoot Camp"; and academic discipline predicted enrollment into "ReBoot Camp" for males.

Second, of those results with common directions across Study 1 and Study 2, we identified those that were analyzed through statistical tests and significant at the $p = .05$ level or lower in both studies. Both schools found significant results of: lower counseling center use by Asian students; higher categorization of Asian students as those with "unmet need"; disproportionately high enrollment of female students; disproportionately high enrollment of Asian undergraduate students; higher enrollment of male students into ReBoot camp; and higher enrollment of male students in physical sciences/engineering into ReBoot Camp.

Finally, the effect sizes of those common significant results across Study 1 and Study 2 were compared. For z-test analyses, the effect was calculated as difference in percentage of the sample from the percentage of the population ($\%_p$), and the 95% confidence intervals were calculated

based on the difference between the two proportions. For chi-squared analyses, the effect size and the 95% confidence intervals were calculated using the conversion formula to produce phi coefficients (Olivier & Bell, 2013). All significant results in Study 1 (outlined above) had an effect size confidence interval that contained the respective effect size from Study 2, with one exception being that Study 1 had even greater disproportionately high enrollment of females. See Table 2.3 for a comparison of effect sizes and confidence intervals for each study's common significant results.

Discussion

This study examined the enrollment of students at two difference campuses into an online universal prevention program for anxiety and depression in response to strategic marketing efforts in order to investigate the impact of intervention name branding on student enrollment. We were able to replicate findings for our research questions related to: (a) baseline elevated symptom rates and prior mental health service rates (b) the representativeness of the enrolled samples relative to respective student populations; and (c) any differences in type of students enrolling in one program name versus the other. Because there was no experimental manipulation in our studies, causality cannot be determined. However, in terms of feasibility, our findings highlight the importance of mental health program branding and advertising, and the replication of findings across two studies suggests external validity of our approach for purposes of reaching more university students.

We offered a highly accessible intervention – an online universal prevention program – for the two most common psychological disorders in university student populations, and we advertised it in such a way as to allay common barriers to help-seeking and attract different students through brand naming. We wanted to know if our intervention successfully attracted

students with elevated symptoms and/or without prior service use, as such findings would support the ability of such an online intervention to serve as an alternative campus resource to those in need, and especially those with unmet need. At both campuses, the majority of students enrolled in the intervention demonstrated elevated symptoms, the majority had not used other formal mental health services, and the majority of those with elevated symptoms had not used the services. The findings remained the same in Study 2 even when a different depression scoring was used and when the mental health services category was more expansive. These results support the ability of online prevention programming advertised with an emphasis on addressing help-seeking barriers to attract students with otherwise unmet need.

We also expected to see a demographically representative sample enrolled, as opposed to a sample with underrepresentation by those students who have been disproportionately deterred by help-seeking barriers. With regard to gender, both samples were composed of disproportionately more females, which is consistent with help-seeking trends by gender (Galdas, Cheater, & Marshall, 2005; Kohl, Crutzen, & De Vries, 2013). Thus, the tested intervention and its advertisement did not seem to attract students equally by gender. With regard to ethnicity, both samples were generally representative except that there were disproportionately more self-identified Asian students enrolled in our intervention relative to the proportion of self-identified Asian students at each campus. The Common Data Set (CDS Advisory Board, 2018) defines "Asian" as a "person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent," thus this definition combines ethnic identities of Asian, Asian America, and otherwise Asian students. Asian students are one of an ethnic minority groups that are otherwise less likely to seek campus mental health services historically (e.g., Eisenberg et al., 2007; Sontag-Padilla et al., 2016). Asian students in both study samples

were less likely to have used mental health services and were more likely to be categorized as having "unmet need" at entry, which remained true in Study 2 when definitions were based on (a) an abbreviated versus a full depression measure or (b) on-campus versus any (on- or off-campus) service use. These findings suggest that the online intervention in Study 1 and 2 was offered in a way that generally attracted students from all ethnic subgroups equally, and perhaps overcame unique barriers for Asian students related to traditional campus in-person services. Ultimately, the description of services in campus advertisements may play an important role in attracting students in need.

Asian students in particular may have been over-represented in our sample because they were for the first time being offered a mental health service that better addressed their particular help-seeking barriers. Personal stigma has been found highest among Asian students (Eisenberg, Downs, Golberstein, & Zivin, 2009). There are also specific barriers for Asian Americans, including: cognitive barriers (e.g., information about mental health and illness is processed differently), affective barriers (e.g., reactions to emotional problems are more automatically negative), and value-related barriers (e.g., there are norms of not discussing emotions with others; Leong & Lau, 2001). Although other ethnic minorities might experience similar cultural values and beliefs, a review found that only for Asian Americans is help-seeking deterred as a result of the cultural values of social norms, collectivism, emotional self-control (Sun, Hoyt, Brockberg, Lam, & Tiwari, 2016). It is possible that the marketing of the online intervention tapped into these cultural values. For example, it was offered for "all students" to do "together" in order to reduce stigma concerns, but this language innately probes values of social norms and collectivism. There was also much language emphasizing the program's self-guided nature, which could appeal as a means of getting help while still retaining a sense of control.

Finally, we examined if "The Happiness Challenge" versus "ReBoot Camp" differentially attracted students of certain demographics, in order to assess in name branding could play a role in reaching student subgroups. As hypothesized, we found that male students in both samples disproportionately selected "ReBoot Camp." Importantly, the samples were 19.7% and 25.6% male, respectively, compared with 10.4% male enrollment when the intervention was previously offered under "The Happiness Challenge" alone (Rith-Najarian, 2015). Therefore, offering the program under a different name seems to have increased the enrollment of male students, despite the samples still being disproportionately female. Again, there was no experimental manipulation of brand name choices, but the successful replication of these findings in two studies with slightly different sign-up pathways supports that there is a relationship between gender preferences and intervention name selection. "ReBoot Camp" may have appealed more to male students by evoking a sense of emotional stoicism through its military tone. Emotion restriction has been conceptualized as a socialized masculine norm that is also associated with military culture and has been found to contribute to stigma against help seeking (Heath, Seidman, Vogel, Cornish, & Wade, 2017). Therefore, a mental health program that has a more emotionally stoic name and tone may be relatively more approachable for male students. In addition, we found that males from the physical sciences/engineering were even more likely to have enrolled in "ReBoot Camp" over "The Happiness Challenge." Given that we developed "ReBoot Camp" to sound more technologically-oriented – a general marketing strategy to convey masculinity – it is not surprising that males in more technologically-oriented fields were particularly preferential towards "ReBoot Camp." Overall, the key takeaway of these branding results is that an intervention's name matters and can be strategically crafted in order to target service engagement by specific population sub-groups.

Our research findings should be interpreted with some limitations in mind. First, because not all program participants completed the supplemental research survey of either study, there is no guarantee that the findings based on survey-collected variables (e.g., reported service use, symptom levels) are representative of students in the respective full samples (for which we had demographic and brand selection data). However, both samples are still relatively large. Second, the under-enrollment of graduate students in Study 2 may have been due to recruitment method differences. Study 1 emailed all students a stand-alone advertisement, whereas for communications in Study 2, undergraduate students were mailed the stand-alone advertisement twice in two weeks, while graduate students received the advertisement first within a newsletter containing other listings, and then in a stand-alone format two weeks later. This temporal delay may have led to diminished effectiveness of the stand-alone advertisement in the graduate students. Thus, we decided post-hoc to analyze undergraduates separately for the sample-population comparisons in Study 2; it is unknown whether full sample results would have been replicated had there been proportionally enrolled graduate students. Third, the sample's demographic data for Study 2 was based on self-report, meaning ethnicity and gender variables may have been self-defined differently than the registrar-defined demographic variables used for sample-population comparisons. Fourth, although Study 2 expanded the definition of "service-user" to include off-campus mental health services, both studies' definitions of "service-user" did not explicitly assess for non-local mental health service use. Given that survey phrasing in both studies did not emphasize a timeline to consider for past service use, both studies may have missed categorizing some students as "service-users" if they did not consider mental health services used in their hometown, abroad, or in any other non-local location. Finally, gender was defined in a binary way, as constrained by university demographic reporting practices, and

therefore students identifying other than "male" or female" were not accurately represented by our gender data, limiting the interpretation and generalizability of our gender-related findings.

The naturalistic data collection is both a limitation and a strength of the current study. Because the data were not collected through an experimental design, we cannot conclude causality of marketing and branding strategies on enrollment. However, given that our current research sought to offer mental health services online in order to attract students that do not use typical in-person services, it would not have been ecologically valid to use an in-person laboratory study design. Randomizing students to encounter different materials virtually was also deemed unrealistic, as we intended to use "all students" marketing language and contact the full study body at campus. If students discussed the recruitment emails with friends or saw fliers for the other program, randomization would be disrupted, and tracking such disrupted cases through self-report would not have been dependable. While experimental design may have been more appropriate for establishing causality, the external validity of our findings for real-world campus implementation would have been limited. Nevertheless, without experimental manipulation, our results do not demonstrate the effects of students' exposure to one strategy or the other, but rather, students could be exposed to many of our marketing materials and both program brands. Regarding the two different names, it is possible that presenting both names simultaneous is an advantage – students appreciate being given a choice – or a disadvantage – students who do have a strong preference are deterred by the juxtaposition. Therefore, we do not yet understand how students would have responded if we offered "ReBoot Camp" alone. It is worth considering experimental approaches that future research might take in investigating differential student response to mental health service branding. For example, randomization of branding conditions at a campus- or cohort-level within a larger multi-site or a longitudinal study could increase

internal validity of findings without sacrificing external validity. Limitations notwithstanding for the current research, the naturalistic data collection and replication of results within two different student populations provides us with greater confidence about the generalizability of our findings for campus-wide implementation efforts.

Overall, our findings highlight the importance of considering not just which services we offer on campuses, but how we advertise them. Although the current study has provided more insight into how students differentially respond to online intervention marketing, there is much that remains to be investigated. First, more research is needed to understand how to reach non-Asian ethnic minority students who are also underserved by formal campus services. To do so, it will be necessary to investigate how unique help-seeking barriers for each ethnic group may be better addressed through a variety of marketing initiatives. Second, future research would benefit by using non-binary gender definitions to investigate the role of gender in help-seeking barriers and response to marketing strategies for students of all gender identities. Third, from an implementation effort standpoint, it will be important to establish at what point there is diminishing return from adding more marketing strategies or branding a program through more names. Based on the current findings, employing programming under two names seems promising for attracting more diverse participation, but there are likely students in-need who will remain un-reachable through such strategies. Fourth, it would be worthwhile to research the application of the same marketing and branding approaches of the current study to formal campus services. Although the advertisements could not promise the same accessibility as online interventions, services could still be branded under different names and described in a way that more explicitly addresses help-seeking barriers. Finally, investigating the replication of findings at universities internationally will elucidate further nuances in student response.

Considering how much time and effort are invested into intervention research and development, we need to ensure that the programs are in fact attracting those young people who are in need. Successful dissemination and implementation of prevention programs for university students will require designing an intervention that is not just effective, but also has effectively considered and addressed help-seeking barriers during the service initiation phase. Given that students of various demographics are differentially deterred from help-seeking, it is important to better understand preferences when advertising for campus mental health resources and services. Reducing disparities in help-seeking and service use among students will require multiple systemic changes as well (e.g., making culturally-tailored service options available; reducing negative public opinions and stigma about mental illness), but branding of services is one strategy that contributes to the solution. The more we understand about how the advertising approach of an intervention impacts the enrollment of its users, the better we can tailor efforts to maximize service use.

Table 2.1

Comparison of Study 1 Sample Demographics to Full Population Demographics

	% Student Body	% Sample
Ethnicity		
American Indian or Alaskan Native	0.2	0.4
Asian	26.2*	30.5*
Black Non-Hispanic	3.2	2.6
International	14.6	11.7
Hispanic	17.8	18.8
Pacific Islander	0.3	0
Two or More Races	4.6	5.7
White Non-Hispanic	30.7	27.9
Unknown/Other	2.5	2.6
Gender		
Female	46.6**	80.3**
Male	53.4**	19.7**
Student Status		
Undergraduate	68.7	71.2
(Mean Age)	(20.6)	(20.7)
Graduate	31.3	28.8
(Mean Age)	(28.0)	(28.3)

z tests of proportions: *$p < .05$; **$p < .01$

Table 2.2

Comparison of Study 2 Sample Demographics to Full Population Demographics

Variable	Student Body		Sample	
Ethnicity	%	% of undergrads	%	% of undergrads
American Indian or Alaskan Native	0.4	0.7	0.1	0.2
Asian	24.4	21.4**	25.5	27.3**
Black Non-Hispanic	6.9	8.0	5.8	7.4
Hispanic	10.3	13.0	8.9	10.7
Pacific Islander	0.1	0.1	0.1	0.2
Two or More Races	4.9	6.6	6.3	7.1
White Non-Hispanic	53.0	50.8	53.2	47.0
Unknown/Other [a]	-	-	-	-
Gender (binary students only)	% by binary	% by non-binary	% by binary	% by non-binary
Female	49.6**	-	74.5**	73.4
Male	50.4**	-	25.5**	25.1
Other	-	-	-	1.5
Student Status	%		%	
Undergraduate	43.9**		60.6**	
Graduate/Professional	56.1		39.4	

[a] This category was omitted from results (and the total denominator adjusted accordingly) because the population-level data were available for only "Unknown," but no students were categorized as "Other," whereas the sample-level data allowed students to omit a response (i.e., "Unknown") or select "Other."

z test of proportions: * $p < .05$. ** $p < .01$.

Table 2.3

Comparison of effect size confidence intervals for Study 1 and Study 2

	Study 1	Study 2	Study 2 test statistics within Study 1 CI?
Positive percent difference of females in sample vs. in population	34.2% [30.9, 37.5]	24.8% [21.5, 28.1]	No; however, both CIs above 0
Positive percent difference of Asian students in sample vs. in population (undergraduate only)	4.8% [0.1, 9.5]	5.8% [1.3, 10.2]	Yes
Asian students disproportionately reporting no counseling center use	r = .13 [.05, .21]	r = .18 [.06, .30]	Yes
Asian students disproportionately representing non-service users with elevated symptoms*	r = .09, [.001, .17]	r = .16 [.04, .28]	Yes
Male students disproportionately selecting “ReBoot Camp”	r = .12 [.03, .20]	r = .11 [.04, .18]	Yes
Students in physical sciences and engineering disproportionately selecting “ReBoot Camp” (males only)	r = .23, [.04, .40]	r = .27 [.01, .50]	Yes

Notes: *For comparison between Study 1 and Study 2 we selected the test statistics produced by definitions relying on PHQ-2 scores and campus counseling center scores to ensure methodological equivalency.

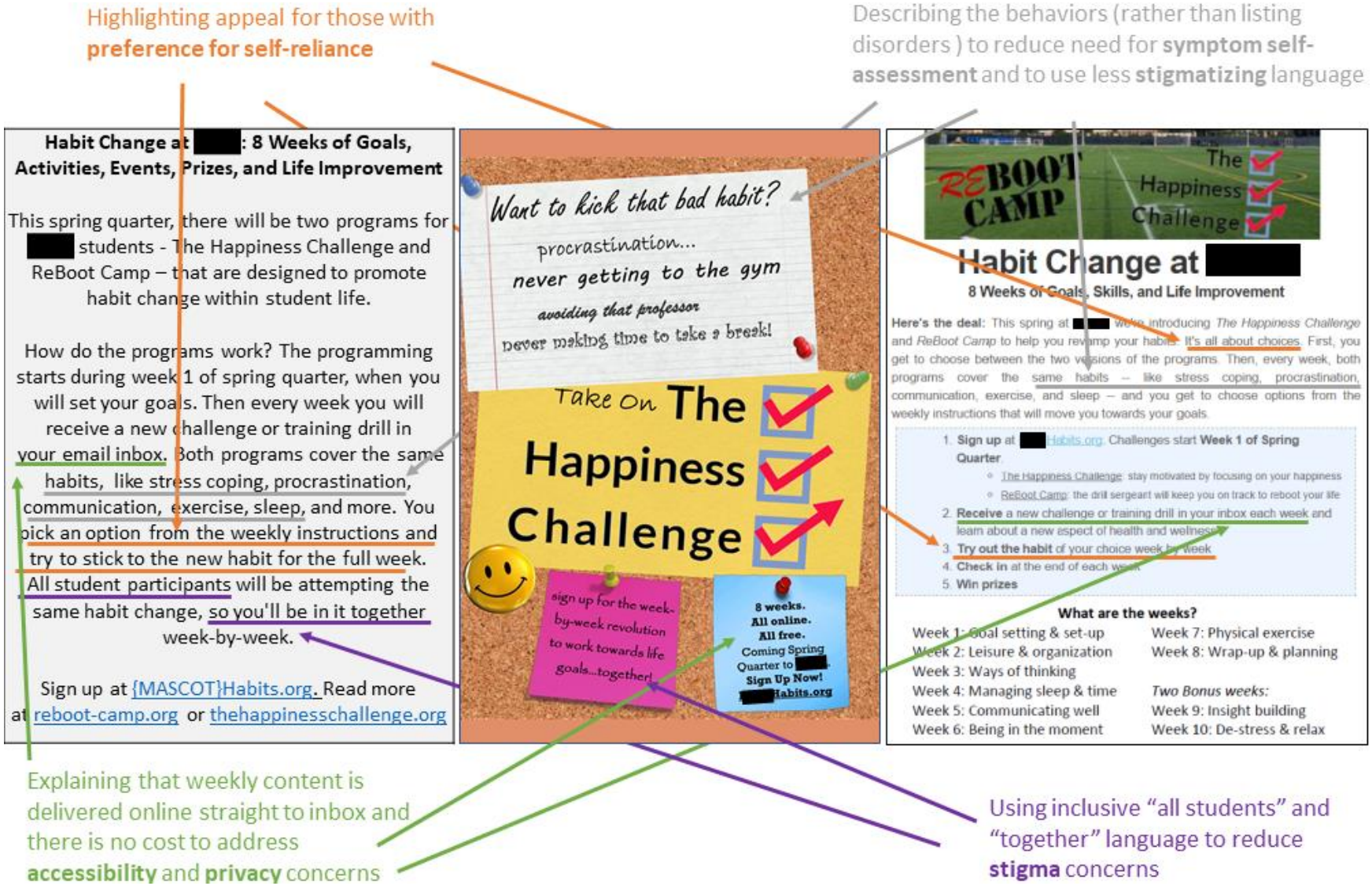


Figure 2.1
Demonstration of how recruitment material language addressed known help-seeking barriers.

Chapter 3: Open Trial Feasibility Study of an Online Universal Prevention Program for Anxiety and Depression in University Students

Abstract

Background: Anxiety and depression prevalence rates are high among university students, yet many students experiencing symptoms at a subclinical or clinical level do not use traditional mental health services. This study investigated a novel online universal prevention program for anxiety and depression that was designed for group-level delivery and was developed with engagement-enhancing features. Feasibility was examined in terms of: (a) recruitment strategy yields, (b) retention and adherence rates, (c) program acceptability, and (d) outcome assessment procedure sensitivity. **Method:** University students enrolled in the online intervention (n = 651) completed a pre-program survey, weekly check-in surveys, and a post-program survey. Recruitment source data and baseline symptom measures (STAI, PHQ, PSS) were collected in the pre-survey. Adherence and use of optional program features were assessed using data from the weekly check-in surveys. Program feedback and post-program symptom measures were collected after the 8-week skills-based online program. **Results:** Of seven recruitment sources, campus-wide recruitment emails were the most effective recruitment strategy (82% of students). There were 72.7% of students who initiated the program, with 10.6% of students fully adhering to the 8 weeks of the program. Program acceptability was demonstrated by high rates of optional dialogue support feature use (e.g., 79.8% of program completers used the module tips and suggestions), high satisfaction rates (e.g., 71.2% endorsed the program as "useful"), and common qualitative themes that were identified. Pre-post changes in anxiety and depression symptom reduction were detected. **Conclusion:** Findings support feasibility of the intervention and research procedures, and implications are relevant for the future development and researching of such online interventions.

Introduction

The developmental transition from adolescence to adulthood is a critical phase for identity formation (Arnett, 2000; Schwartz et al., 2005) and potentially for the onset of psychopathology (Conley et al., 2014; Schulenberg et al., 2004). During this transition, both undergraduate and graduate students report significant mental health concerns associated with the stressors and expectations specific to the university environment (Eisenberg, Gollust, Golberstein, & Hefner, 2007; Hyun, Quinn, Madon, & Lustig, 2006). Given that anxiety and mood disorders are the most prevalent mental health concerns among university students (e.g., Auerbach et al., 2018), interventions that target such internalizing symptoms have the potential to positively impact a large number of students. However, despite increasing student demand at campus counseling centers (Reetz, Bershad, LeViness, & Whitlock, 2016), there is evidence that student service users are still just a fraction of those students struggling with psychological disorders (Blanco et al., 2008; Ketchen Lipson, Gaddis, Heinze, Beck, & Eisenberg, 2015). Therefore, to target student anxiety and depression effectively, one way to complement traditional campus mental health services it to explore other far-reaching intervention options such as universal prevention programs. Prevention programming at a universal-level is intended to promote health behaviors that generally produce more benefits than costs for everyone in the population (Gordon, 1983). Given that all enrolled students constitute a clearly defined population, a campus setting is well positioned for universal prevention programming.

Group-Level and Online Delivery

For universal prevention programming, there are a number of advantages of a group-level delivery approach, rather than one-on-one delivery. Offering intervention to all students at the same time as a campus-wide campaign allows for lower stigma concerns, social cohesion and

support, and ease of synchronized implementation. There is some evidence to suggest that universal prevention programs elicit less stigma than indicated prevention programs (Rapee et al., 2006). Moreover, group-level delivery allows participants to sign up at the same time as peers and friends, which can provide social facilitation for engaging with the program. Research has indeed found that university students are more willing to seek help and use mental health services if they know a friend or relative who has used services (Disabato, Short, Lameira, Bagley, & Wong, 2018; Vogel, Wade, Wester, Larson, & Hackler, 2007). Finally, from an administrator perspective, offering the same resources, to all students, at the same time, on a regular schedule allows for more efficient planning and resource allocation. Accordingly, delivering universal prevention programs at a group-level may maximize their reach.

Another delivery approach that would have unique advantageous for campus-wide universal prevention is offering programming online, rather than in-person. Recently online and computer-based prevention programs for university students have been increasingly developed for depression, anxiety, and stress (Davies et al., 2014; Louise Farrer et al., 2013). Given university students propensity to use the internet for health information (Escoffery et al., 2005; Rennis, McNamara, Seidel, & Shneyderman, 2015), online-delivered interventions are a natural extension of existing student habits. Online-delivery offers easier accessibility (regarding time and location), privacy, lower cost, and allowance for self-reliance, all benefits which are particularly important for addressing common help-seeking barriers among young adults (Gulliver et al., 2010). In terms of implementation benefits, online programming has the potential for larger reach and scalability. Importantly, there is also evidence of effectiveness of online or computer-based prevention programs for anxiety, depression, and stress prevention in university students; a review of such programs found that the prevention groups outperformed

inactive comparison groups (Davies et al., 2014). Together, online-delivery is promising in terms of implementation benefits as well as effectiveness for symptom improvement.

Group-level delivery and online-delivery each offer advantages for universal prevention programming on campuses, yet existing online mental health programs for university students are typically designed for non-concurrent individual delivery while the existing group-level prevention programs use in-person group formats (**Chapter 1**). Online-delivered programs for university student anxiety, depression, and stress have been found to include peer engagement less often (**Chapter 1**), and when they do it is in the form of peer stories, testimonials, or message boards (e.g., Chiauzzi, Brevard, Thum, Decembrele, & Lord, 2008; Ellis, Campbell, Sethi, & O'Dea, 2011; Frazier et al., 2015) rather than all peers receiving intervention content that is temporally synchronistic campus-wide. Therefore, an intervention that combines concurrent group-level delivery with an online format remains to be tested for the universal prevention of anxiety, depression, and stress for university students.

Intervention Design Based on Prior Research

Though group-level online prevention programming may be novel in terms of delivery format, the development can still be informed by existing online programs that have already demonstrated effectiveness. To do so, we can capitalize on existing research about design principles, content, and engagement strategies for effective online mental health interventions.

In terms of general intervention framework, reviews of online health intervention programs have found that typical online mental health intervention designs organize content into independent modules and last for 9-11 weeks (Clarke, Kuosmanen, & Barry, 2014; Davies et al., 2014; Kelders et al., 2012). Reviews have also found that skills-based programs produce larger effects sizes for health behavioral change, especially when more skills are taught (Webb, Joseph,

Yardley, & Michie, 2010; Clarke et al., 2014). A skills-based intervention using modular design might be particularly advantageous for an online intervention with group-level delivery. Modular design allows program content to be structured into independent units that can be rearranged or omitted, and thus comprehension of one module's content is not contingent on other modules' content (Chorpita, Daleiden, & Weisz, 2005b). The benefit of such "information hiding" (i.e., participants need not know what is in other modules in order to benefit from the current one) is that even if students miss or skip some weeks, they can still make use of content from past and future weeks. Such flexibility allows for variability in participant adherence during synchronized group delivery, which is necessary given that monitoring/enforcing individual-level adherence would be impractical for campus-wide intervention. The benefit of modular design's "partial decomposability" (i.e., ordering of modules can be changed based on implementation needs) could facilitate intervention adaptation for different campus contexts such that content can be delivered in a different order or with certain modules omitted based on unique student population needs. Overall, it seems that a promising program design for effective prevention of anxiety and depression in young adults would incorporate the best relevant skills-based practices into 9-11 weekly modules.

In terms of the content for intervention modules, such decisions can be informed by the existing evidence base. It has been argued that there is no "one way" or "best way" for intervention design, yet part of the reason why interventions have positive effects has to do with whether they have a high concentration of practice elements that are commonly derived from relevant treatment literature (Chorpita & Daleiden, 2018). Accordingly, skill modules for an original intervention could correspond to those practice elements that are common among previously tested prevention programs targeting on anxiety, depression, and/or stress in

university students. One review of such prevention programs found that certain practice elements are indeed relatively common (e.g., psychoeducation, cognitive restructuring, mindfulness; **Chapter 1**), providing an initial "ingredient list" to consider. To compliment knowledge from the evidence base with local consumer-derived knowledge (e.g., Chorpita & Daleiden, 2018), final selection and prioritize of modules from that "ingredient list" could be informed by feedback from focus groups of intended target audience.

Finally, in addition to designing the intervention with skills for symptom reduction, it is also important to build an intervention that is sufficiently engaging. An important downside of online interventions is that they typically have worse participant adherence and moderate to high non-completion rates (Christensen, Griffiths, & Farrer, 2009; Clarke et al., 2014; Kelders et al., 2012). A review on persuasive technology in web-based programs for physical and mental health explored a number of adherence-enhancing features, some of which could be incorporated into online anxiety and depression prevention programs for university students (Kelders et al., 2012). Higher adherence to web-based programs was associated with online programs that had more frequent: (a) contact with a professional clinician, (b) intended usage (once a week being standard), (c) updates (new content or lessons becoming available), and (d) dialogue support (Kelders et al., 2012). Kelders and colleagues (2012) considered dialogue support for online interventions as: (1) praise, (2) rewards, (3) reminders, (4) suggestions, (5) descriptions seeming specific to participants, (6) aesthetic appeal, and (7) a system that adopts a social role (e.g., buddy, coach). Universal prevention programs preclude frequent clinician interaction, which is not feasible for large-scale implementation. However, it would be practical for online programs to incorporate more dialogue support features, encourage at least weekly use, and provide regular content updates. Many of these features have been missing in existing prevention programs for

university student anxiety/depression (Cavanagh et al., 2013; Musiat et al., 2014), however given existing evidence, their inclusion could facilitate adherence in universal prevention programs.

The Current Study

The current study is an open trial study, designed to examine the feasibility of an 8-week online modular program for universal prevention of anxiety and depression in university students. The intervention was designed to incorporate: (1) practice elements for skills training, as identified in existing effective anxiety/depression programs for university students (**Chapter 1**), (2) sequential weekly access of content (i.e., weekly modules released one at a time, rather than all content freely available at one time), (3) use of multiple dialogue support strategies, and (4) delivery at the group-level.

Although an intervention's effectiveness cannot be established without scientifically rigorous RCTs, there is much value in initially conducting an open trial, such as the current study. Prior to studying novel interventions through expensive and more time-intensive effectiveness research trials, it is important to first investigate feasibility through pilot studies (Leon, Davis, & Kraemer, 2011). It has been argued that a pilot study design is appropriate for the evaluation of a novel intervention's feasibility in terms of recruitment strategies, retention, outcome assessment procedures, and satisfactory implementation (Leon et al., 2011). More recently, recommended research questions for feasibility studies include (Orsmond & Cohn, 2015):

(1) Can we recruit appropriate participants? (e.g., recruitment rates are sufficient);

(2) How appropriate are the data collection procedures and outcome measures for the intended population and purpose of the study? (e.g., outcome measures are seemingly sensitive to the effects of the intervention);

(3) Are study procedures and intervention suitable for and acceptable to participants? (e.g., intervention adherence rates are adequate);

(4) Does the research team have the resources and ability to manage the study and intervention? (e.g., study can be conducted with the designed budget);

(5) Does the intervention show promise of being successful with the intended population? (e.g., qualitative data suggest the intervention is promising)

Both of these approaches to feasibility study design (Leon et al., 2011; Orsmond & Cohn, 2015) have guided researchers investigating novel online mental health interventions.

Accordingly, feasibility components to be assessed in the current open trial include: (a) recruitment yield from various strategies; (b) retention rates and treatment adherence; (c) program acceptability (i.e., platform feature use, participant satisfaction, and qualitative feedback), and (d) outcome assessment procedure sensitivity (i.e., detection of pre-post change through online data collection of self-reported symptom measures). Although pre-post data collection in a non-randomized design is insufficient for assessment of intervention effectiveness, if there is no detection of pre-post outcome change then further piloting would be necessary to determine if different outcome measures are necessary or if the intervention's current design cannot produce preliminary outcome improvements. Feasibility results will ultimately inform what changes should be made for a future large-scale RCT of the intervention. The findings from this open trial also have the potential to provide useful information for other prevention programs that aim to improve implementation feasibility.

Methods

The current study was approved by the appropriate Institutional Review Board.

Sample

Eligible participants were undergraduate, graduate, and professional school students at UCLA. Inclusion criteria were: (a) age 18 or older, (b) currently enrolled at UCLA, and (c) enrolled in the prevention program. There were no pre-study exclusion criteria.

Procedure

Program recruitment. Recruitment for the prevention program was separate from recruitment of research participants (described below). Students were recruited as program participants through multiple strategies. Across recruitment strategies, the group-level delivery of the intervention was emphasized by using language like "we'll be in it together" or "a program for all students." Program advertisements were emailed (a) at the campus-level (i.e., department email announcements, a mass email sent from the registrar, student health newsletters) and (b) at the organization-level for any student groups or Greek organizations that agreed to forward our materials. We also composed eight different flyer designs that were posted (a) in-print throughout dormitories and central campus locations and (b) online through social media announcements by partner organization's pages. Finally, counseling center intake screeners were provided information about the intervention and enrollment procedures so that the staff could refer subclinical students waiting for initial counseling appointments. Regardless of recruitment source, students could directly enroll themselves in the program by visiting the sign-up website. To increase engagement during the recruitment process, the sign-up website (a) included a promotional video about the program and (b) offered students the choice to opt into one of two names of the program - "The Happiness Challenge" or "Reboot Camp" - based on their preference (see **Chapter 2** for details). There was no randomization of students into conditions, as the program was offered to all students at UCLA, regardless of their research participation. After signing-up online for the program, students received an automated program welcome email

and were subscribed to the email list. The open enrollment period lasted six weeks prior to the first week of the program, and 1,655 students enrolled in the program (3.8% of the campus student body).

Research recruitment. All students who signed up for the program were then sent an email four weeks and two weeks prior to the start of the intervention informing them of the opportunity to participate in the research study (as described below). Participation was voluntary and interested students could click the emailed link to read full study information and then provide their consent online to have their survey responses, student record information, and weekly online program activity used as data for research purposes. Of the students who enrolled in the prevention program online, 671 students (40.5% of program participants) consented to participate in the research study, of which 651 comprised the final sample. See Figure 3.1 for participant exclusion and research attrition.

Data collection. Research participants first completed a pre-program survey which was purposefully administered in the last weeks of an academic quarter, as the prevention program was scheduled to also end in the last weeks of the subsequent quarter. The intention was to minimize time-of-quarter confounds on any observed pre-post symptom changes as reported by students. Once the program started, all participants received weekly emails directing them to visit the website and read the weekly skills instructions. Program participants reported their own weekly use of the program by completing brief check-in surveys at the end of each week. These weekly check-in surveys were intended to be part of the program itself, but also served to collect self-reported program acceptability and adherence data. Research participants were sent an email with the post-program survey link after the eighth week of the program. They had two weeks to complete the survey and were sent a reminder email four days before the survey closed.

All symptom outcomes were assessed using self-report measures, completed entirely online via the pre- and post-survey. Given the online nature of the prevention program, this approach to data collection was deemed appropriate. However, both self-report measures and online surveys have a number of limitations. Self-report measures – as opposed to alternatives like behavioral, physiological, or observer-rated measures – are subject to respondent errors related to social desirability concerns, memory biases, and over/under-reporting of symptoms, among others (Fan et al., 2006; Schwarz, 2007; Tourangeau & Yan, 2007). Online-collected surveys – as opposed to alternatives like pen-and-pencil surveys or interviews – are subject to respondent errors related to inattentiveness or insufficient effort resulting in data for some subjects that is straight-lined, incomplete, or random (Cheung, Burns, Sinclair, & Sliter, 2016; Curran, 2016). Considering these concerns, it was important to plan on assessing the feasibility of this approach for detecting potential pre-post symptom change in the context of a population seeking out the current intervention.

We also preemptively shortened the online surveys in order to reduce survey burn-out. We did so by eliminating collection of self-report demographic data and by using trigger logic with the depression questionnaire (described below). Demographic data was instead collected directly from university student records, for which students provided their student ID numbers

We employed a number of strategies to encourage survey completion. Research participants were incentivized to complete the pre- and post- survey in exchange for entries into a gift card drawing. Program participants were encouraged to complete all weekly surveys regardless of module completion in order to earn a "Life Skills Training" certificate. We did not offer payment or study credit for survey completion as direct financial or academic incentivization risked artificially inflating adherence rates for universal prevention

implementation. Moreover, given the expected large enrollment it would have been cost-prohibitive to offer payment to all research participants.

Intervention

The intervention program had eight main modules, with two optional modules at the end. To select the skills content for these modules, we referred to a systematic review of practice elements in prevention programs for anxiety, depression, and stress in university students (**Chapter 1**) and selected practice elements based on those that were common to programs that were either (a) universal or (b) online/computer-based. See Appendix 3A for more details on how we selected content for each module and our procedure for developing the written content. Information on the modules is provided in Table 3.1. The main program ended after module eight, and program participants could continue with the optional modules after they completed their post-program survey.

Each module was released sequentially, to provide the program participants with updated and focused material week-by-week. The intervention also included dialogue support features associated with higher adherence. Reminders to continue practice were sent via email toward the end of every week. Program participants also had the option to download an electronic calendar with preloaded reminders that could sync to calendars associated with their email or smart phone. Each module presented an array of different options for practicing skills for the respective week, in order to allow flexibility and customization for each individual participant. In addition, every module had a section for tips, materials, and suggestions for how to practice each module's skill. The introduction to every module was written to specifically refer to student life. Aesthetic appeal was addressed by cohesive graphic design themes for "The Happiness Challenge" and "Reboot Camp" versions separately. Each version of the program had identical content, but used

different color schemes, logos, and graphics. Finally, program participants were incentivized by the chance to earn rewards at the end of every week. For every module that a participant logged their skills practice in the weekly check-in survey, they were entered into a drawing for a relevant prize (e.g., yoga mats for the physical exercise week; movie theater tickets for the communication skills week).

There were a number of ways group-level delivery was emphasized. First, a welcome email was sent reminding participants that all students were signed up "together" and encouraged them to invite friends to enroll. Then, the first week empathized the importance of social support and peer engaging for promoting skills learning. Throughout the program, all emails made references to each new module starting this week campus-wide. Finally, optional in-person events happening on campus were recommended in the weekly email if they related to the respective weekly module's skills (e.g., drop-in mediation workshops, free movie screenings).

Assessment of Feasibility

Recruitment source questionnaire. In one section of the entry survey, students were first asked "What referral sources did you encounter?" Listed options included all program recruitment strategies as well as potential other referral sources (e.g., friends) and a write-in "other" field. Students were also asked "What was important to your decision to sign-up?" Checkbox options included promotional video or program name choice.

Program adherence and completion rates. The weekly check-in surveys were used to operationalize participant's completion of each week's skills practice. In the check-in survey, program participants logged "reading", "attempting", or "completing" the weekly skills practice, as well as checking off which parts of the module they used (e.g., extra materials, scheduling instructions). They also wrote responses to three questions: (1) how they specifically applied the

skills for that week, (2) how the skills practice related to their overall program goal, and (3) if they plan to continue using the skill. All responses for all modules were reviewed for data cleaning. Responses were not counted towards program adherence if (a) the student only checked "reading" the module, (b) the response was evaluated as poor quality, or (c) there were duplicate responses submitted from the students for the same module. Quality of responses was coded by two separate research assistants who flagged inadequate responses, and discrepancies were resolved by the principal investigator. Poor quality responses were defined as those that demonstrated the participant had not actually practiced the skills (e.g., "I don't really know, I was so busy with midterms"), and such responses were not counted towards module adherence. Each participant was credited with the number of modules for which they had logged adherence. Although this assessment of program usage was self-reported, which is less optimal than a behavioral measure of adherence, the data was collected during the course of the program, rather than retrospectively at the end of the program, which should minimize memory biases in reporting.

From the module adherence counts, we defined various adherence and completion statuses to each participant. In terms of adherence, we categorized: "initiators" as those students who submitted a check-in; "partial completers" as those who logged activity for at least one module; "mostly adherent completers" as those who logged activity through the final week and missed two modules or less; and "fully adherent completers" as those who logged activity for all modules. Overall "program completers" for subsequent analyses were defined by the "mostly adherent completers" definition.

Program acceptability. Program acceptability was assessed through use rates of optional dialogue support features, satisfaction rates, and open-ended feedback. First, to assess use of

optional dialogue support, the weekly check-in surveys asked program participants to indicate if they had: (1) used the tips/suggestions for the respective module, and (2) used the extra electronic calendar reminders. Any student that had indicated using either feature at least once during the course of the intervention was credited with use of that type of dialogue support. Second, the post-program survey assessed satisfaction by asking the participant to agree or disagree with a checklist of descriptors for the program, including: "useful," "interesting," "relevant to me", "supportive," "too easy," and "too hard." Finally, the post-program survey asked for feedback through open-ended questions (What did you like about the program? Why did it work for you? What did you not like about the program? What could be improved?) that were qualitatively coded for themes. Responses to these questions were combined to serve as the unit of analysis for each student. See coding procedure details below in the Data Analysis.

Outcome assessment procedure sensitivity. Feasibility of the outcome assessment procedure that relied on self-reported measures completed online was assessed by determining if pre-post symptom changes could be detected regardless of data type (e.g., intent-to-treat analysis, completer analysis). There were three symptom outcome measures administered in both the pre-survey and post-survey: the Patient Health Questionnaire 9 (PHQ-9), the Spielberger State-Trait Anxiety Inventory – State scale (STAI-S), and the Perceived Stress Scale (PSS). The PHQ-9 and its two-item version (PHQ-2) were used to assess symptoms of depression within the past two weeks. The full PHQ-9 was auto-administered to a participant only if they answered "more than half of days" to either item on the PHQ-2 regarding depressed mood or anhedonia. Total scores range from 0 – 6 for the PHQ-2 and from 0 – 27 for the PHQ-9. The PHQ-9 has high internal consistency and good test-retest reliability (Kroenke et al., 2001), and the PHQ-2 has shown good sensitivity and specificity for detecting major depressive disorder (Löwe et al., 2005). The

STAI-S was used to assess symptoms of anxiety that are current/recent, as opposed to more trait-like ongoing anxiety. The STAI-S has 20 items and total scores range from 20 – 80. The STAI-S has shown good test-retest reliability and construct validity, and is commonly used with university students (Spielberger, 2010). The PSS was used to assess the degree of stress perceived in one's life. The PSS has 14 items, and total scores range from 0 – 56. PSS has shown good internal consistency and test-retest reliability and has been validated with university students (Cohen, Kamarck, & Mermelstein, 1983).

Data Analyses

Recruitment strategy yields. Descriptive analyses were used to examine rates of reported recruitment source and strategies.

Retention and adherence rates. Descriptive analyses were used to examine retention and adherence rates. Chi-squared analyses were used to compare program completers versus non-completers on research completion status (full sample, $n = 651$).

Program acceptability. Descriptive analyses were used to examine self-reported (a) dialogue support feature use rates and (b) satisfaction rates. Program acceptability was also examined using qualitative feedback theme frequency. Coding themes were not defined a priori. Potential themes were identified through an exploratory review of responses, with some common barrier themes in mind (e.g., accessibility). Final themes were selected through exploratory tallying of recurrent topics and keywords throughout responses. A qualitative coding manual was developed with definitions and examples for all coding themes. Qualitative coding was conducted by two undergraduate students who had served as research assistants for the study for one year. Training involved these two coders practicing on a subset of responses and then receiving feedback from the principal investigator about common discrepancies. The codebook

was revised to clarify exceptions to codes that were related to coder discrepancies during practice coding. All participant responses were then double coded, and the principal investigator resolved any discrepancies.

Outcome assessment procedure sensitivity.

Missing data. Data that was missing on the PHQ-2, STAI-S, or PSS at the item-level within a measure was addressed by calculating the participant's average item score and weighting it by total number of items in measure. For data missing at the measure-level, multiple imputation was used only for research participants that had completed the measure at one of the assessment times. See Appendix 3B for details of full procedure for multiple imputation.

Outlier identification. Linear regressions were run with pre-program measure scores predicting post-program measure scores. Scatter plots, student's deleted standardized residual, leverage, and other outlier analyses were examined. No outliers surpassed the guideline cut-off values, so none were excluded.

Repeated-measures MANOVAs. Two sets of four versions of RM-MANOVA were run predicting change from pre- to post-program scores on the STAI-S, PHQ-2, and PSS, with completer status as the between-group variable. The first set of RM-MANOVAs analyzed: (1) only cases with complete measures for pre- and post-program surveys, (2) the intent-to-treat (ITT) sample with last observation carried forward, (3) multiple imputation (MI) of missing post-program measures for research participants who completed pre-survey, and (4) MI of missing pre- or -post program measures for all research participants that had completed one or the other. A second set of identical RM-MANOVAs was run, this time excluding 24 students who reported receiving outside clinical services. The goal of running these multiple RM-ANOVA versions was to determine if changes could be detected from the self-report measures without sensitivity

to analysis type (i.e., completer, intent-to-treat, multiple imputation). Although any pre-post changes should not be interpreted as direct evidence of intervention effects such findings would provide preliminary feasibility to support further testing the intervention's effectiveness.

Results

Sample

Our sample of 651 was comprised of 71.2% undergraduate students and 28.8% graduate or professional school students, ages 18-61 (M = 22.9, SD = 5.5, median 21, mode 19). The sample was 80.3% female. Based on registrar-defined ethnicity, our sample was 30.5% Asian, 27.9% White, 18.8% Hispanic, 11.7% international, 5.7% multiethnic, 2.6% Black, 0.4% American Indian or Alaskan Native, and 2.6% unknown/other. The sample contained disproportionately more female and Asian students relative to the UCLA student population but was otherwise representative (**Chapter 2**).

Recruitment Response

The most commonly encountered recruitment sources were campus-level emails (i.e., department email announcements, email sent from the registrar, student health newsletter email), as reported by 82% of research participants. All other recruitment sources were less commonly encountered: printed flyers (5.8%), social media announcements (4.6%), organization-level emails (i.e., student groups, Greek organizations, 3.2%), and being referred by campus counseling center (0.6%). In addition to organized recruitment efforts, 7.2% reported being referred by a friend and 4.1% reported being referred by a coach, advisor, or supervisor. With regard to additional engagement strategies, 16% of cited the choice between the two programs as an important factor in signing-up. Of research participants, 59.9% enrolled in The Happiness Challenge" version, and 40.1% enrolled in "ReBoot Camp." The video promo was cited as an

important factor by only 2.7% of students.

Retention and Completion

There were 473 program initiators (72.7%), 408 partial completers (62.7%), 119 mostly adherent completers (18.3%), and 69 fully adherent completers (10.6%). Completion rates of skills modules were (as ordered by delivery): 31.0% for Make It Happen, 22.7% for Challenge Your Thinking, 22.9% for Life Troubleshooting, 18.7% for Decide How to Say It, 20.0% for Pause, 21.5% for Physical Exercise, 10.6% for Find Your Challenge Zone (Bonus Week), and 8.9% for De-Stress (Bonus Week). In terms of survey completion, 220 (32.7%) of research participants fully completed both surveys. Program completers were significantly more likely to complete the post-program survey, χ^2 (1, 651) = 200.03, $p < .001$, with 95.7% of 119 program completers providing responses compared with 25.9% of 532 program non-completers.

Program Acceptability

Use of Optional Dialogue Support. Of program initiators (at least one module completed, n = 408), 43.4% used the module tips and suggestions, and 32.1% opted to use the extra electronic calendar reminders. Examining program completers only (n = 119), 79.8% used the module tips and suggestions, and 74.8% opted to use the extra electronic calendar reminders.

Post-program feedback. Of the 250 research participants who completed the post-program survey (regardless of their completion status), 71.2% endorsed that the program was "useful," 64.8% endorsed that the program was "interesting," 65.2% % endorsed that the program was "relevant to me," with only 6% endorsing that the program was "too easy" and only 7% endorsing that the program was "too hard." However, less than half of students (43.6%) endorsed that the program was "supportive." Looking at completers only, 81.5% endorsed that the program was "useful," 74.8% endorsed that the program was "interesting," 76.5% endorsed

that the program was "relevant to me," with only 8.4% endorsing that the program was "too easy" and only 0.8% endorsing that the program was "too hard." Still only half of completers (51.3%) endorsed that the program was "supportive."

Qualitative feedback. In response to the open-ended questions, 239 students provided feedback. The qualitative coding procedure produced good to excellent inter-rater reliability for the identification of each theme ($\kappa = 0.64 - 0.91$). The most common themes ($n = 10$ or more) and detailed qualitative results are presented in a table in Appendix 3C.

The most common feedback related to many of the dialogue support features, which were mentioned primarily as positive feedback: "reminders" (33%), the "extras" section of each module (i.e., tips, resources, suggestions, materials; 23%), and "incentives/rewards" (5%). Many students also commented on "challenge difficulty" (22%), with responses ranging from "too easy" to "sufficient" to "too difficult". In terms of challenge "flexibility" or "customization," 9% found it to be sufficient, whereas 7% reported needing more. Similarly, although there was positive feedback for "reminders," many students requested either more or less frequent reminders or requested specific days/times. Finally, there were three common themes that related to desired program features that were missing: (a) more involved "goal-setting" (16%), (b) more "peer facilitation" within the program (11%), and (c) personalized progress tracking (9%). There was also a desire for the website to be "more interactive" (11%).

Many of the themes related to the accessibility afforded by the program's online modular design, which was mostly positive but highlighted some challenges. The "online delivery" method was mentioned in 15% of responses, either expressing satisfaction with the delivery via email and website or suggesting other online mediums (e.g., apps, videos). The program's "accessibility" or "self-guided" nature was identified in 12% of responses. In 11% of responses,

the “modular sequencing design” was either mentioned as a strength or facilitated actionable recommendations (e.g., order of modules). In contrast to students discussing the benefits of an online program’s accessibility, 28% of responses expressed the resulting difficulty with “self-motivation” or “accountability” (e.g., no accountability beyond one’s self; easy to forget the program when busy).

Assessment Procedures

All eight versions of the RM-MANOVAs produced significant univariate completer status x time effects for STAI-S (*p*s ranging < .001 to .03) and PHQ-2 (*p*s ranging < .001 to .02), but not for PSS (*p*s ranging .28 to .99). Thus, for the STAI-S and PHQ-2, results were reliable regardless of data type used. See Appendix 3D for details.

Of the eight versions, the RM-ANOVA that used multiply imputed data for pre- and post-program data and excluded outside-service users had the largest sample size (compared with complete case data or MI data for post-data only) and thus more power, although more clinically conservative (compared with ITT or including students with other service use). This RM-ANOVA produced a significant multivariate group by time interaction, $F(3, 584) = 5.84$, $p = .001$. There was no significant univariate group x time interaction for PSS scores. There were significant univariate group x time effects for STAI-S scores, $F(1, 586) = 9.51$, $p = .002$, with the program completers experiencing more decreases in STAI-S on average (partial eta squared = 0.02, power = .87). There were also significant univariate group x time effects for PHQ-2 scores, $F(1, 586) = 11.28$, $p = .001$, with the program completers experiencing more decreases in PHQ-2 on average (partial eta squared = 0.02, power = .92).

Discussion

Findings from this study suggest that the group-level online program for universal

prevention of anxiety and depression in university students is feasible in terms of recruitment strategies, retention, program acceptability, and outcome assessment procedures. In addition, our findings highlight strategies and components that are relatively more worthwhile, while also identifying areas that will need further improvement.

In terms of the recruitment, the overall large number of students enrolled in the program (n = 1,655) demonstrates the feasibility of this online intervention for large-scale intervention. Importantly, only 39% consented to the research study and were eligible, which is a reminder that many service users are not necessarily interested in participating in research. Looking at the recruitment yield from various strategies, 82% of research participants reported encountering campus-level emails, whereas other recruitment sources (i.e., printed flyers, social media announcements, organization-level emails, referral by campus counseling center) were far less commonly encountered. With regard to other engagement strategies, the video promo was cited as an important factor by only 2.7% of students, whereas 16% of students cited the choice to choose between the two programs as an important factor in signing-up. These findings suggest that mass emails are most worthwhile for recruitment efforts into such an online universal prevention program, and that offering branding choice is a self-identified factor influencing decision to enroll. That is not to say that other low-resource recruitment efforts should be abandoned entirely, as visibility on campus could still contribute to student interest (e.g., seeing flyers around campus prior to receiving a sign-up email). However, there seems to be diminishing return for recruitment efforts that are more time intensive (e.g., training campus counseling center staff in referral procedures; producing promo videos). In addition to organized recruitment efforts, 7.2% reported being referred by a friend and 4.1% reported being referred by a coach, advisor, or supervisor. This finding suggests that some recruitment occurs naturally

beyond planned strategies, and it could also speak to the importance of peer facilitation in the sign-up process for universal programming on campuses.

Regarding program adherence and completion, 72.7% of students at least initiated the program, 62.7% partially adhered (or better), and 10.6% fully adhered. The partial adherence rate is higher than similar programs for university students (Cavanagh et al., 2013; Musiat et al., 2014), and similar for full adherence, despite the fact that our prevention program was longer (Cavanagh et al., 2013). Given that the bonus skills modules (weeks 9 and 10) had more than 50% lower use than the final skills week, extending the program beyond eight weeks may not be useful for the online program in the future. These low intervention adherence rates should be considered within the context the study design. Open trial designs may provide a closer approximation of "real world" adherence and retention for online interventions given the naturalistic and uncontrolled aspects of these studies, whereas an RCT design has been associated with higher participant adherence rates in online interventions (Kelders et al., 2012). A review of real world adherence for digital self-help depression and anxiety interventions (Fleming et al., 2018) found that minimal use rates for interventions ranged from 21% – 88% (median: 61%) and moderate use rates (completing two to four modules) ranged from 7% to 39% (median: 11%). However, there is substantial variation in how online intervention studies define program adherence and completion (**Chapter 1**), making it difficult to compare relative adherence performance across studies. Thus, findings from the current open trial demonstrate intervention feasibility in terms of adherence and retention, though we hope making changes based on the program acceptability data can inform how to optimize the intervention and its adherence rates for future research and dissemination.

In terms of research completion, approximately one third of sample completed both

surveys, and program completers were significantly more likely to complete the post-program survey. Future research on online interventions should keep such drop-out in mind and aim to enroll far more students than the necessary final sample size for adequate power. These findings are also a reminder that research dropout is not the same thing as program dropout, and thus survey completion is not an appropriate proxy for intervention completion rates.

Program acceptability was demonstrated in terms of rate of use of optional dialogue support features, user-reported satisfaction, and qualitative feedback. There was good use of module tips and suggestions as well as electronic calendar reminders, especially by research participants who completed the program. The majority of users endorsed the program as "useful", "interesting", and "relevant to me." Some of the most common qualitative themes were positive feedback about the reminders, the "extras" section of each module (i.e., tips, resources, suggestions, materials), the incentives/rewards, the modular sequencing, and the online delivery. Thus, for some participants, the adherence-supporting features (e.g., reminders, sequential release of program content) may have been important for their ongoing participation in the program, especially given that this feedback was provided without specific prompts about these features.

The satisfaction survey and the qualitative feedback also helped identify some concerns and some program features that may need to be modified or added. The fact that a minority of research participants endorsed the program as "supportive" may reflect the fact that online delivery does not facilitate the same interpersonal support as traditional face-to-face interventions. (Alternatively, there is the possibility of measurement artifact and we may need to consider more valid measures of program satisfaction). Relatedly, a common theme in the qualitative feedback was about the lack of peer facilitation within the program. Further

development of the online program will need to consider how to increase participants' sense of being supported and connected. Solutions should focus on increasing a sense of community while protecting students from confidentiality concerns and potential iatrogenic peer effects. For example, an anonymous campus activity stream on the platform might facilitate a sense of community among student users without requiring administrator monitoring. A second improvement to consider based on the qualitative feedback is refining the goal-oriented aspects of the program (e.g., completing a goal plan at the beginning of the program that can be referenced throughout). A final feature commonly requested in the feedback was some kind of personalized progress tracking; as such online platforms might benefit from incorporating something like a dashboard that will allow users to review what they logged over time. For example, if users had individual accounts that displayed their weekly check-ins on some kind of dashboard, then they could review their progress over time.

Another important takeaway from the qualitative feedback was the need for flexibility and user customization. "Flexibility" and "customization" were discussed directly in 16% of student feedback responses. Many students commented on "challenge difficulty" (22%), with responses ranging from "too easy" to "sufficient" to "too difficult". Similarly, even with the positive feedback for "reminders," many students requested having either more frequent or less frequent reminders, or requested specific days/times. Future development of the online platform could provide students with multiple options for how many skills they want to practice in each module or how many reminders (and when) they want sent. This feedback supports the importance of the dialogue support feature of "descriptions seeming specific to participants." Allowing such customization might also address students desire, per qualitative feedback, for the website to be "more interactive." In summary, there is not an ideal one-size-fits solution to user

preferences, and therefore online interventions may be more feasible if they are designed to be responsive to individual needs.

Finally, an important trade-off of the benefits of online accessibility was highlighted. The program's "accessibility" or "self-guided" nature was identified in 12% of responses. In contrast to students discussing the benefits of an online program's accessibility, 20% of responses expressed the resulting difficulty with "self-motivation" or "accountability" (e.g., no accountability beyond one's self; easy to forget the program when busy). This feedback reveals that although online interventions may address some engagement barriers (e.g., accessibility issues), new barriers that are more unique to online interventions also need to be considered.

Regarding outcome assessment procedures, we were interested in determining if pre-post data collected through self-reported measures via online surveys would be sufficient for detecting symptom changes, regardless of analysis type (e.g., completer versus intent-to-treat). In this study, pre-post changes in measures of anxiety and depression showed significant decreases over time and for the program completers compared with the program non-completers. Although there was substantial survey drop-out for the post-program survey, the complete case, ITT, and MI analyses were all significant for change in PHQ-2 and STAI-S for program completers. Based on these findings, it appears that our assessment procedure was sufficient to detect change for anxiety and depression. Regarding stress, it must be further explored if the lack of symptom change is due to the intervention less effectively targeting stress or the PSS not validly detecting subjective stress level changes related to the intervention tested in this study.

The detected pre-post symptom changes suggest that the intervention is promising for promoting symptom improvement and the assessment procedure design is suitable for testing intervention effectiveness. Although these findings could be considered preliminary support for

the intervention's effectiveness, symptom improvement results could be explained by non-intervention causes, given the open trial design. For example, the passage of time often leads to elevated symptoms regressing towards the mean. Alternatively, we might interpret between-group findings with reverse directionality such that students experiencing symptom improvements were more likely to complete the program. Therefore, further testing of the online intervention with randomization to conditions is necessary.

Limitations

There are a number of study limitations to note when considering these feasibility findings. First, because our study was an open trial and did not include a control group or experimental manipulation, causality or directionality cannot be interpreted. Second, our acceptability data may be positively biased by program completers, as they were more likely to complete the post-program survey. Additionally, the lower response rate from non-completers means that some other qualitative themes may not have emerged, making it more difficult to adequately assess reasons for non-adherence. Third, our research sample may not be fully representative of all students using the intervention, and so results may not generalize to students who enrolled in the online program but not in the research study. Fourth, although outcome data from the self-report symptom measures did demonstrate pre-post change, the effect sizes were small, and thus it may still be difficult to detect symptom improvement in future research on such an online intervention. Finally, our sample was disproportionately female, which is consistent with studies finding service use to be higher in females (e.g, Eisenberg, Golberstein, & Gollust, 2007), but limits generalizability of these findings to non-female identifying students.

Limitations notwithstanding, the various implications discussed above can be used to directly guide the future development of the tested intervention. The lessons learned can also

inform the development of other prevention programs. Though we have established basic feasibility and acceptability for the program, we have yet to establish effectiveness, and thus a future RCT is warranted. It will also be important to determine for whom these types of programs work. Given that a minority of students fully adhered to the program, it is important to identify who is more likely to improve in response to the program. Coupled with the qualitative findings that indicated self-motivation to be a major theme for student participants, baseline motivational variables would be good candidates to test as moderators of intervention response. Finally, future research should further investigate nuances to how drop-out and adherence are defined, as rates may greatly vary by different criteria, and we do not yet know what the necessary "dose" of an online intervention is.

Conclusion

The current online prevention program for anxiety and depression in university students was novel and delivered at a group-level, and it was important to establish feasibility before further developing and researching such a program. Given the number of university students with subclinical and clinical anxiety and depression symptoms who do not use traditional mental health services, it is crucial that we develop other intervention options that will be appealing alternatives, such as online universal prevention program. The current study supports further experimental investigation of group-level online programming with engagement-facilitating features for universal prevention of anxiety and depression in university students.

Table 3.1

Descriptions of content in program modules

Module	Content
1. Getting Set Up	Explains how to set a program-related goal, and has student write and submit their goal. Introduces the practice of self-monitoring with recommendation materials (e.g., printable logs). Introduces the concept of peer engagement (e.g., talk to friends about learned skills) to reinforce skills
2. Make It Happen	Explains intentionally engaging in activities for the purpose of mood and stress management. Instructs participant to proactively schedule at least three pleasant activities, mastery activities, or value-drive activities for the current week
3. Challenge Your Thinking	Explains the relationship between thoughts and feelings. Describes three common types of unhelpful thinking habits. Encourages participants to identify such thoughts and reframe them to be less extreme throughout the week
4. Life Troubleshooting	Provides the option to either focus on time management or sleep hygiene. Outlines a list of suggested strategies for both skills sets. Instructs participant to pick three strategies that are most related to their specific difficulties and test them for the remainder of the week
5. Decide How to Say It	Explains the purpose of practice in order to improve communication skills. Provides strategies for pre-conversation preparation and for during conversations. Strategies address issues of assertiveness, social anxiety, and anger. Instructs participant to use skills in three conversations, starting with an easy conversation and ending with one that they expect to be more difficult
6. Pause	Explains the purpose of mindfulness as being moment-present, observant, and non-judgmental. Provides a list of mindfulness options with practice explanations. Instructs participant to practice mindfulness three times
7. Physical Exercise	Explains the benefits of exercise and how to overcome barriers to motivation. Provides practice options for increasing physical activity, trying new sports, or focusing on yoga/stretching. Instructs participant to exercise three times
8. Looking Ahead	Prompts participant to reflect on goal progress and on the most the skill(s) that they have found most helpful. Provides materials to plan for skills maintenance

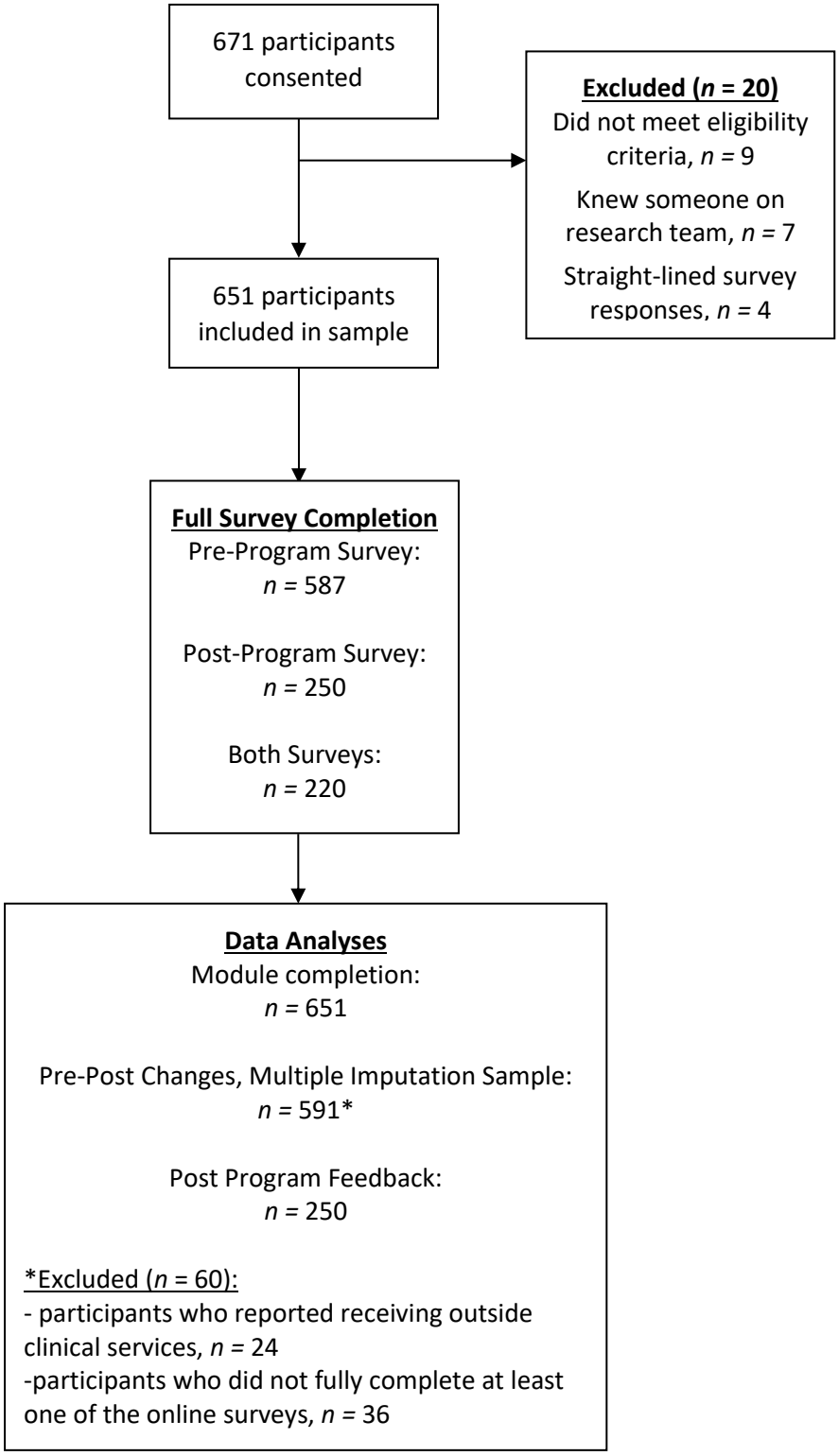
9. Find Your Challenge Zone (optional)	Explains the benefits of better understanding one's emotional, behavioral, cognitive, and physiological responses. Instructs participant to select one emotion and monitor their responses to three situations during the week
10. De-Stress (optional)	Explains the benefits of engaging in relaxation. Provides materials and suggestions for relaxation exercises. Instructs participant to practice relaxation three times

Modules 2 – 7 and 9 – 10 all contain psychoeducation about the benefits of the particular skills, information about relevant campus resources/services, a reminder to continue self-monitoring, and a reference to goal setting

---

Figure 3.1

CONSORT Flow Diagram



Chapter 4: Randomized Controlled Trial of an Online Program for the Universal Prevention of Anxiety and Depression in University Students

Abstract

Background: Although many online-delivered programs for the prevention of anxiety, depression, and stress in university students have proven efficacious, few are designed for universal-level prevention, and none are designed for synchronized group-level delivery. Given promising results from a non-randomized open trial testing such an online program, more experimentally designed studies are warranted to establish effectiveness. Moreover, it is important to understand for whom such an intervention is most effective. **Method:** This randomized controlled trial compared symptom changes for university students (n = 1607) in an 8-week online intervention condition versus a waitlist condition. Participants completed baseline, posttest, and 3-month follow-up assessments online. The primary outcome measure (Depression, Anxiety, & Stress Scale – 21; DASS-21) was administered at all time points and the moderator variable measures (internal motivation for treatment, grit) were administered at baseline. **Results:** Piecewise linear mixed effect models were run to examine group by time interactions and group by time by moderator interactions. The intervention group showed significantly more improvement than the waitlist in depression, anxiety, and stress outcomes at posttest, and the effects were maintained through 3-month follow-up. Subsequently, when waitlisted students accessed the online intervention their improvements in symptoms replicated those of the original intervention group. Effects were moderated by internal motivation – but not by grit – such that d effect sizes for the intervention condition relative to the waitlist condition were 0.40 to 0.50 for students with high motivation, 0.16 to 0.17 for students with average motivation, and -0.06 to 0.01 for students with low motivation. **Conclusion:** Implications, future directions, and limitations regarding effective implementation of such online interventions for university students are discussed.

Introduction

Increasingly, online programs are being developed for the prevention of depression, anxiety, and stress in university students (Davies, Morriss, & Glazebrook, 2014; Farrer et al., 2013). Compared with face-to-face options, online prevention programs offer a number of benefits for university students – accessibility (regarding time and location), privacy, fewer stigma concerns (Kauer et al., 2014) – many of which directly relate to common help-seeking barriers experienced by young people (Gulliver et al., 2010). Online programming also provides a more feasible delivery method for a universal-level prevention program offered to an entire campus. Given that anxiety and depression are the most commonly reported mental health concerns by college students in the United States (American College Health Association, 2018) and internationally (Auerbach et al., 2018), offering universal prevention programming for these types of symptoms is relevant a largest portion of the student population.

Dozens of prevention programs have demonstrated effectiveness for reducing anxiety, depression, and/or stress outcomes through randomized controlled trials (RCTs) with university students (e.g., Conley, Durlak, & Kirsch, 2015; Davies et al., 2014), but there are only a handful of programs offered through an online format and at universal level (e.g., 3 of 68 anxiety, depression, and/or stress prevention programs for university students as reviewed in **Chapter 1**). There is thus a need for additional research in this area. Moreover, no existing such universal prevention programs are designed for synchronized group-level delivery (**Chapter 3**). Group-level delivery offers a number of advantageous – for example, lowered barriers to service initiation through advertising the program to "all students" (**Chapter 2**), or ease of implementation for intervention administrators being able to deliver and monitor programming during designated times rather than ongoing delivery. An open trial tested group-level delivery

design of an online universal prevention program for university student anxiety, depression, and stress and found promising feasibility (**Chapter 3**), but now more rigorous experimental testing is warranted.

Despite studies generally supporting the effectiveness of online mental health programs for university students, there are also challenges that come with it: increased accessibility via online methods has come with problematically low rates of intervention adherence and retention. Online mental health programs typically have worse participant program completion rates than interventions that are delivered partially or entirely face-to-face (Clarke et al., 2014; Kelders et al., 2012; Van Ballegooijen et al., 2014). As systematically reviewed in **Chapter 1**, adherence has been measured in variable ways and is often low for RCTs of online anxiety, depression and stress prevention programs for university students. Consequently, any detected symptom improvement favoring online interventions for university students may be driven by those who actually remain engaged with the intervention. We are thus left with uncertainty when understanding the effects of such prevention-level online programs: does the intervention itself lead to improved symptoms, or do individual characteristics lead to help-seeking and resource use in general that in turn facilitate symptom improvement? In the latter case, online interventions might outperform control conditions "on average" because a subset of help-seeking and engaged students are provided with an accessible mental health service. Given the variability in individual use of online mental health by university students, it is important to understand who is more likely to have a positive response to such interventions. Thus, there is also a need for studies of such online interventions to test relevant moderators.

Given the self-guided nature of online interventions, motivation is likely an important moderator to consider. According to self-determination theory (Deci & Ryan, 1985), there is a

spectrum of motivation ranging from external to intrinsic: *external* (i.e., drive to reduce negative consequences/punishments or secure tangible rewards), *introjective* (i.e., to reduce internalized pressure, shame, guilt, etc.), *identifying* (i.e., drive towards goals and values that are identified as personally importance), *integrated* (i.e., drive to act in congruence with values that are core to one's sense of self); and *intrinsic* (i.e., drive by interest, curiosity, satisfaction, and enjoyment). In the application of self-determination to mental health interventions, research evidence suggests that more self-motivation (i.e., less external) predicts greater treatment adherence and progress (Ryan & Deci, 2000; Ryan & Deci, 2008). A study that tested positive psychology self-help interventions found that students with more intrinsic motivation experienced greater improvement in their outcomes post-program (Lyubomirsky, Dickerhoof, Boehm, & Sheldon, 2011). The study authors accordingly argued that in order for self-help interventions to be maximally effective, users need to both (a) be receiving empirically supported intervention practices and (b) have their own motivation to use the intervention. Indeed, according to student feedback from the open trial study in **Chapter 3**, qualitative findings showed that motivational difficulty was one of the most common self-reported barriers to using the online intervention.

Unfortunately, research that examines motivation as a moderator of online intervention response is scarce with university student populations and has mixed findings for other adult populations. In a study of online depression prevention programming for university students it was found that self-identified need for help – a related index of motivation – predicted program adherence, which subsequently predicted improvement (Lintvedt et al., 2013). In a study of an online relaxation intervention for adults, there was no relationship between baseline internal motivation for treatment and post-intervention reduction of stress symptoms; however, high baseline external motivation (e.g., feeling pressured by others to get help) did predict worsened

stress symptoms at post-intervention (Alfonsson, Olsson, & Hursti, 2016). Paradoxically, a study of an online CBT intervention for adults also found that "high motivation" for treatment predicted relatively less depression improvement (Farrer et al., 2014). However, in this study the treatment motivation measure (i.e., Nijmegen Motivation List for Prevention Scale; Allart-Van Dam, Hosman, & Keijsers, 2004) assessed both internal and external reasons for change. Further investigation of internal motivation as a moderator for symptom improvement in online mental health programs for university students is warranted.

Although baseline motivation for treatment may be important, motivation alone may be insufficient without the actual ability to persevere towards goals. Indeed, studies have found that online mental health intervention adherence and retention are predicted by one's ability to adhere to a program, as assessed by measures of: internal locus of control (Geraghty, Wood, & Hyland, 2010); behavioral control specific to expected ability to stick to program (Hebert, Vincent, Lewycky, & Walsh, 2010; Wojtowicz, Day, & McGrath, 2013); and self-report of being in "preparation" or "action" phase of stages of change (Al-Asadi, Klein, & Meyer, 2014). A related construct worth considering is grit, which is one's individual capacity for long-term goal attainment and perseverance. Accordingly, we might expect having more grit to help individuals persist through online interventions and thus benefit from them. Grit has already been predictive of retention in other contexts such as military training, workplace employment continuance, and high school graduation (Eskreis-Winkler, Shulman, Beal, & Duckworth, 2014). Although grit has been increasingly examined in psychology studies, it has not yet been tested as a moderator of intervention response. Given the low adherence rates for online interventions, testing moderation of intervention response by a construct such as grit seems worthwhile.

The Current Study

The current study proposes to examine the effectiveness of an original online anxiety and depression prevention program intended for universal delivery to university students using an RCT design. The program was designed with eight modules to be delivered with sequential weekly access of content and use of more dialogue support strategies, as such features have been recommended to improve adherence (Kelders et al., 2012). We are interested in symptom change, and moderators of symptom change, in the active intervention group as compared with a waitlist group. These results will help us better understand the relative benefit of online mental health interventions on university campuses, as well as individual differences (i.e., who benefits more from what).

Research questions. First, we aimed to evaluate the effectiveness of this new group-level online prevention program. The use of a waitlist condition allowed us to assess experimentally if the intervention condition is responsible for change in anxiety, depression, and stress symptoms over time. Given the previously established effectiveness of online skills-based mental health programs (e.g., **Chapter 1**), we expected students in the intervention condition would experience significantly more improvement in symptoms compared with those in the waitlist group. Second, to strengthen assessment of effectiveness, we also examined: (a) maintenance of any intervention effects through follow-up, and (b) replication of intervention effects by the waitlist once delayed access to the intervention was granted. Third, we aimed to explore how individual characteristics, specifically (a) internal motivation for treatment and (b) grit, may moderate symptom change: does symptom change improve because of condition, because of individual characteristics, or a combination of both? Finally, we aimed to evaluate program acceptability in terms of satisfaction and participation rates.

Methods

The current study was approved by the appropriate Institutional Review Board.

Participants

We recruited undergraduate, graduate, and professional students, who were at least 18 years old, from a large public university in southern California. Recruitment materials were distributed via a mass email to all enrolled students, printed flyers, social media, announcements over emails lists (e.g., academic departments), and announcements in psychology courses requiring research study participation credits. The only a priori exclusion criteria after baseline survey was concurrent enrollment in a similar online anxiety and depression treatment study, given the overlapping therapeutic content. Students were offered two forms of compensation for survey completion, based on their preference: entry into prize drawing for gift cards (one $100 gift card and ten $10 gift cards) or research study course credits. Post hoc exclusion procedures removed students with invalid data reporting (e.g., straight-lined responses to surveys, high inconsistency in responses to cross-validation item pairs selected based on content similarity). See details of participant inclusion, withdrawal, and exclusion flow in Figure 4.1.

Power analyses. Although we planned to use mixed linear effect (MLE) models, a priori power analyses for complex MLEs are less established. Thus, we ran power analyses for repeated-measures ANOVA, a similar but more conservative analysis. G*Power software was used to calculate the necessary sample size to adequately power a repeated-measure ANOVA with two assessment periods and four groups (2 conditions x 2-level moderators). Based on results from an open trial of the program which compared program completers to non-completers (**Chapter 3**), parameters were set at: effect size, $F = 0.10$; correlation among repeated measures $r = 0.52$; error probability = 0.05; power = 0.95; and nonsphericity correction = 1. The necessary sample size estimated was 420 complete cases. Given that drop out from online programs can be

quite high, we aimed to enroll 800 students, with at least 400 assigned to each intervention condition.

Design and Procedure

Prior to beginning any research procedures, students provided informed consent online and then completed the baseline survey. In order to complete the baseline survey, students had to provide their student ID number, which (a) ensured only registered students could enroll and (b) allowed us to link demographic data from student records. The baseline survey was open for three weeks, during which participants were notified of their randomly assigned condition via email within 24-72 hours of completing the survey. Randomization to either the intervention group or the waitlist group was conducted using a random number generator. Students assigned to the intervention group also received a verification code in the email, allowing them to access the online platform and set-up their account. The intervention group was then active for eight weeks, after which the posttest survey was open to the intervention group and waitlist group for two weeks. There was then a down period of two weeks during winter break, after which students assigned to waitlist group gained access to the online intervention, if they completed the posttest survey. There was then another 8-week period during which the original waitlist group completed the intervention while the original intervention group was in a maintenance phase. After this period, all participants who completed the posttest survey were invited to complete the 3-month follow-up survey, which was open again for two weeks.

For the duration of the study, participants in either group were allowed to access other on- or off-campus mental health services and resources, and we assessed for such use in each survey. This design was intentional in order to maximize generalizability of findings to real-world campus environments in which students can access services as usual.

Although this study and intervention was intended to be a prevention study, because all students were invited to participate regardless of symptom level, we were aware of the possibility that some enrolling students might need higher levels of care. To this end, we included the Patient Health Questionnaire-9 measure at each survey assessment in order to assess for suicidal ideation. A safety follow-up protocol was enacted for any student that endorsed suicidal ideation. This protocol included outreach by phone or email, further assessment of risk, and referral and case management to connect the student with mental health services as needed.

Intervention group. An earlier version of this online intervention had been tested in an open trial (**Chapter 3**) and common themes from participant feedback informed adaptations to the current intervention (see Appendix 4A for informed adaptations). When participants first accessed the online platform they were guided through the account set-up process. First, the user selected either brand of the program – "The Happiness Challenge" or "ReBoot Camp." Each version of the program has identical content, but uses different color schemes, logos, and graphics, and thus allows users to be involved in their own intervention decision making by selecting the brand they prefer (see **Chapter 2** for details). Second, the user customized their settings for timing and frequency of weekly reminder emails. Finally, the user engaged in goal-setting, which included selection of one of seven goal areas (e.g., mood management, anxiety reduction) related to intervention progress, motivational enhancement strategies (e.g., writing reasons this goal was important to them), and identification of a person as their go-to support.

The intervention program was eight weeks long, including an introduction week module, six skills modules, and a maintenance planning module. Details of module contents are presented in Table 4.1. Each module was released sequentially, to provide the participants with updated and focused material week-by-week. The introduction to every module was written to

specifically refer to student life. Each module presented a checklist of activities for participants to complete as well as tips and suggestions for how to practice each module's skill. At the beginning of the week, participants made a "plan" indicating which activities/skills they intended to try and any barriers they expected to encounter.

Throughout the week, participants could log any time they completed an activity along with ratings of their mood and stress level, which was charted in their dashboard graph. Activities could be logged within a seven-day grace period. At the end of each module participants were prompted to complete a weekly check-in which asked participants to reflect on what they liked about this module and if it moved them toward their program goal. Weekly prize drawings (items of $10-40 value) were offered to all participants submitting a weekly check-in. Students who had logged activity for all eight modules were eligible for the completer prize drawing (items of $100-400 value). There was also a "Campus" section of their dashboard that provided students with info about relevant campus resources, a notification center, and an anonymous livestream of all campus users' activity, to provide a sense of community.

Waitlist group. Participants in this group received no intervention or communications between their baseline randomization and the invitation for posttest survey. They were provided access to the online program after completing the posttest survey.

Measures

Primary outcome measure. Our primary symptom outcome measure was the 21-item version of the Depression Anxiety and Stress Scale (DASS), which assesses self-reported symptoms related to depression, anxiety, and stress. The DASS-21 has demonstrated high internal consistency (.83 - .90) and good construct validity in university student samples (Norton,

2007). Internal consistency of the DASS-21 subscales using our sample at baseline was good (depression: α = .89; anxiety: α = .79; stress: α = .82).

Moderator variables.

Internal motivation. The Treatment Motivation Questionnaire (Ryan, Plant, & O'Malley, 1995) is a measure that assesses participants' reasons for initiating treatment and their expectations for completing the program. The TMQ has two motivation scales: (a) internal motivation (e.g., "I really want to make some changes in my life.") and (b) external motivation (e.g., "I came to treatment now because I was under pressure to come."). The TMQ has significantly predicted intervention completion in other research studies, for example in-person alcohol treatment (Ryan et al., 1995) and online stress treatment (Alfonsson et al., 2016). For this study, questions were minimally adapted to apply to an online mental health promotion program instead of an in-person treatment (see Appendix 4B for adapted measure). Questions from the help-seeking subscale were removed, as they relate directly to expectations about interacting with other treatment participants, which was not applicable to our study design. The TMQ was included in the baseline survey only. For the purposes of this study, we focused on the internal motivation subscale. Internal consistency for our adapted TMQ internal motivation measure was good (α = .87), and dropping none of the 11 items would have produced a larger alpha.

Grit. The Short Grit Scale is a 12-item self-report measure assesses trait-level perseverance and motivation for long-term goals (Duckworth & Quinn, 2009). Items are rated on a 5-point scale, representing "not at all like me" to "very much like me" in response to statements (e.g., "Setbacks don't discourage me", "I finish whatever I begin"). Grit-S has produced good internal consistency (0.73 - 0.83), test-retest reliability (r = 0.68 one year apart), and predictive validity with other measures of successful goal attainment (Duckworth & Quinn,

2009). The Grit-S was included in the baseline survey only. Internal consistency of the Grit-S measure using our sample was good ($\alpha = .82$).

Other self-report measures.

Demographic information. Demographic data was based on self-report according to student records, and included: gender, ethnicity/race, age, and student status (i.e., undergraduate vs. graduate/professional). Student record ethnicity/race data was based on self-reported identify per ethnic categories defined by the Common Data Set (CDS Advisory Board, 2018): nonresident alien; Hispanic/Latino; Black or African American; White; American Indian or Alaskan Native; Asian; Native Hawaiian or other Pacific Islander; two or more races; and unknown. Gender was defined as binary because the student record system only allows students to select "Male or "Female" for reporting purposes.

***Self-reported resource and service use*.** Students were asked to indicate past and current use of health-related services on- and off-campus using a checklist of common resources/services as well as a write-in "other" option. These questions were included in the all three surveys.

Patient Health Questionnaire-9 (PHQ-9). The PHQ-9 assesses symptoms of depression during the previous two weeks. The PHQ-9 has demonstrated high internal consistency (.86 - .89) and criteria validity by predicting likelihood of diagnosis by a mental health professional (Kroenke et al., 2001). Each item is rated on a 0-3 Likert scale ("Not at all" to "Nearly Every Day"). We used PHQ-9 as a screening measure to assess suicidal ideation and overall depression severity. Students endorsing anything other than "not at all" on the ninth item about suicidal ideation were flagged for follow-up protocol. Given that students contacted for the safety follow-up protocol were functionally receiving addition intervention regardless of their assigned condition, we recorded which students did versus did not endorse suicidal ideation at baseline.

Students were included in the full study regardless of suicidal ideation as a universal prevention program is intended for all students in a population regardless of symptom level.

Measures of intervention acceptability.

Program satisfaction. The posttest and follow-up surveys included questions about experiences with the program for the respective group that just completed the intervention. The questions were adapted from the Client Satisfaction Questionnaire (Larsen, Attkisson, Hargreaves, & Nguyen, 1979), which has been found to be an appropriate measure of satisfaction with high internal consistency and concurrent validity across a broad range of intervention contexts (e.g., adult residential treatment setting, Kelly et al., 2018; adult outpatient setting, De Wilde & Hendriks, 2005; children's outpatient mental health services, Copeland, Koeske, & Greeno, 2004). See Appendix 4C for questionnaire.

Program participation. Behavioral data from participants' accounts was collected from their online program account. Users with at least one logged activity for a given module were categorized as having participated in that module.

Data Analysis Plan

Psychometric assessment of primary outcome measure. To assess the psychometric properties of the DASS-21 in our sample, we examined descriptive data and a factor analyses of DASS-21 scores at baseline. Although the DASS-21 has been validated across ethnic/racial groups of university students (e.g., Norton, 2007), some studies have found inconsistencies in the factors when using the DASS with individuals from certain cultures (e.g., Camacho, Cordero, & Perkins, 2016; Oei, Sawang, Goh, & Mukhtar, 2013). Therefore, factor analysis of the DASS-21 in this study seemed warranted, given that it was administered in English only with a sample from a diverse student body, including many international students and first-generation

American students. Exploratory factor analysis was conducted through principal axis factoring with direct Oblimin and promax rotations to identify factors with an Eigenvalue ≥ 1, using a factor loading cut-off of .4 (Stevens, 1992).

Analysis and software selection. To assess intervention effects and moderation effects, linear mixed effects (LME) models were run in R (R Core Team, 2013) using the multilevel package (Bliese, 2016). By using a LME model approach, missing data were estimated using restricted maximum likelihood estimation and therefore all participants (n = 1607) could be included in the models. This specific imputation approach is favorable over complete case analysis as it does not reduce power due to decreased sample size nor introduce bias if research drop-out is related to treatment non-response. This approach is also favorable over simpler intent-to-treat approaches (e.g., last observation moved forward) that are based on the unlikely assumption that missing data are completely at random. R^2 equivalent statistic was calculated for all mixed models using the r2glmm package (Jaeger, 2017). Simple slopes of interactions were calculated using the reghelper package (Hughes, 2018). All other analyses (i.e., factor analyses, t-tests, one-way ANOVAs, binary logistic regressions) were run in SPSS version 24 (IBM Corp, 2016).

Assessment of randomness in group assignment and drop-out. Differences by condition or drop-out status were assessed through independent t-tests or Chi-squared analyses, depending on if variables were continuous (e.g., DASS-21 scores) or nomological (e.g., gender, ethnicity, SI status). For t-tests, if Levene's test was significant, then we reported the test statistic for which equal variance was not assumed.

Covariate selection. We identified covariates for models by running independent t-tests or ANOVAs to test for between-group differences in baseline DASS-21 scores by demographic

variables and by baseline suicidal ideation (see justification in methods above). Any demographic variables that related to DASS-21 outcomes at baseline were included as covariates in the subsequent mixed effect models. For further rationale for this approach and to see the unadjusted model results, see Appendix 4D.

Intervention effects. The basic model of intervention effects included group, time, group by time interaction, covariates, and any covariate's respective group by time interaction as fixed effects as predicting DASS-21 subscale scores. Participant intercepts and outcome slopes were allowed to vary randomly. Separate models were run predicting depression, anxiety, and stress as outcomes. Models were run as piecewise LME models in order to treat time effect slopes separately for (a) baseline to posttest survey (i.e., intervention versus waitlist response), versus (b) posttest to 3-month follow-up survey (i.e., intervention group in maintenance phase; waitlist group had delayed access to intervention). To do so, dummy coding was using to create a Time A variable (baseline = 0, posttest = 1, follow-up = 1) and a Time B variable (baseline = 0, posttest = 0, follow-up = 1), ensuring that the resulting test statistics represented only the slope for its respective time phase. All models treated the intervention condition as the reference group. Assessment of intervention effects was done by examining group by time interactions with the Time A variable followed by examining simple slopes of time.

Intervention effect maintenance. Using the same piecewise models described above, assessment of intervention effect maintenance was done by probing group by time interactions for the Time B variable (i.e., posttest to follow-up). In the presence of significant interaction, effect maintenance was determined if simple slopes of time for the intervention group (i.e., during the 3-month maintenance phase) were non-significant.

Replication of intervention effects. To access if the waitlist group replicated intervention effects once granted access to the online intervention, we collapsed data provided by the immediate intervention group from baseline to posttest with data provided by the waitlist group from post-waitlist (i.e., posttest) to post-delayed-intervention (i.e., follow-up). LME effects models were then run including group, time, group by time interaction, covariates, and any covariate's respective group by time interaction as fixed effects predicting DASS-21 subscale scores. In this case, a non-significant group by time interaction was assessed as indicating comparable magnitude of intervention effects between the immediate intervention group and the delayed intervention group.

Moderation effects. Each moderator of pre-post effects (TMQ, GRIT) was tested in its own model. Moderation models included the moderator by group by time interaction, all respective lower order interactions and main effects, and the same covariates as the basic model. Moderator models were also run separately for predicting depression, anxiety, and stress as outcomes.

Robustness checks and Type I error corrections. To assess robustness of results, if the respective intervention or moderator model produced a significant interaction, then the model was run using four different sample versions. The main version (n = 1607) included the full final sample. The second version (n = 1367) excluded students who reported receiving other mental health services at baseline. The third version (n = 661) excluded students who reported receiving other mental health services at posttest. The fourth version was an initiator sample (n = 947) and included only students from both groups who had at least initiated the intervention. Results are reported for the main version sample, with results robustness check versions provided as a range within parentheses for every model run with robustness checks.

Because our analyses tested intervention and moderation effects for multiple outcomes, it was important to address Type I error across findings. We controlled the false discovery rate for main models by applying Benjamini and Hochberg's (1995) approach to calculating critical *p*-values for the respective interaction term.

Effect sizes. Effect sizes were calculated based on estimated marginal means using the emmeans package for R (Lenth, Singmann, Love, Buerkner, & Herve, 2018). All within-group effect sizes were calculated using this formula:

$$d = \frac{M_1 - M_2}{\sqrt{\frac{(n_1 - 1){SD_1}^2 + (n_2 - 1){SD_2}^2}{n_1 + n_2 - 2}}}$$

Between-group effect sizes were then calculated by subtracting the waitlist condition's *d* from the intervention condition's *d.* For ease of interpretation of effect sizes for any significant moderator, models were run again treating the moderator as a categorical rather than continuous variable. To transform the moderator into a categorical variable, scores that were below or equal to one standard deviation below the mean were coded as "low", scores above or equal to one standard deviation above the mean were coded as "high", and all scores in between were coded as "average."

Program acceptability. Finally, satisfaction and program use rates were examined to assess program acceptability.

Results

Participants. Our sample of 1607 was comprised of 63.0% undergraduate students and 37.0% graduate or professional school students, with a mean age of 22.8 (*SD* = 5.4; median = 21). The sample was 74.2% female. Based on registrar-defined ethnicity, our sample was 31.8%

White, 26.4% Asian, 19.3% Hispanic, 12.6% international, 5.0% multiethnic, 1.9% Black, and 2.9% other/unknown. See Table 4.2 for demographic details of students by condition.

Preliminary Analyses

Psychometric properties of DASS-21. At baseline, DASS-21 scores for depression (M = 5.43, SD = 4.60) and anxiety (M = 4.68, SD = 3.77) were in the mild range and scores for stress (M = 7.28, SD = 4.21) were in the normal range. All DASS-21 subscale scores were comparable to means found for university students (Norton, 2007). Factor analysis with either rotation produced the same three factors: factor 1 (eigenvalue = 8.13) with all seven depression subscale items loading into it; factor 2 (eigenvalue = 1.87) with six anxiety subscale items and one stress subscale item loading into it; and factor 3 (eigenvalue = 1.17) with three stress items loading into it. DASS-21 items 2, 6, 14, and 18 did not have factor loading of ≥ .4 into any factor. Given these results, the DASS-21 stress scale in particular may not have been as reliably measuring stress in our sample as it has in prior research.

Assessment of differences by condition. There were no differences in DASS-21 scores at baseline between conditions (p = .3 to .9). There were also no differences between conditions for all demographic variables (p = .3 to .6), baseline SI status (p = .3), resource use at baseline (p = .9), resource use by posttest (p = .9), nor moderator variables at baseline (treatment motivation, p = .6; GRIT, p = .3).

Covariate selection. There were significant between-group differences by gender for baseline anxiety (t = 2.62, p = .009) and baseline stress (t = 3.97, p < .001), but not baseline depression (t = -.52, p = .6). Gender was thus included as covariate in all LME models to maintain equivalency across models predicting depression, anxiety, and stress. There were no significant differences by ethnicity for depression (F = 1.38, p = .2), anxiety (F = 1.16, p = .3),

or stress (F = 2.11, p = .03; post hoc Bonferroni tests, p = .13 to .99), therefore ethnicity was not selected as a covariate. No significant differences by international status (p .10 to .65) were found, therefore it was not selected as a covariate. Finally, there were significant between-group baseline differences by binary suicidal ideation status for depression (t = 18.12, p < .001), anxiety (t = 4.80, p <.001), and stress (t = 11.30, p < .001), justifying its inclusion as a covariate in all LME models. The inclusion of the SI variable functionally allowed us to control for any differences due to unique severity and/or additional intervention offered to students endorsing SI.

Assessment of differences by drop-out status. There were no significant differences on baseline DASS-21 scores (depression, p = .2; anxiety, p = .5; stress, p = .6) or moderator variables (treatment motivation, p = .8; Grit, p = .6) between participants who completed the posttest survey versus those who had dropped out. There were more research drop-outs by the gender covariate (χ^2 = 15.14, p < .001; 35.5% of female students vs. 46.3% of male students), but not by the baseline SI covariate (p = .9). There were also more research drop-outs from the intervention group (47.1%) than from the waitlist group (38.5%) such that waitlist participants were disproportionately retained at posttest, χ^2 (1, 1607) = 12.30, p < .001. Examining participants in each condition separately, there were no differences in baseline DASS-21 scores or moderator variables between posttest survey completers and drop-outs in the intervention group (p = .3 to .7) nor in the waitlist group (p = .2 to .8).

Assessment of Intervention Effects

Immediate intervention effects. For depression, R^2 = .15 (robustness checks: R^2 .11 to .14), there was a significant group by time interaction, t = -3.05, p = .002 (robustness checks: t = -1.98 to -3.51; p = < .001 to .048). Greater symptom reduction was indicated for those in intervention condition, d = 0.23, compared with the waitlist group, d = 0.05, with a simple slope

of time being significant for the intervention group (t = -2.06, p = .04), but not for the waitlist group (p = .6). For anxiety, R^2 = .08 (robustness checks: R^2 = .06 to .07) there was a significant group by time interaction, t = -3.01, p = .003 (robustness checks: t = -2.42 to -3.60; p = < .001 to .02). Greater symptom reduction was indicated for those in intervention condition, d = 0.26, compared with the waitlist group, d = 0.08, with a simple slope of time being significant for the intervention group (t = -3.34, p < .001), but not for the waitlist group (p = .4). For stress, R^2 = 0.07 (robustness checks: R^2 .05 to .07) there was a significant group by time interaction, t = -2.92, p = .004 (robustness checks: t = -.98 to -3.29; p = .001 to .3). Greater symptom reduction was indicated for those in intervention condition, d = 0.16, compared with the waitlist group, d = -0.01, with a simple slope of time being significant for the intervention group (t = -2.35, p = .02), but not for the waitlist group (p = .9). Overall, the intervention showed more reduction in depression, anxiety, and stress from baseline to posttest, with group by time interaction terms' significance tests surpassing the calculated Benjamini-Hochberg critical values (p < .02, .03, and .05, respectively).

Maintenance of intervention effects. We next examined if immediate intervention effects were durable from post-treatment to follow-up. For depression, there was a significant group by time interaction, t = 3.48, p < .001 (robustness checks: t = 2.20 to 3.27; p = < .001 to .03), with students in the maintenance phase demonstrating no significant change in symptoms, d = -0.01 (simple slope of time: p = .3). For anxiety, there was a significant group by time interaction, t = 3.24, p = .001 (robustness checks: t = 1.96 to 3.20; p = .001 to .050), with students in the maintenance phase demonstrating no significant change in symptoms, d = 0.11 (simple slope of time: p = .5). For stress, there was a significant group by time interaction, t = 2.69, p = .007 (robustness checks: t = 1.62 to 3.07; p = .002 to .11), with students in the

maintenance phase demonstrating no significant change in symptoms, d = 0.14 (simple slope of time: p = .9).

See Figure 4.2 for symptom change by condition from baseline to pre-test to follow-up. See Appendix 4E for full robustness model results. See Appendix 4F for comparison of effect sizes calculated from raw means versus from estimated marginal means.

Replication of intervention effects. With data collapsed across the immediate intervention group and the delayed intervention group (n = 1,291), group by time interactions were examined. There was no difference in intervention effects between the immediate intervention group and the delayed intervention group for depression (p = .26) or anxiety (p = .11). For stress, there was a significant group by time effect, t = 2.70, p = .007, with the simple slope of time being steeper for the delayed intervention group (t = - 2.99, p = .002) relative to the immediate intervention group (t = -0.74, p = .46).Therefore, intervention effects for the delayed intervention group were comparable in magnitude for depression and anxiety, and greater in magnitude for stress, relative to the immediate intervention group.

Based on the original LME models (n = 1607), outcome means of intervention group and the waitlist group did not differ by follow-up assessment, with the 95% confidence intervals for the between-group effect sizes containing zero for depression, d = 0.14 [-0.04, 0.31], anxiety, d = 0.12 [0.00, 0.24], and stress, d = 0.07 [-0.13, 0.27]. Thus, after delayed intervention access, the waitlist group caught up to – but did not surpass – the intervention group during its maintenance phase.

Moderation Effects

Treatment Motivation. Prior to testing three-way interaction, we first ran models with the internal motivation variable and its respective two-way interactions added. There were

significant interactions of motivation by time for depression, t = -3.36, p < .001, anxiety, t = -3.72, p < .001, and stress, t = -3.36, p < .001.

We examined moderation of intervention by time effects from baseline to posttest. For depression, R^2 = .20 (robustness checks: R^2 = .17 to .20), there was a significant group by time by motivation interaction, t = -2.67, p = .008 (robustness checks: t = -2.27 to -2.75; p = .006 to .02), with simple slopes of time being significant only for the intervention group at average motivation levels (t = -2.28, p = .02) and at high (+1 SD) motivation levels (t = -4.05, p < .001). For anxiety, R^2 = .12 (robustness checks: R^2 = .11 to .13) there was a significant group by time by motivation interaction, t = -2.69, p = .007 (robustness checks: t = -2.76 to -2.3.28; p = .001 to .006), with simple slopes of time being significant only for the intervention group at average motivation levels (t = -3.66, p < .001) and at high (+1 SD) motivation levels (t = -5.33, p < .001). For stress, R^2 = .12, there was not a significant group by time by motivation interaction, t = -1.78, p = .07.[2] Thus, moderation of intervention effect was indicated for depression and anxiety, with the interaction terms' significance tests surpassing the calculated Benjamini-Hochberg critical values (p < .03 and .02, respectively), whereas the moderation of intervention effect for stress did not meet conventional standards for significance (i.e., $p \leq .05$). The two-way motivation by time interaction terms became insignificant (ps .2 to .4) in the presence of the three-way interactions predicting change in depression and anxiety.

See Table 4.3 for estimated marginal means and effect sizes by group by time by motivation. To calculate estimated marginal means at different motivation levels, we ran the model treating motivation as a categorical variable. Results remained the same for the depression

[2] Despite this interaction term being non-significant, simple slopes of time were significant for students in the intervention group with average motivation (t = -3.99, p < .001) and high motivation (t = -4.00, p < .001).

model (R^2 = .19; group by time by motivation interaction: t = -2.55, p = .01), anxiety model (R^2 = .12; group by time by motivation interaction: t = -2.65, p = .008), and stress model (R^2 = .14; non-significant group by time by motivation interaction: t = -1.81, p = .07).

Given that visual examination of marginal means in Table 4.3 suggested the possibility that the motivation variable was confounded by baseline severity, we ran a post hoc analysis to rule out baseline symptom severity. We ran another set of LME models including pre-intervention DASS scores as predictors in each respective model. We were interested in examining the group by motivation interaction, as the Time A predictor was not included. Because baseline DASS scores were no longer included as part of the repeated measures outcome, fewer cases could be imputed for missing scores and thus the sample size was smaller (n = 992). We found that motivation still moderated the group effect, over and above baseline symptom differences for the depression model (R^2 = .24; group by motivation interaction: t = 2.50, p = .01) and the anxiety model (R^2 = .24; group by motivation interaction: t = 3.34, p < .001), and now there was also moderation within the stress model (R^2 = .23; group by motivation interaction: t = 2.50, p = .01). The significance tests for these interaction terms surpassed the calculated Benjamini-Hochberg critical values for depression (p < .03), anxiety (p < .02), and stress (p < .05)

Grit. There was not a significant group by time by grit interaction for the depression model (R^2 = .19, interaction: t = 1.39, p = .2), anxiety model (R^2 = .09 interaction: t = 0.86, p = .4), or stress model (R^2 = .08, interaction: t = -0.64, p = .5).

Program Acceptability

Satisfaction. Of students in either group who initiated the intervention, 560 provided feedback after completing the program (posttest or follow-up survey). The quality of the online

platform and program content was rated as "good" or "excellent" by 72% of respondents. When asked if they would recommend the program to a friend looking for similar help, 62% said they would. A slight majority (58%) said they were "very" or "mostly" satisfied with the program. However, only 49% said the program helped them ("yes, definitely" or "yes, I think so") more effectively cope with stress, anxiety, and/or depressed mood.

Participation rates. There 947 students who initiated the intervention (i.e., finished setting up their online account; 587 from immediate intervention, 360 from waitlist condition). For these students, the average number of modules with logged activity for these students was 3.72 (*SD* = 3.16). Individual module participation rates declined over time, with 73% for the first module and 31% for the last module. Examining adherence rates just by the immediate intervention group, 73% initiated the intervention, 57.3% completed practice for at least one module, 33.5% completed practice for at least half of the modules, and 15.2% completed practice for all modules. See Table 4.4 for full details of program adherence rates. We ran a post hoc correlation test to examine if number of modules completed related to responses on the satisfaction item and found that they were positively correlated, $r = 0.42$, $p < .001$.

Discussion

This study aimed to evaluate an online universal prevention program for anxiety, depression, and stress symptoms in university students, in terms of: effectiveness in symptom reduction, maintenance of effects, replication of effects, moderation of effects, and program acceptability. Given the growing number of online mental health interventions for university students, it is important to establish which ones are effective for symptom reduction, especially for highly prevalent mental health concerns of students like anxiety, depression, and stress. As more of such online interventions have become available, with many demonstrating favorable

results (**Chapter 1**), it is imperative to further refine such interventions for large-scale implementation. The tested online program was the first of its kind to be designed for group-level delivery with university students and it had already shown promising feasibility in an open trial (**Chapter 3**). Given the variability in rates of adherence to online-delivered programs (**Chapter 1**) and the self-motivation and self-accountability barriers (**Chapter 3**), it is also essential to understand for whom these interventions are most effective. To this end, the current study also tested internal motivation and grit as moderators of an online prevention program for university student internalizing symptoms.

Regarding effectiveness in terms of symptom reduction, we evaluated initial changes in depression, anxiety, and stress between groups. Examining group by time changes, the intervention condition had significantly decreasing slopes and produced small effects for depression, anxiety, and stress, relative to the waitlist. Small effects are consistent with the universal prevention level of the program, given that not all students are experiencing elevated symptoms at entry, and thus many participants have limited range of symptom reduction. Other online universal prevention programs that have examined overall effect sizes for depression and/or stress outcomes have also found small effects (Cavanagh et al., 2013; Frazier et al., 2015). Considering that students were allowed to access services in both conditions, the waitlist in some ways was a services-as-usual condition, and results remained significant whether including or excluding students using services from either condition. Overall, these findings provide support for the effectiveness of the online intervention in preventing symptoms of anxiety and depression from becoming clinically elevated.

In terms of maintenance and replication of effects, we evaluated if there was maintenance of changes by the intervention group and replication of changes by the original waitlist group.

Given that surpassing an inactive comparison group at post-intervention can be considered a minimum standard for demonstrating effectiveness, establishing durability and replicability of effects is a way to strengthen evidence in favor of an intervention (Rith-Najarian et al., 2017). By three-month follow-up, the intervention group had no significant changes in depression, anxiety, or stress from post treatment, indicating that reduced symptom outcomes were maintained. For students in the original waitlist condition, after accessing the intervention they had significantly decreasing slopes and small effects for depression, anxiety, and stress, demonstrating replication of symptom change in response to the online intervention. By follow-up assessment, each condition had non-significantly different outcome means from each other. Thus, the intervention initially produced significantly better outcomes relative to the waitlist group, then also showed that outcome changes were replicated and maintained, overall demonstrating strong evidence for the online intervention's effectiveness.

Examining the moderation of the intervention effect findings, internal motivation did interact with condition, but grit did not. Internal motivation for treatment moderated intervention effects such that students with high motivation experienced moderately sized effects in the intervention condition relative to the waitlist condition for depression (d = 0.5) and anxiety (d = 0.5). For students with average motivation, they also did significantly better in the intervention condition relative to waitlist, but effect sizes were minimal (depression: d = 0.16; anxiety: d = 0.16). For students with low motivation, there were no differences in symptom change between conditions. Results indicate that internal motivation alone does not result in symptom change but combined with condition it did predict even greater anxiety and depression symptom improve for students in a skills-based online intervention. We ran post hoc analyses to rule out the possibility that the effect of internal motivation was confounded by baseline severity, which could have led

to the appearance of moderation due to floor effects. Although internal motivation predicted even greater intervention response over and above baseline symptoms, we acknowledge that higher baseline symptoms are likely a requisite for an individual to have motivation for change through intervention. Still, motivation appears to be unique from severity such that not all individuals with higher baseline symptoms necessarily have higher motivation. The moderation of intervention effect was not significant for the model predicting change in stress outcome (p = .07); given that the stress subscale of DASS-21 had atypical item-factor loadings in our sample, it is unclear if moderation would have reached significance with a measure that more reliably assessed stress symptoms in this sample. Grit did not moderate intervention effects. Thus, the ability to persevere towards a goal as a trait (i.e., grit) was not as important as having internal motivation for the specific goal of wanting to get help and improve. Expanding on the mixed results across research on the role of motivation on symptom improvement within online interventions, the current study provides favorable support for the role of internal motivation specifically in facilitating more anxiety and depression improvement.

With regard to program acceptability, the evidence from satisfaction rates and participation rates provided mixed support. The majority of student users practiced skills at least once a week and the majority of feedback from respondents showed: (a) positive ratings of the online program quality, (b) likelihood to recommend it to a friend, and (c) endorsement of being "very" or "mostly" satisfied with the program. About half of students completed 50+% of the modules and about half of feedback respondents endorsed the program as being helpful for coping with symptoms. However, only about one fourth of students participated in all program weeks. Thus, although the majority of students reported program acceptability, many did not and there were variable adherence rates, suggesting that the current online skills-based intervention

was not positively received by all users. These rates are not entirely surprising given the low participant adherence for online interventions in general; for example, a review of self-guided online interventions for depression found that 17% of participants complete all intervention modules (Karyotaki et al., 2015). Still, given that module participation rate was correlated with responses to the satisfaction item, individuals who completed fewer modules were less satisfied, or vice versa. Therefore, it will be important to investigate in the future what is driving acceptability rates and adherence, and how to increase them.

Overall the intervention produced small effects, which is consistent with expected effect sizes for a universal prevention program. However, moderator analyses suggested that the overall effects are likely driven by individuals with high internal motivation, such that the moderately sized effects experienced by students with high internal motivation increased the overall effect by being averaged with the much lower effects experienced by the rest of the sample. In other words, the intervention produced moderate change for self-motivated individuals who may have been reached in larger numbers by offering an accessible and universal skills-based intervention.

Future Directions

The findings of the current study have some important implications. A clinical implication is that campus services can more strategically refer and recruit students to such programs by targeting those with self-driven motivation. Alternatively, clinicians and researchers can further investigate how to enhance motivation efficiently for any student interested in online or otherwise self-guided interventions, as this may maximize treatment effectiveness. Future research with online interventions should consider moderators that reflect individual differences such as internal motivation in order to contextualize effect sizes. It will also be important to test the role of motivation in the context of other important moderators, such as severity. Other

online mental health intervention studies have found a moderation effect by baseline symptom severity, but as far as we are aware it has never been tested along with treatment motivation. It makes conceptual sense that having higher motivation for treatment would involve some level of symptom severity, but because we did not have some baseline measure of global symptom severity, we were unable to assess the potentially unique roles of severity and internal motivation. Next, studies should begin to investigate how motivation may relate to potential mediating variables such as adherence, especially in light of data suggesting a dose-response relationship for online mental health interventions (Donkin et al., 2013). It will also be important to further develop the online intervention and test changes that might increase participant adherence regardless of motivation. Better understanding who responds to these online interventions and why will further enable clinical psychologists to maximize impact of such intervention delivery.

Limitations

The findings of the current research should be considered with some limitations in mind. First, non-payment of participants is both a strength and a limitation: although we likely had more accurate estimates of real-world adherence, there was less incentive for individuals to complete surveys which resulted in less complete data due to high drop-out. Second, although high drop-out is common with similar online intervention studies, problematically in this study there were more research drop-outs from the intervention group compared with the waitlist group by posttest assessment. It is possible that those participants who dropped out were not experiencing improvement and thus discontinued, possibly leaving a higher proportion of responders to drive intervention group effects to appear larger. Therefore, our results may have a slight positive bias, but we cannot be certain. Fortunately, there were no differences in pre-

intervention scores nor in baseline moderator variables by drop-out status within either condition. Another explanation for higher research assessment completion by the waitlist group could be that there was lower burden for them in terms of email communications, and thus they were more willing to engage with emails by the time the posttest survey was sent. Future research would benefit by using an attentional control group with parallel email volume and time requirements. Third, we relied solely on self-report measures for assessment of intervention outcomes. Findings would be more robust if findings were replicated across other kinds of measures, such as measures of functioning (e.g., academic performance), behavioral change (e.g., sleep habits), and physiological change (e.g., cardiovascular reactivity to stress). Given the online nature of the intervention, future studies could also incorporate mobile health sensory data to collect other types of outcomes. Fourth, the larger p-values across analyses for the stress scale relative to the depression and anxiety scales may have been due to the DASS-S items not cohesively loading into the original measure's stress factor. Thus, changes in stress as an outcome of the tested online intervention may be as validly estimated in our sample. Fifth, despite the sample being diverse in terms of ethnicity, there was disproportionate representation of students by gender, with females being over-represented while non-binary gendered students were not accurately represented at all. Finally, the use of a waitlist group as our control condition, as opposed to a more active control (e.g., attentional control, alternative intervention), limits our ability to disentangle if the observed outcome effects are attributable uniquely to the tested intervention or to the provision of support in general.

Conclusion

This study provided preliminary support for the effectiveness of an online skills-based program delivered at a universal-level for the prevention of depression, anxiety, and stress in

university students. In efforts to further university student mental health, it is essential that we provide students with accessible evidence-based intervention options. The intervention had small effects overall, moderate effects for students with high motivation, and adequate user acceptability. Identifying predictors – such as motivation – of symptom improvement is particularly relevant for online mental health interventions, given that drop-out is high. With the rise of personalized medicine, understanding individual differences in prevention program response will help us better tailor programs to engage and help more individuals.

Table 4.1
Descriptions of content in program modules

Module	Strategies/Skills	Example activities
1. Getting Set Up	● Treatment rationale and expectations ● Self-monitoring ● Goal Setting	● Watch the welcome video ● Review your goal plan ● Check out the "Progress" section of your dashboard
2. Make It Happen	● Behavioral activation ● Activity scheduling ● Value-driven behavior	● Do something social ● Spend time on a hobby ● Clean and organize ● Good deeds
3. Change Your Thinking	● Cognitive distortions ● Cognitive restructuring	● Identify any unhelpful thinking habits ● Adjust extreme language ● Identify evidence for and against the thought ● Shift your attention
4. Life Troubleshooting	● Problem solving ● Sleep hygiene ● Time management	● Do relaxing activities pre-sleep ● Make a worry list ● Replace your main time waster with refreshers ● Set SMART goals
5. Decide to Say It	● Interpersonally-focused graduated exposure ● Communication skills	● Go to a social event or party where you don't know everyone ● Honestly say how you're feeling ● Say “no” when you don’t have time or interest ● Respond to emails or texts that you’re tempted to avoid
6. Physical Exercise	● Physical exercise	● Go to the gym with friends ● Try a new sport, fitness activity, or class ● Yoga ● Go for a walk

7. Pause	● Mindfulness	● Eat mindfully ● Listen mindfully ● Meditate mindfully
8. Looking Ahead	● Maintenance planning	● Review progress ● Print out copies of favorite materials ● Plan for triggers

Table 4.2

Demographic Composition by Condition

	Intervention Group (n = 804)	Waitlist Group (n = 803)
Gender		
Male	203 (25.2%)	212 (26.4%)
Female	601 (74.8%)	591 (73.6%)
Ethnicity		
White non-Hispanic	276 (34.3%)	235 (29.3%)
Black non-Hispanic	12 (1.5%)	19 (2.4%)
Hispanic	133 (18.5%)	161 (20.0%)
Asian	196 (24.4%)	229 (28.5%)
Pacific Islander	1 (0.1%)	1 (0.1%)
Native American or Alaska Native	2 (0.2%)	1 (0.1%)
Two or More Races	41 (5.1%)	39 (4.9%)
International	107 (13.3%)	96 (12.0%)
Unknown	20 (2.5%)	22 (2.7%)
Type of Student		
Undergraduate	498 (61.9%)	515 (64.1%)
Graduate	306 (38.1%)	288 (35.9%)
Age		
Mean	M = 22.9, SD = 5.7	M = 22.8, SD = 5.2
Median	21	21

Table 4.3

Effect Sizes (Estimated Marginal Means) by Condition and Internal Motivation Level

		Intervention				Waitlist				
Outcome	Motivation		pre	post	*d*		pre	post	*d*	$d_1 - d_2$
Depression	High	*N* = 148	8.3 (0.31)	5.93 (0.42)	0.63	*N* = 131	8.26 (0.33)	7.77 (0.43)	0.13	0.50
	Average	*N* = 524	5.24 (0.16)	4.54 (0.22)	0.19	*N* = 526	5.12 (0.16)	5.01 (0.21)	0.03	0.16
	Low	*N* = 128	3.42 (0.33)	3.36 (0.46)	0.02	*N* = 143	3.57 (0.31)	3.35 (0.41)	0.06	-0.04
Anxiety	High	*N* = 148	6.96 (0.28)	4.81 (0.36)	0.64	*N* = 131	6.98 (0.3)	6.51 (0.37)	0.14	0.50
	Average	*N* = 524	4.46 (0.15)	3.69 (0.19)	0.23	*N* = 526	4.51 (0.15)	4.27 (0.18)	0.07	0.16
	Low	*N* = 128	3.5 (0.3)	3.4 (0.39)	0.03	*N* = 143	2.87 (0.29)	2.59 (0.35)	0.09	-0.06
Stress	High	*N* = 148	9.35 (0.32)	7.56 (0.43)	0.47	*N* = 131	9.78 (0.34)	9.5 (0.44)	0.07	0.40
	Average	*N* = 524	7.18 (0.17)	6.71 (0.23)	0.12	*N* = 526	7.21 (0.17)	7.39 (0.21)	-0.05	0.17
	Low	*N* = 128	5.63 (0.34)	5.53 (0.47)	0.03	*N* = 143	5.22 (0.32)	5.14 (0.41)	0.02	0.01

Table 4.4

Program Adherence Rates

	Overall	Immediate Intervention	Delayed Intervention
Initiated	947 (58.9 %)	587 (73.0 %)	360 (44.8 %)
1+ module	736 (45.8 %)	461 (57.3 %)	275 (34.2 %)
2+ modules	595 (37.0 %)	371 (46.1 %)	224 (27.9 %)
3+ modules	502 (31.2 %)	304 (37.8 %)	198 (24.7 %)
4+ modules	447 (27.8 %)	269 (33.5 %)	178 (22.2 %)
5+ modules	381 (23.7 %)	228 (28.4 %)	153 (19.1 %)
6+ modules	338 (21.0 %)	201 (25.0 %)	137 (17.1 %)
7+ modules	297 (18.5 %)	174 (21.6 %)	123 (15.3 %)
8 modules	224 (13.9 %)	122 (15.2 %)	102 (12.7 %)

Figure 4.1

CONSORT Flow Diagram

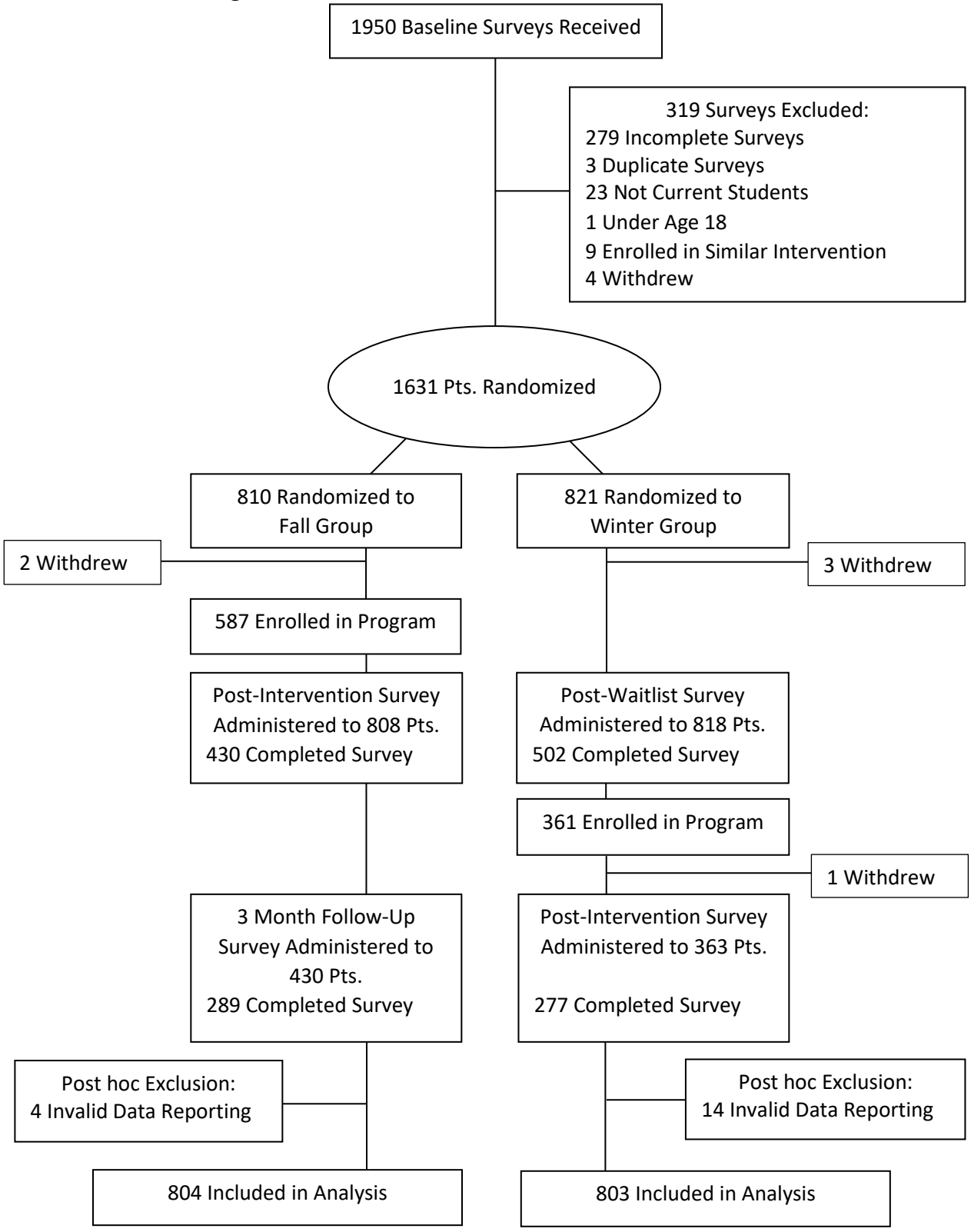
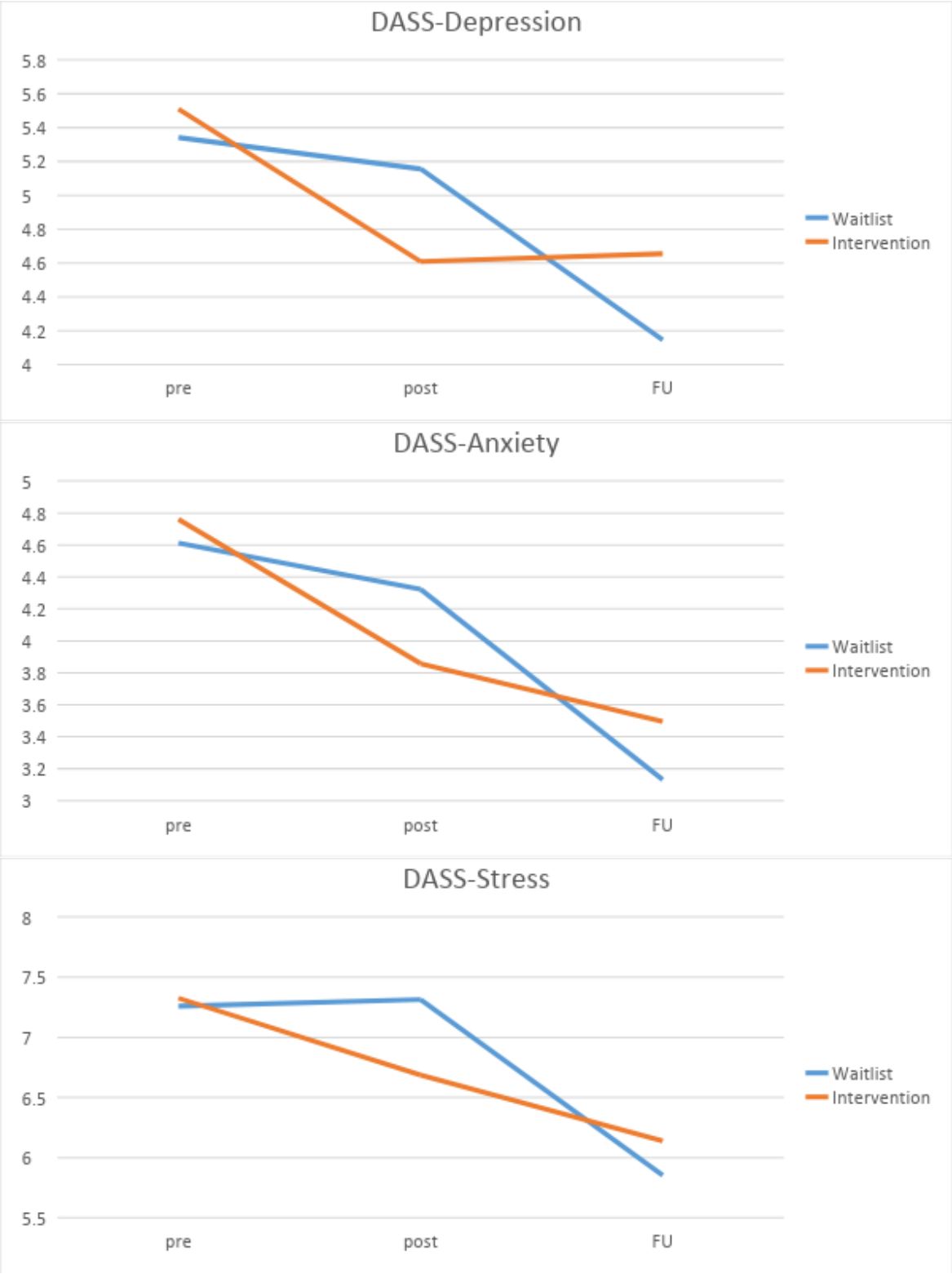


Figure 4.2

Symptom change by condition from baseline to pre-test to follow-up.



Dissertation Discussion

This dissertation involved the development and implementation of an online prevention program for anxiety and depression in university students, as researched through a series of studies. In Chapter 1, the systematic review found that: effect sizes for programs were moderate, there were practice elements that emerged as more common, and many programs were limited by disproportionately female samples and inconsistently reported adherence data. In Chapter 2, results from recruitment phases at two campuses were replicated and found that a strategic recruitment approach produced a generally representative sample students but also attracted students with "unmet need" for services and students of traditionally underserved demographics. In Chapter 3, results from an open trial found initial feasibility in terms of: ample recruitment, adequate adherence (with some variability), high program acceptability, common feedback themes, and detection of pre-post symptom changes. In Chapter 4, results from a randomized controlled trial found that the intervention program produced: adequate acceptability, small effects overall that were replicated and durable through follow-up, and moderate effects for students with high motivation. Thus, the online intervention that was developed through this research process successfully included evidence-based practices, was able to reach students in need, was feasible for implementation, and resulted in significant improvements in anxiety, depression, and stress symptoms.

Taken together, this dissertation highlights that offering online intervention programming for university students can effectively reduce symptoms of anxiety, depression, and stress, yet there is variability in who enrolls, adheres, and responds. Existing research reviews had already established that online prevention programs could be effective in symptom reduction with university students (e.g., Conley et al., 2015; Davies et al., 2014), while also identifying that they

are limited by issues such as high drop-out (e.g., 80% for unguided online mental health interventions; Cuijpers, 2018). Findings of the systematic review in **Chapter 1** were convergent with these prior reviews' findings, but also expanded meta-analytic insights into common practice elements and other limitations (e.g., service disparities by gender). The systematic review also revealed that there are few universal-level online-delivered prevention programs for university student anxiety, depression, and stress. This dissertation addressed this gap by developing and testing such an online prevention program, and simultaneously provided insights into how programs can address help-seeking barriers and improve intervention uptake, such as being strategic with branding strategies (**Chapter 2**) or building specific online platform features requested by students (**Chapters 3 and 4**). Although poor adherence rates have been a known problem for a while, the reasons for poor usage of technology-based interventions are understudied (Torous et al., 2018). The qualitative data collection in **Chapter 3** elucidated some of the barriers to engaging with online interventions, in particular the importance of self-motivation and self-accountability. There are many common barriers to students engaging with prevention programming, all further complicated by individual differences in preferences, culture, life demands, etc. Accordingly, this dissertation also attempted to further our understanding of relevant individual differences by examining differential enrollment response to branding by ethnicity and gender (**Chapter 2**), as well as moderation of symptom improvement by students' baseline motivation level (**Chapter 4**) Ultimately, this dissertation research demonstrates that online universal prevention programming can be an effective option for some otherwise underserved students. However, findings simultaneously suggest that there are still many students for whom these interventions are not the solution.

Strengths and Future Directions

A particular strength of this dissertation is its focus on how to translate research into action by taking a development for dissemination and implementation (D4D) approach and by using multiple methods (i.e., systematic review, observation data collection with replication, open trial design with quantitative and qualitative data, and a randomized controlled trial). Through a D&I-informed intervention design approach, this dissertation developed an online intervention that was grounded in research evidence, and iteratively informed by study findings and participant feedback. Thus, the dissertation research incorporated aspects of a research-to-practice model (e.g., innovations are developed through non-community research and then disseminated) and a community-centered model (e.g., innovations are locally developed and piloted), both of which should be considered when designing interventions to improve their adoption (Emshoff, 2008).

One commonly used evaluation framework of health promotion interventions is called RE-AIM (Glasgow, Vogt, & Boles, 1999), which was developed to encourage "the sustainable adoption and implementation of effective, generalizable, evidence-based interventions." The framework has five components – Reach, Effectiveness/Efficacy, Adoption, Implementation, Maintenance – that can be used a guidelines for planning or assessing programs to improve their chances of success in real-world settings. In service of assessing the strengths and remaining work of this dissertation research, I will briefly evaluate which components of RE-AIM were and were not addressed by my studies.

The first component is "Reach the target population", which means intervention development should attend to the absolute number and representativeness of users, which unfortunately few studies do. Intervention samples of UCLA students in the open trial (**Chapter 3)** and RCT (**Chapter 4**) showed that about 4% of the population enrolled and the enrollment

study (**Chapter 2**) demonstrated that samples were generally representative. Although recruitment strategies helped reach some traditionally underserved populations (i.e., Asian students, male students), further work remains to engage male students who were still underrepresented.

The second component is "Effectiveness or efficacy," which refers to the impact of an intervention on important outcomes (e.g., symptom change, quality of life, economic outcomes). Prior to developing the intervention tested in this dissertation, the systematic review (**Chapter 1**) of prevention programs established that moderate effect sizes are produced on average. The open trial (**Chapter 3**) provided important data to support the feasibility of the interventions and the RCT (**Chapter 4**) established the effects on symptom improvement were small overall but moderate for those with higher motivation. Future research should examine intervention effectiveness by assessing change in other important outcomes (e.g., functioning). Additionally, future research needs to consider how to increase the effects of such online prevention programming.

The third component is "Adoption by target staff, settings, or institutions," which refers to the need to attend to not just individual-level adoption of the intervention but institution-level adoption (e.g., how many settings are interested in delivering the intervention). This step was the least well addressed by this dissertation research. At a preliminary level, the intervention was implemented at two different campuses – UCLA and Yale. Although not reported in the dissertation, interest by other campuses is promising as indicated by the dozens of submitted requests from other universities and colleges via the intervention website's interest form. It will be important to test adaption of the program on other campuses in order to further establish implementation effectiveness for the online intervention.

The fourth component is "Implementation consistency, costs and adaptions made during delivery," which includes assessment of setting-level factors (e.g., time and cost required to implement intervention) and individual-level factors (e.g., adherence rates, participant feedback). The dissertation research included such assessment at the individual-level in terms of assessing participant adherence rates (**Chapter 3** and **Chapter 4)**. In addition, the open trial (**Chapter 3**) assessed participant feedback and then adaptations were made to the platform directly informed by that feedback, after which the new version of the online intervention was tested in the RCT (**Chapter 4).** At the individual level, qualitative work remains to determine if the adaptations made to the online intervention did in fact improve acceptability. Assessment of implementation adaptations and costs at the setting-level is still necessary. Future research could investigate how successfully the online intervention is implemented when it is being coordinated by campus administrators or health services staff, rather than a research investigator. Given the ease of content adaptation afforded by the online intervention's modularity, it would be worth investigating if intervention outcomes change depending on which modules are offered and in what order. It will also be important for cost effectiveness studies to determine if the financial costs of study implementation are lower than alternative prevention efforts, considering the relative student reach and impact.

The fifth component is "Maintenance of intervention effects in individuals and settings over time," which applies at the setting-level (i.e., if an intervention's delivery become part of an institutional routine) and at the individual-level factors (i.e., long-term effects after the intervention ends). At the individual level, the RCT (**Chapter 4**) established that intervention effects are maintained through 3-month follow-up. Examination of effect maintenance through longer-term follow-up (i.e., 1+ year) would be even more informative. At the setting-level, there

has not yet been opportunity to assess the intervention's adoption into institutional routine. If future research were to involve implementation at other campuses by non-researchers, it would be important to collect feedback from campus administrators about usability and ease of implementation, given their campus' ongoing needs.

Overall the dissertation studies address many of the points from the RE-AIM framework. However, there are many interesting and worthwhile future directions for research that could investigate the remaining gaps in assessment of the intervention's implementation effectiveness and generalizability.

Conclusion

Although many similar online intervention programs have produced significant results in research studies, evidenced-based prevention programs are by and large not being disseminated on university campuses. This dissertation provided insight into how such online programing can be offered and implemented in order to shrink the service need-use gap for young people. Although the developed intervention is not a one-size-fits all solution, it successfully engaged previously unserved students and prevented worsening of anxiety, depression, and stress symptoms for many of its users. At a larger level, the recommendations and implications from each of the studies provide branding and delivery insights that could be more widely inform other mental health prevention programs that hope to reduce the burden of mental health problems in young populations.

Appendix 1A

Review Sample References

Anshel, M. (1996). (1996). Effect of chronic aerobic exercise and progressive relaxation on motor performance and affect. *Behavioral Medicine, 21,* 186–205.

Arpin-Cribbie, C., Irvine, J., & Ritvo, P. (2012). Web-based cognitive-behavioral therapy for perfectionism: A randomized controlled trial. *Psychotherapy Research, 22,* 194–207. https://doi.org/10.1080/10503307.2011.637242

Baghurst, T., & Kelley, B. C. (2014). An Examination of stress in college students over the course of a semester. *Health Promotion Practice, 13,* 438-447. https://doi.org/10.1177/1524839913510316

Bearman, S. K., Stice, E., & Chase, A. (2003). Evaluation of an intervention targeting both depressive and bulimic pathology: A randomized prevention trial. *Behavior Therapy, 34,* 277–293. https://doi.org/10.1016/j.biotechadv.2011.08.021.Secreted

Braithwaite, S. R., & Fincham, F. D. (2007). ePREP: Computer Based Prevention of Relationship Dysfunction, Depression and Anxiety. *Journal of Social and Clinical Psychology, 26,* 609–622. https://doi.org/10.1521/jscp.2007.26.5.609

Cavanagh, K., Strauss, C., Cicconi, F., Griffiths, N., Wyper, A., & Jones, F. (2013). A randomised controlled trial of a brief online mindfulness-based intervention. *Behaviour Research and Therapy, 51,* 573–578. https://doi.org/10.1016/j.brat.2013.06.003

Chen, Y., Yang, X., Wang, L., & Zhang, X. (2013). A randomized controlled trial of the effects of brief mindfulness meditation on anxiety symptoms and systolic blood pressure in Chinese nursing students. *Nurse Education Today, 33,* 1166–1172. https://doi.org/10.1016/j.nedt.2012.11.014

Chiauzzi, E., Brevard, J., Thum, C., Decembrele, S., & Lord, S. (2008). MyStudentBody-Stress: an online stress management intervention for college students. *Journal of Health Communication, 13,* 555–572. https://doi.org/10.1080/10810730802281668

Cukrowicz, K. C., & Joiner, T. E. (2007). Computer-based intervention for anxious and depressive symptoms in a non-clinical population. *Cognitive Therapy and Research, 31,* 677–693. https://doi.org/10.1007/s10608-006-9094-x

Danitz, S. B., & Orsillo, S. M. (2014). The mindful way through the semester: An investigation of the effectiveness of an acceptance-based behavioral therapy program on psychological wellness in first-year students. *Behavior Modification, 38,* 549–566. https://doi.org/10.1177/0145445513520218

Day, V., McGrath, P. J., & Wojtowicz, M. (2013). Internet-based guided self-help for university students with anxiety, depression and stress: A randomized controlled clinical trial. *Behaviour Research and Therapy, 51,* 344–351. https://doi.org/10.1016/j.brat.2013.03.003

Deckro, G. R., Ballinger, K. M., Hoyt, M., Wilcher, M., Dusek, J., Myers, P., … Benson, H. (2002). The evaluation of a mind/body intervention to reduce psychological distress and perceived stress in college students. *Journal of American College Health, 50,* 281–287. https://doi.org/10.1080/07448480209603446

Ellis, L. a, Campbell, A. J., Sethi, S., & O'Dea, B. M. (2011). Comparative randomized trial of an online cognitive-behavioral therapy program and an online support group for depression and anxiety. *Journal of CyberTherapy and Rehabilitation, 4,* 461–467. https://doi.org/10.1037/t00735-000

Fehring, R. J. (1983). Effects of biofeedback-aided relaxation on the psychological stress symptoms of college students. *Nursing Research, 32*, 362–366. https://doi.org/10.1097/00006199-198311000-00009

Fontana, A. M., Hyra, D., Godfrey, L., & Cermak, L. (1999). Impact of a peer-led stress inoculation training intervention on state anxiety and heart rate in college students. *Journal of Applied Biobehavioral Research, 4,* 45–63. https://doi.org/10.1111/j.1751-9861.1999.tb00054

Frazier, P., Meredith, L., Greer, C., Paulsen, J. A., Howard, K., Dietz, L. R., & Qin, K. (2015). Randomized controlled trial evaluating the effectiveness of a web-based stress management program among community college students. *Anxiety, Stress and Coping, 28*, 576–586. https://doi.org/10.1080/10615806.2014.987666

Frögéli, E., Djordjevic, A., Rudman, A., Livheim, F., & Gustavsson, P. (2016). A randomized controlled pilot trial of acceptance and commitment training (ACT) for preventing stress-related ill health among future nurses. *Anxiety, Stress and Coping, 29*, 202–218. https://doi.org/10.1080/10615806.2015.1025765

Gaab, J., Blättler, N., Menzi, T., Pabst, B., Stoyer, S., & Ehlert, U. (2003). Randomized controlled evaluation of the effects of cognitive–behavioral stress management on cortisol responses to acute stress in healthy subjects. *Psychoneuroendocrinology, 28,* 767–779.

Galante, J., Dufour, G., Vainre, M., Wagner, A. P., Stochl, J., Benton, A., … Jones, P. B. (2018). A mindfulness-based intervention to increase resilience to stress in university students (the Mindful Student Study): a pragmatic randomised controlled trial. *The Lancet Public Health, 3,* e72–e81. https://doi.org/10.1016/S2468-2667(17)30231-1

Gallego, J., Aguilar-Parra, J. M., Cangas, A. J., Langer, Á. I., & Manãs, I. (2015). Effect of a Mindfulness Program on Stress, Anxiety and Depression in University Students. *Spanish Journal of Psychology, 17*, E109. https://doi.org/10.1017/sjp.2014.102

Gardenswartz, C. A., & Craske, M. G. (2001). Prevention of panic disorder. *Behavior Therapy, 32*, 725–737. https://doi.org/10.1016/S0005-7894(01)80017-4

Geisner, I. M., Neighbors, C., & Larimer, M. E. (2006). A randomized clinical trial of a brief, mailed intervention for symptoms of depression. *Journal of Consulting and Clinical Psychology, 74*, 393–399. https://doi.org/10.1037/0022-006X.74.2.393

Grassi, A., Gaggioli, A., & Riva, G. (2009). The green valley: the use of mobile narratives for reducing stress in commuters. *Cyberpsychology & Behavior, 12*, 155–161. https://doi.org/10.1089/cpb.2008.0156

Grassi, A., Gaggioli, A., & Riva, G. (2011). New technologies to manage exam anxiety. *Annual Review of CyberTherapy and Telemedicine, 9*, 26–49. https://doi.org/10.3233/978-1-60750-766-6-57

Greeson, J. M., Juberg, M. K., Maytan, M., James, K., & Rogers, H. (2014). A randomized controlled trial of Koru: A mindfulness program for college students and other emerging adults. *Journal of American College Health, 62*, 222–233. https://doi.org/10.1080/07448481.2014.891595

Hamdan-Mansour, A. M., Puskar, K., & Bandak, A. G. (2009). Effectiveness of cognitive-behavioral therapy on depressive symptomatology, stress and coping strategies among Jordanian university students. *Issues in Mental Health Nursing, 30*, 188–196. https://doi.org/10.1080/01612840802694577

Hammerfald, K., Eberle, C., Grau, M., Kinsperger, A., Zimmermann, A., Ehlert, U., & Gaab, J. (2006). Persistent effects of cognitive-behavioral stress management on cortisol responses to acute stress in healthy subjects--a randomized controlled trial. *Psychoneuroendocrinology, 31*, 333–339. https://doi.org/10.1016/j.psyneuen.2005.08.007

Hazlett-Stevens, H., & Oren, Y. (2016). Effectiveness of mindfulness-based stress reduction bibliotherapy: A Preliminary randomized controlled trial. *Journal of Clinical Psychology, 73*, 626–637. https://doi.org/10.1002/jclp.22370

Heaman, D. (1995). The quieting response (QR): a modality for reduction of psychophysiologic stress in nursing students. *The Journal of Nursing Education, 34*, 5–10.

Jain, S., Shapiro, S. L., Swanick, S., Roesch, S. C., Mills, P. J., Bell, I., & Schwartz, G. E. R. (2007). A randomized controlled trial of mindfulness meditation versus relaxation training: effects on distress, positive states of mind, rumination, and distraction. *Annals of Behavioral Medicine, 33*, 11–21. https://doi.org/10.1207/s15324796abm3301_2

Johansson, N. (1991). Effectiveness of a stress management program in reducing anxiety and depression in nursing students. *Journal of American College Health, 40*, 125-129 https://doi.org/10.1080/07448481.1991.9936268

Jones, M. C., & Johnston, D. W. (2000). Evaluating the impact of a worksite stress management programme for distressed student nurses: A randomised controlled trial. *Psychology & Health, 15*, 689–706. https://doi.org/10.1080/08870440008405480

Kang, Y. S., Choi, S. Y., & Ryu, E. (2009). The effectiveness of a stress coping program based on mindfulness meditation on the stress, anxiety, and depression experienced by nursing students in Korea. *Nurse Education Today, 29*, 538–543. https://doi.org/10.1016/j.nedt.2008.12.003

Kanji, N., White, A., & Ernst, E. (2006). Autogenic training to reduce anxiety in nursing students: Randomized controlled trial. *Journal of Advanced Nursing, 53*, 729–735. https://doi.org/10.1111/j.1365-2648.2006.03779.x

Kenardy, J., McCafferty, K., & Rosa, V. (2003). Internet-delivered indicated prevention for anxiety disorders: A randomized controlled trial. *Behavioural and Cognitive Psychotherapy, 31*, 279–289. https://doi.org/10.1017/S1352465803003047

Kim, K. B., Cohen, S. M., Oh, H. K., & Sok, S. R. (2004). The effects of meridian exercise on anxiety, depression, and self-esteem of female college students in korea. *Holistic Nursing Practice, 18*, 230–234. https://doi.org/10.1097/00004650-200409000-00004

Lintvedt, O. K., Griffiths, K. M., Sørensen, K., Østvik, A. R., Wang, C. E. a, Eisemann, M., & Waterloo, K. (2013). Evaluating the effectiveness and efficacy of unguided internet-based self-help intervention for the prevention of depression: A randomized controlled trial. *Clinical Psychology & Psychotherapy, 20*, 10–27. https://doi.org/10.1002/cpp.770

McEntee, D. J., & Halgin, R. P. (1999). Cognitive group therapy and aerobic exercise in the treatment of anxiety. *Journal of College Student Psychotherapy, 13*, 37–55.

Musiat, P., Conrod, P., Treasure, J., Tylee, A., Williams, C., & Schmidt, U. (2014). Targeted prevention of common mental health disorders in university students: Randomised controlled trial of a transdiagnostic trait-focused web-based intervention. *PLoS ONE, 9*, e93621. https://doi.org/10.1371/journal.pone.0093621

Oman, D., Shapiro, S. L., Thoresen, C. E., Plante, T. G., & Flinders, T. (2008). Meditation lowers stress and supports forgiveness among college students: a randomized controlled trial. *Journal of American College Health, 56*, 569–578. https://doi.org/10.3200/JACH.56.5.569-578

Orbach, G., Lindsay, S., & Grey, S. (2007). A randomised placebo-controlled trial of a self-help Internet-based intervention for test anxiety. *Behaviour Research and Therapy, 45*, 483–496. https://doi.org/10.1016/j.brat.2006.04.002

Peden, A. R., Rayens, M. K., Hall, L. A., & Beebe, L. H. (2001). Preventing depression in high-risk college women: A report of an 18-month follow-up. *Journal of American College Health, 49*, 299–306.

Perna, F. M., Antoni, M. H., Kumar, M., Cruess, D. G., & Schneiderman, N. (1998). Cognitive-behavioral intervention effects on mood and cortisol during exercise training. *Annals of Behavioral Medicine, 20*, 92–98. https://doi.org/10.1007/BF02884454

Phang, C. K., Mukhtar, F., Ibrahim, N., Keng, S.-L., & Mohd. Sidik, S. (2015). Effects of a brief mindfulness-based intervention program for stress management among medical students: the Mindful-Gym randomized controlled study. *Advances in Health Sciences Education, 20*, 1115–1134. https://doi.org/10.1007/s10459-015-9591-3

Philpot, V. D., & Bamburg, J. W. (1996). Rehearsal of positive self-statements and restructured negative self-statements to increase self-esteem and decrease depression. *Psychological Reports, 79*, 83–91. https://doi.org/10.2466/pr0.1996.79.1.83

Ratanasiripong, P., Ratanasiripong, N., & Kathalae, D. (2012). Biofeedback intervention for stress and anxiety among nursing students: A Randomized controlled trial. *International Scholarly Research Network Nursing, 2012*, 1–5. https://doi.org/10.5402/2012/827972

Rose, R. D., Buckey, J. C., Zbozinek, T. D., Motivala, S. J., Glenn, D. E., Cartreine, J. a., & Craske, M. G. (2013). A randomized controlled trial of a self-guided, multimedia, stress management and resilience training program. *Behaviour Research and Therapy, 51*, 106–112. https://doi.org/10.1016/j.brat.2012.11.003

Rose, R. L., & Veiga, J. F. (1984). Assessing the sustained effects of a stress management intervention on anxiety and locus of control. *Academy of Management Journal, 27*, 190–198. https://doi.org/10.2307/255967

Seligman, M. E. P., Schulman, P., DeRubeis, R. J., & Hollon, S. D. (1999). The prevention of depression and anxiety. *Prevention & Treatment, 2*, 1–24. https://doi.org/10.1037/1522-3736.2.1.28a

Sethi, S., Campbell, A. J., & Ellis, L. A. (2010). The use of computerized self-help packages to treat adolescent depression and anxiety. *Journal of Technology in Human Services, 28,* 144–160. https://doi.org/10.1080/15228835.2010.508317

Shankarapillai, R., Nair, M. A., & George, R. (2012). The effect of yoga in stress reduction for dental students performing their first periodontal surgery: A randomized controlled study. *International Journal of Yoga, 5,* 48–51. https://doi.org/10.4103/0973-6131.91714

Shapiro, S. L., Schwartz, G. E., & Bonner, G. (1998). Effects of mindfulness-based stress reduction on medical and premedical students. *Journal of Behavioral Medicine, 21*, 581–599. https://doi.org/10.1023/A:1018700829825

Shearer, A., Hunt, M., Chowdhury, M., & Nicol, L. (2016). Effects of a brief mindfulness meditation intervention on student stress and heart rate variability. *International Journal of Stress Management, 23*, 232. https://doi.org/10.1037/a0039814

Song, Y., & Lindquist, R. (2014). Effects of mindfulness-based stress reduction on depression, anxiety, stress and mindfulness in Korean nursing students. *Nurse Education Today, 35,* 86–90. https://doi.org/10.1016/j.nedt.2014.06.010

Steinhardt, M., & Dolbier, C. (2008). Evaluation of a resilience intervention to enhance coping strategies and protective factors and decrease symptomatology. *Journal of American College Health, 56*, 445–453. https://doi.org/10.3200/JACH.56.44.445-454

Stephens, R. L. (1992). Imagery: a treatment for nursing student anxiety. *The Journal of Nursing Education, 31*, 314–320.

Taylor, B. L., Strauss, C., Cavanagh, K., & Jones, F. (2014). The effectiveness of self-help mindfulness-based cognitive therapy in a student sample: a randomised controlled trial. *Behaviour Research and Therapy, 63*, 63–69. https://doi.org/10.1016/j.brat.2014.09.007

Warnecke, E., Quinn, S., Ogden, K., Towle, N., & Nelson, M. R. (2011). A randomised controlled trial of the effects of mindfulness practice on medical student stress levels. *Medical Education, 45*, 381–388. https://doi.org/10.1111/j.1365-2923.2010.03877.x

Whitehouse, W. G., Dinges, D. F., Orne, E. C., Keller, S. E., Bates, B. L., Bauer, N. K., ... & Orne, M. T. (1996). Psychosocial and immune effects of self-hypnosis training for stress management throughout the first semester of medical school. *Psychosomatic Medicine, 58*, 249–263.

Wolitzky-Taylor, K. B., & Telch, M. J. (2010). Efficacy of self-administered treatments for pathological academic worry: A randomized controlled trial. *Behaviour Research and Therapy, 48*, 840–850. https://doi.org/10.1016/j.brat.2010.03.019

Yusoff, M. S. B., & Esa, A. R. (2015). A DEAL-based intervention for the reduction of depression, denial, self-blame and academic stress: A randomized controlled trial. *Journal of Taibah University Medical Sciences, 10,* 82–92. https://doi.org/10.1016/j.jtumed.2014.08.003

Zuroff, D. C., & Schwarz, J. C. (1978). Effects of transcendental meditation and muscle relaxation on trait anxiety, maladjustment, locus of control, and drug use. *Journal of Consulting and Clinical Psychology, 46,* 264–271. https://doi.org/10.1037/0022-006X.46.2.264

Appendix 1B
Descriptive Characteristics of Prevention Programs

Article	Prevention Level	Target(s)	Study Sample	Winning group(s)	Primary Delivery	Measure(s) with Significant Outcomes	Adherence
Anshel (1996)	S	S	*n* = 60; male university students; 0% female; age: 19.3 - 25.6; *M* = 21.9	Aerobic exercise	S	POMS-anxiety, POMS-depression, heart rate, blood pressure	100% completed all self-training sessions
				Progressive relaxation	S	heart rate	100% completed all self-training sessions
Arpin-Cribbie et al. (2012)	I	A, D	*n* = 83; university students in a psychology course; Canada; 70% female; age: 18 - 48; *M* = 20.14	Cognitive behavioral intervention	T	CES-D, BAI	57.7 - 92.3 %: rates of reading each full module
				General stress management	T	BAI	50.0 - 90.0%: rates of reading each full module
Baghurst & Kelley (2014)	U	S	*n* = 601; undergraduate students; 45% female	Physical activity	G	PSS	
				Stress management	G	PSS	

Bearman et al. (2003)	S	D	*n* = 74; female university students with body image concerns; 100% female; 48% Caucasian; 33% Asian / Pacific Islander; 14% Hispanic; 4% mixed/other; 1% Black; age: 17 - 20; *M* = 18.9	CBT intervention	G	BDI, PANAS-negative	
Braithwaite & Fincham (2007)	S	A, D	*n* = 91; undergraduate psychology students in a romantic relationship; 59% female; 61% Caucasian; 19% Asian; 6% African American; 14% Other	CBASP	T	BAI, BDI, PANAS	
				ePREP	T	BAI, BDI, PANAS	
Cavanagh et al. (2013)	U	A, D, S	*n* = 104; university students; United Kingdom; 88% female; age: 19 - 51	Learning Mindfulness Online	T	PSS, PHQ-4	100% read some emails; 87% read most/all emails; 87% practiced more than weekly; 26% practiced daily or more
Chen et al. (2013)	U	A, D	*n* = 60; nursing students; China; 87% female; 100%	Meditation group	G	SAS, blood pressure, heart rate	

			Han ethnicity; age: 18 - 22; *M* = 19.5				
Chiauzzi et al. (2008)	I	S	*n* = 240; undergraduate students at four-year colleges; 51% female; 54% White/Caucasian; 16% Asian; 12% Black/African American; 8% Hispanic; 8% Other	MyStudent Body-Stress	T	CAS-anxiety	96% "received the intervention"; 126 minutes using online program on average
Cukrowicz & Joiner (2007)	U	A, D	*n* = 152; undergraduate students in psychology; 74% female	CBASP	T	BAI, BDI, PANAS-positive	100% attended the single computer session; 60% completed 1+ practice worksheet
Danitz & Orsillo (2014)	S	A, D, S	*n* = 98; undergraduate prelaw students and law students; 80% female; 69% White; 12% Hispanic, 12% multi-racial; 8% Asian/Pacific Islander; 4% African American; 4% Other; age: 18 - 39; *M* = 21.13	Workshop	G	DASS-21-depression	43% attended the workshop; subsequent self-guided practice rates not reported
Day et al. (2013)	I	A, D, S	*n* = 66; university students; Canada; 89% female; age: *M* = 24	Immediate access CBT	T	DASS-21-depression,	

						DASS-21-anxiety, DASS-21-stress	
Deckro et al. (2002)	U	A, S	*n* = 128; undergraduate and graduate students; 60% female; age: 17 - 60; *M* = 24	Maximize Your Potential	G	GSI, PSS, STAI-state	Attendance: 73% 1+ session, 65% 3+ sessions, 32% all 6 sessions
Ellis et al. (2011)**	I	A, D	*n* = 39; university students; Australia; 77% female; age: 18 - 25; *M* = 19.67	MoodGarden	T	DASS-21-anxiety	
Fehring (1983)	U	S	*n* = 90; undergraduate students; 78% female; age: 18 - 35; *M* = 22.8	Biofeedback-aided relaxation	S	POMS, STAI-state	4.9 average biofeedback practice per week
Fontana et al. (1999)	U	A, S	*n* = 36; undergraduate students in psychology; 58% female; 86% White; 6% Black; 5% Latino; 3% Asian; age: *M* = 18.75	Stress inoculation training group	G	STAI-state, heart rate	100% attended all 6 sessions
Frazier et al. (2015)	U	A, D, S	*n* = 194; undergraduate community college students; 75% female; 71% European American/White; age: 18 - 45+; Modal 18–21	Present control intervention	T	DASS-21-depression, DASS-21-anxiety, DASS-21-stress	63-83%; adherence rates for the three individual modules, based on stress logs

Frögéli et al. (2016)	U	S	*n* = 113; nursing students; Sweden; % female not reported	ACT intervention	G	PSS	42% attended half (3 of 6) or more of sessions
Gaab et al. (2003)	U	S	*n* = 48; male students at a science, technology, engineering and mathematics university; Switzerland; 0% female; age: *M* = 24	Stress inoculation training group	G	PSS, lower integrated salivary free cortisol response	
Galante et al. (2018)	U	S	*n* = 616; university students; United Kingdom; 63% female; 69% White; 20% Asian; 7% Mixed; 3% Other; 1% Black; age: 18 - 31+	Mindfulness Skills for Students	G	CORE-OM-distress, WEMWBS	86% started the course; 59% received 4 sessions, < 40% completed all 8 weeks
Gallego et al. (2015)	U	A, D, S	*n* = 125; undergraduate students; Spain; 58% female; age: 18 - 43; *M* = 20.07	Mindfulness	G	DASS-21-depression, DASS-21-anxiety, DASS-21-stress	
Gardenswartz & Craske (2001)	I	A	*n* = 121; undergraduate psychology students; 69% female; 39% Caucasian; 30% Asian; 10.6% Hispanic; 5.7% African American; 10.6% other; age: 18 - 39; *M* = 20.3	Workshop group	G	FQ, Panic attack occurrence	85% of those assigned attended the massed 5-hour workshop

Geisner et al. (2006)	I	D	*n* = 177; undergraduate students in psychology; 70% female; 49% White or Caucasian; 48% Asian or Asian America; age: 18+; *M* = 19.28	Intervention	S	DSM-Depression scale	79% reported receiving the mailed materials; 57% reported reading them carefully
Grassi et al. (2009)	S	A, S	*n* = 120; university students who commute; Italy; 50% female; age: 20 – 25	Vidnar	S	STAI-state	
Grassi et al. (2011)	U	A	*n* = 75; female university students; Italy; 100% female; age: 20 - 23; *M* = 20.86	Audio and video narrative on mobile phone	S	STAI-state	
Greeson et al. (2014)	U	S	*n* = 90; undergraduate, graduate, and professional students; 66% female; 62% White; 26% Asian American; 6% African American; 4% Other; 2% Prefer not to answer; age: 18 - 59; *M* = 25.4	Koru mindfulness training program	G	PSS	93% attended at least one session; 89% attended 2+ sessions; 33% attended all 4 sessions

Hamdan-Mansour et al. (2009)	I	D, S	*n* = 84; university students; Jordan; 45% female	Modified Teaching Kids to Cope	G	BDI, PSS	
Hammerfald et al. (2006)	U	S	*n* = 83; undergraduate students in psychology; Switzerland; 71% female	Cognitive behavioral stress management	G	PASA-perceived stress index, cortisol	
Hazlett-Stevens & Oren (2016)	U	A, D, S	*n* = 92; university and community college students; 75% female; 63% Caucasian; 21% Hispanic or Latino; 7% Asian or Asian American; 6% multiracial; 3% other; age: *M* = 22.1	Bibliotherapy	S	PSS, DASS-21-depression, DASS-21-anxiety, DASS-21-stress	79% read half+ of first chapter and 72% did half+ its exercises; 34% read half+ of last chapter and 22% did half+ its exercises
Heaman (1995)	S	S	*n* = 45; undergraduate nursing students; 100% female; age: 20 - 50; *M* = 28.9	Experimental group	G	STAI-state	
Jain et al. (2007)	I	S	*n* = 104; medical, graduate nursing, and undergraduate premed students; 81% female; 63% White; 16% Hispanic; 5% Native American; 7% Asian/ Pacific Islander; 2.5% mixed; 6.2% unknown; age: 18 - 61; *M* = 25	Mindfulness intervention	G	GSI	5.7 average hours of practice outside sessions
				Somatic relaxation	G	GSI	5.7 average hours of practice outside sessions

Johansson (1991)	U	A, D, S	*n* = 76; female undergraduate nursing students; 100% female; 92% Caucasian; age: 19 - 30; *M* = 21.5	Experimental group	G	STAI-state, IPAT-Depression	
Jones & Johnston (2000)	I	S	*n* = 79; nursing students; United Kingdom; 85% female; age: *M* = 27.3	Treatment group	G	BDI, STAI-trait	
Kang et al. (2009)	S	A, D, S	*n* = 41; nursing students; South Korea; 100% female; age: *M* = 22	Experimental group	G	STAI, PWI	76% attended 6+ of 8 sessions
Kanji et al. (2006)	S	A	*n* = 93; nursing students; United Kingdom; 91% female; age: 19 – 49	Autogenic training	G	STAI-state, STAI-trait, blood pressure, pulse rate	57% reported practicing at 5-month follow-up, with 21% practicing at least once daily
Kenardy et al. (2003)	I	A	*n* = 83; undergraduate students in psychology; Australia; 62% female; age: *M* = 20.73	Evaluation group	T	CES-D	3.36 of 6 lessons completed on average; 90.37 minutes spent on the online program

Kim et al. (2004)	S	A, D	*n* = 54; female undergraduate students; South Korea; 100% female; age: 19 - 24	Meridian exercise intervention	G	STAI, DSI	
Lintvedt et al. (2013)	I	D	*n* = 163; university students; Norway; 77% female; age: *M* = 28.2	Internet intervention	T	CES-D	77% used the online program
McEntee & Halgin (1999)***	I	A, S	*n* = 80; undergraduate students in psychology; 68% female; age: *M* = 20.2	Cognitive group counseling combined with aerobic exercise	G	STAI-state, STAI-trait	
Musiat et al. (2014)	I****	A, D	*n* = 1047; university students; United Kingdom; 70% female; 69% White; 22% Asian; 1% Black; 7% Other; age: 18 - 57; *M* = 21.8	PLUS	T	PHQ-9, GAD-7	47% of students "completed a module at 12-week follow-up after starting it"
Oman et al. (2008)	U	S	*n* = 47; undergraduate students; 80% female; 73% White; 27% non-White; age: 18 – 24	Mindfulness-based stress reduction	G	PSS	88% participated; 35% attended all sessions

Orbach et al. (2007)	I	A	*n* = 58; undergraduate, graduate, and professional students; United Kingdom; 72% female; age: *M* = 23	CBT group	T	TAI	100% "looked" at all modules; 33% did last module; average time = 157 mins
Peden et al. (2001)	I	D	*n* = 92; female undergraduate students; 100% female; age: 18 - 24; *M* = 19.3	Cognitive behavioral group intervention	G	BDI, CES-D	
Perna et al. (1998)	S	S	*n* = 34; undergraduate student athletes; 59% female; age: *M* = 19	Cognitive behavioral stress management program	G	POMS-depression, cortisol	100% attended 4+ sessions; 79% attended 6+ sessions; 50% attended all 7 sessions; average audio tapes use = 2.11 time/week
Phang et al. (2015)	U	S	*n* = 75; medical students; Malaysia; 76% female; 53% Malay; 37% Chinese; 9% Indian; age: *M* = 21	Mindful Gym	G	PSS	100% attended 2+ sessions; 49% attended all 5 sessions
Philpot & Bamburg (1996)	I	D	*n* = 60; undergraduate students in psychology; 63% female; 80% Caucasian; 20% Black; age: *M* = 21.4	Rehearsal group	S	BDI	

Ratanasiripong et al. (2012)	S	A, S	*n* = 60; nursing students; Thailand; 100% female; age: 18 - 21; *M* = 19.27	Biofeedback intervention	S	STAI-state	
Rose & Veiga (1984)	U	A, S	*n* = 48; undergraduate business students; 46% female; age: *M* = 22	Experimental group	G	STAI-trait	
Rose et al. (2013)	I	S	*n* = 66; graduate students; 50% female; 52% Caucasian; 32% Asian or Asian American; 9% Hispanic or Latino; 7% other; age: *M* = 27.32	SMART-OP	T	PSS, amylase recovery, blood pressure	88% completed all 6 self-guided training sessions
Seligman et al. (1999)	I	A, D	*n* = 231; undergraduate students; 52% female	Prevention workshop	G	BDI, HDRS, HARS	
Sethi et al. (2010)	I	A, D	*n* = 38; undergraduate students; Australia; 66% female; age: 18 - 23; *M* = 19.47	MoodGYM	T	DASS-21-anxiety, K10	
Shankarapillai et al. (2012)	S	A	*n* = 100; undergraduate dentistry students; India; 43% female; age: *M* = 22	Yoga intervention group	S	STAI-state	81% listened to the audio tape and practiced lessons

Shapiro et al. (1998)	S	A, D, S	*n* = 78; premedical undergraduate students and medical students; 56% female; 79 % Caucasian; 8% Hispanic; 5% Indian; 3% African American; 3% Asian American	Active intervention	G	STAI-state, STAI-trait, GSI, SCL-90-depression	97% "completed the intervention"
Shearer et al. (2016)	U	A, D, S	*n* = 74; undergraduate students in psychology; 57% female; 43% Asian, 41% Caucasian, 7% Hispanic, 3% African American, 3% other, 1% Native American, 1% Pacific Islander, and 1% unidentified	Brief mindfulness meditation group	G	heart rate variability, PANAS	
Song et al. (2014)	S	A, D, S	*n* = 50; nursing students; South Korea; 82% female; age: *M* = 19.5	Mindfulness-based stress reduction	G	DASS-21-depression, DASS-21-anxiety, DASS-21-stress	92% "received the intervention"
Steinhardt & Dolbier (2008)	U	S	*n* = 64; undergraduate and graduate students; 82% female; 43.9% Caucasian; 26.3% Asian; 19.3% Hispanic; 5.3% African American; 5.2% other	Resilience intervention	G	CES-D, PANAS, PSS	

Stephens (1992)	S	A	*n* = 159; female nursing students; 100% female; age: 18 - 53	Imagery/relaxation group	S	STAI-state	
Taylor et al. (2014)	U	A, D, S	*n* = 80; undergraduate and graduate students; United Kingdom; 81% female; 86% White; 14% Non-white; age: *M* = 28.61	Mindfulness-based cognitive therapy self-help	S	DASS-21-anxiety, DASS-21-depression	85% read at least half of the book; 58% read the whole book; 58% practiced at least weekly; median of 2-3x weekly practice for 10-20 minutes
Warnecke et al. (2011)	S	S	*n* = 66; medical students; Australia; 65% female; age: *M* = 23.92	Mindfulness-based stress reduction	S	PSS, DASS-21-anxiety	65% completed a practice record; average 26.7 days of practice during the 8 weeks (vs. 54 days recommended)
Whitehouse et al. (1996)	S	S	*n* = 35; medical students; 60% female; age: *M* = 24.8	Self-hypnosis training condition	G	BSI-anxiety	self-hypnosis practiced 5x/week on average; group attendance not reported
Wolitzky-Taylor & Telch (2010)	I	A	*n* = 113; undergraduate and graduate students; 75% female; 45%	Audio-photic stimulation	S	PSS, PSWQ	70% completed at least 1/3 of home sessions; 6.36

			Caucasian; 25% Asian American; 17% Hispanic; 34% African-American; 8% biracial/multiracial; 1% Native American; .1% Pacific Islander				home sessions completed on average
				Worry exposure	S	PSS, PSWQ	83% completed at least 1/3 of home sessions; 7.93 home sessions completed on average
Yusoff & Esa (2015)	U	D, S	*n* = 171; medical students, Malaysia; 65% female; 78% Malay; 22% Non-Malay	DEAL-based intervention	G	BDI, MSSQ	100% received the 1-day workshop; subsequent self-guided practice rates not reported
Zuroff & Schwarz (1978)	U	A	*n* = 61; undergraduate students in psychology; 51% female	Transcendental meditation	G	SRIA	

Notes: Prevention level: U = universal; S = selective; I = indicated. Primary delivery: G = in-person group format; S = self-administered materials; T = technology format (online or computer-delivered). Target(s): A = anxiety, D = depression, S = stress. Measure name abbreviations: BAI=Beck Anxiety Inventory; BDI=Beck Depression Inventory; BSI=Brief Symptom Inventory; CAS=College Adjustment Scale; CES-D=Center for Epidemiologic Studies Depression Scale; CORE-OM=Clinical Outcomes in Routine Evaluation Outcome Measure; DASS=Depression, Anxiety, & Stress Scale; DSI=Depression Status Inventory; FQ=Fear Questionnaire; GAD=Generalized Anxiety Disorder; GSI=Global Severity Index; HARS=Hamilton Anxiety Rating Scale; HDRS=Hamilton Depression Rating Scale; IPAT=Institute for Personality and Ability Testing; K10=Kessler Psychological Distress Scale ; MSSQ=Medical Student Stressor Questionnaire ; PANAS=Positive and Negative Affect Scale; PASA=Primary Appraisal

Secondary Appraisal; PHQ=Patient Health Questionnaire; POMS=Profile of Mood States; PSS=Perceived Stress Scale; PSWQ=Penn State Worry Questionnaire; PWI=Psychosocial Wellbeing Index; SAS=Self-Rating Anxiety Scale; SCL=Symptom Checklist; SRIA=S-R Inventory of Anxiousness; STAI=State-Trait Anxiety Inventory; TAI=Test Anxiety Inventory; WEMWBS=Warwick–Edinburgh Mental Wellbeing Scale

*If country is not indicated, study took place in the United States
**MoodGym not coded as it was already coded in Sethi et al. (2010) article
***There were two other groups that were subcomponents of the coded intervention, and thus were not coded as separate groups
****enrollment was at the universal level, but significant between-group change results were only found for high risk students

Appendix 2A

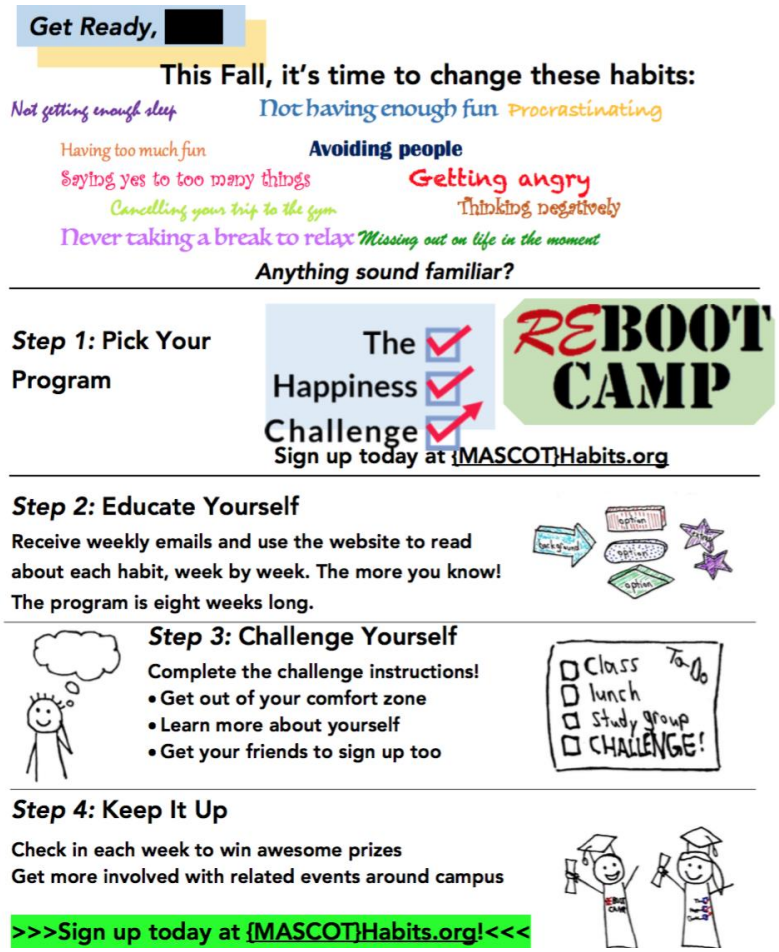
Supplemental Figures: Comparison of Example Recruitment Materials

Campus 1:

Campus 2:

Figure S1. Comparison of graphic ads specific to each program.

Campus 1:



Campus 2:

Figure S2. Comparison of short graphic ad displaying both programs.

Campus 1 (public university in CA)	Campus 2 (private university in CT)
Habit Change at [redacted]: 8 Weeks of Goals, Activities, Events, Prizes, and Life Improvement This spring quarter, there will be two programs for [redacted] students - The Happiness Challenge and ReBoot Camp – that are designed to promote habit change within student life. Sign up at {MASCOT}Habits.org How do the programs work? The programming starts during week 1 of spring quarter, when you will set your goals. Then every week you will receive a new challenge or training drill in your email inbox. Both programs cover the same habits -- like stress coping, procrastination, communication, exercise, sleep, and more. You pick an option from the weekly instructions and try to stick to the new habit for the full week. All student participants will be attempting the same habit change, so you'll be in it together week-by-week. Why two programs? You get to choose between two versions of the same program, based on which one appeals more to you. The Happiness Challenge keeps you motivated by focusing on your happiness. ReBoot Camp has a "drill sergeant" who will keep you on track as you reboot your life. What are the perks? You'll develop better habits and learn about the science behind them. There will be tons of campus events and resources to keep you motivated and engaged throughout the quarter. Every week you complete = entry into the weekly prize drawing. Students that complete all 8 weeks will be entered for the big prize drawing (iPad mini, FitBits) and earn an official program certificate of healthy life skills training. Read more about the programs at reboot-camp.org or thehappinesschallenge.org	Habit Change at [redacted]: 8 Weeks of Goals, Activities, Events, Prizes, and Life Improvement This fall semester, there will be two programs for [redacted] students - The Happiness Challenge and ReBoot Camp – that are designed to promote habit change within student life. Sign up at {MASCOT}Habits.org How do the programs work? The programming starts during the first week of October, when you will set your goals. Then every week you will receive a new challenge or training drill in your email inbox. Both programs cover the same habits -- like stress coping, procrastination, communication, exercise, sleep, and more. You pick an option from the weekly instructions and try to stick to the new habit for the full week. All student participants will be attempting the same habit change, so you'll be in it together week-by-week. Why two programs? You get to choose between two versions of the same program, based on which one appeals more to you. The Happiness Challenge keeps you motivated by focusing on your happiness. ReBoot Camp has a "drill sergeant" who will keep you on track as you reboot your life. What are the perks? You'll develop better habits and learn about the science behind them. There will be tons of campus events and resources to keep you motivated and engaged throughout the quarter. Every week you complete = entry into the weekly prize drawing. Students that complete all 8 weeks will be entered for the big prize drawing (one-month membership at [redacted], valued at $165) and earn an official program certificate of healthy life skills training. Read more about the programs at striveweekly.com

Figure S3. Comparison of text in email ads providing information about both programs.

Campus 1:

Habit Change at [redacted]

8 Weeks of Goals, Skills, and Life Improvement

Here's the deal: This spring at [redacted], we're introducing *The Happiness Challenge* and *ReBoot Camp* to help you revamp your habits. It's all about choices. First, you get to choose between the two versions of the programs. Then, every week, both programs cover the same habits -- like stress coping, procrastination, communication, exercise, and sleep -- and you get to choose options from the weekly instructions that will move you towards your goals.

1. **Sign up** at {MASCOT}Habits.org. Challenges start **Week 1 of Spring Quarter**.
    - The Happiness Challenge: stay motivated by focusing on your happiness
    - ReBoot Camp: the drill sergeant will keep you on track to reboot your life
2. **Receive** a new challenge or training drill in your inbox each week and learn about a new aspect of health and wellness
3. **Try out the habit** of your choice week by week
4. **Check in** at the end of each week
5. **Win prizes:**
    1. **Complete a weekly check-ins to be entered for the weekly prizes** -- gift cards, food, and free stuff!
    2. **Complete all 8 weeks for a chance to WIN the biggest prizes of all: FitBits or an iPad!**
    3. **Completes all 8 weeks to receive an official certification of healthy life skills training** -- resume building, anyone?

What are the weeks?

Week 1: Goal setting & set-up
Week 2: Leisure & organization
Week 3: Ways of thinking
Week 4: Managing sleep & time
Week 5: Communicating well
Week 6: Being in the moment
Week 7: Physical exercise
Week 8: Wrap-up & planning

Two Bonus weeks:
Week 9: Insight building
Week 10: De-stress & relax

Campus 2:

Habit Change at [redacted]

8 Weeks of Goals, Skills, and Life Improvement

Here's the deal: This fall at [redacted], we're introducing *The Happiness Challenge* and *ReBoot Camp* to help you revamp your habits. It's all about choices. First, you get to choose between the two versions of the programs. Then, every week, both programs cover the same habits—like stress coping, procrastination, communication, exercise, and sleep—and you get to choose options from the weekly instructions that will move you towards your goals.

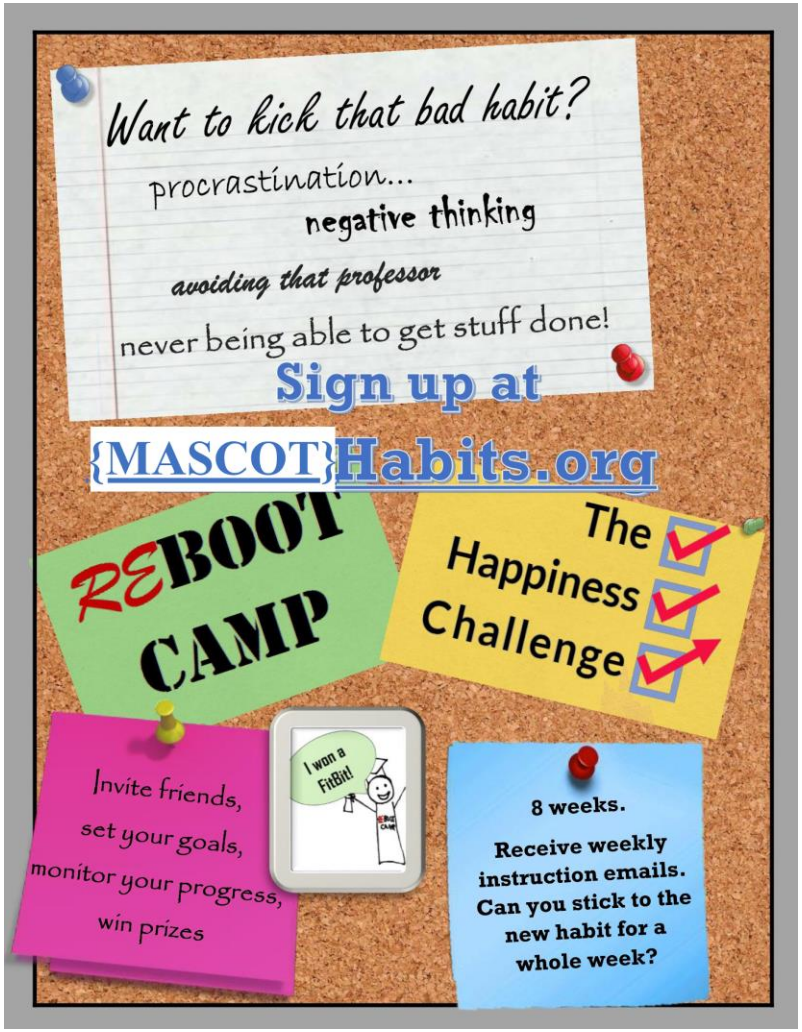
1. **Sign up at** {MASCOT}Habits.org. Challenges start week 1 of fall semester.
    - The Happiness Challenge: stay motivated by focusing on your happiness
    - Reboot Camp: training drills will keep you on track to reboot your life
2. **Receive** a new challenge or training drill in your inbox each week and learn about a new aspect of health and wellness
3. **Try out the habit** of your choice week by week
4. **Check in** at the end of each week
5. **Win prizes:**
    - [redacted] Gift Cards
    - Unlimited month pass to [redacted]
    - Chick fil-A Sandwich Passes
    - [redacted] Gift Cards

What are the weeks?

Week 1: Welcome week
Week 2: Make it happen
Week 3: Challenge your thinking
Week 4: Life troubleshooting
Week 5: Decide to say it
Week 6: Physical exercise
Week 7: Pause
Week 8: Wrap-up week

Figure S4. Comparison of ads displaying both programs and week themes.

Campus 1:



Campus 2:

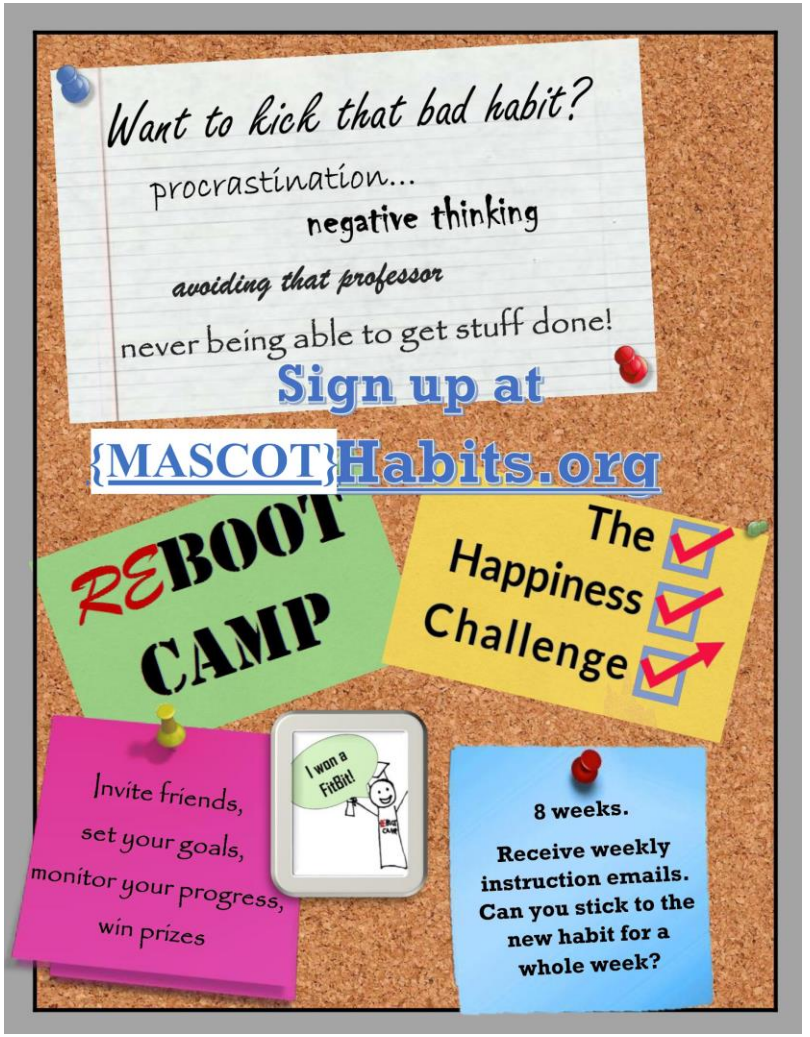


Figure S5. Comparison of attachment for email ads displaying both programs.

Appendix 2B

Table S1. Summary of all statistical analysis results comparison between Study 1 and Study 2

Statistic Reported	UCLA Number	Yale number	Study 1 Direction / Significance	Study 2 Direction / Significance
z-test of undergraduate proportion (sample vs population)			+, non-significant	**+, significant**
z-test of female proportion (sample vs population)			**+, significant**	**+, significant**
z-test of Asian proportion (sample vs population)			**+, significant**	+, non-significant
z-test of Asian proportion (sample vs population) - undergraduate only			**+, significant**	**+, significant**
z-test of American Indian or Alaskan Native proportion (sample vs population)			=, non-significant	=, non-significant
z-test of Black Non-Hispanic proportion (sample vs population)			=, non-significant	-, non-significant
z-test of Hispanic proportion (sample vs population)			+, non-significant	-, non-significant
z-test of Pacific Islander proportion (sample vs population)			=, non-significant	=, non-significant

z-test of Multiracial proportion (sample vs population)			+, non-significant	+, non-significant
z-test of White Non-Hispanic proportion (sample vs population)			-, non-significant	=, non-significant
z-test of Unknown proportion (sample vs population)			=, non-significant	not tested
z-test of International proportion (sample vs population)			-, non-significant	not tested
Mean age of undergrad students (sample vs population)			=, non-significant	not tested
Mean age of graduate students (sample vs population)			+, non-significant	not tested
% clinically significant symptoms (PHQ-2/STAI)	**60.80%**	**69.40%**	**above 50%**	**above 50%**
% clinically significant symptoms (PHQ-9/STAI)		69.40%	not tested	above 50%
% never used campus counseling center	**69.40%**	**72.30%**	**above 50%**	**above 50%**
% never used on-campus or off-campus counseling services		70%	not tested	above 50%

% never used campus counseling center of those who had elevated symptoms	**63.40%**	**66.80%**	**above 50%**	**above 50%**
% never used off-campus counseling center of those who had elevated symptoms		63.40%	not tested	above 50%
Being Asian on use of counseling center use (chi-squared)			**-, significant**	**-, significant**
Being Asian on use of any mental health service (chi-squared			not tested	**-, significant**
Gender (female) on use of counseling center (chi-squared)			+, non-significant	**+, significant**
Gender as a predictor of any services (chi-squared)			not tested	**+, significant**
Black on use of counseling center (chi-squared)			=, non-significant	=, non-significant
Black as a predictor of any services (chi-squared)			not tested	=, non-significant
White on use of counseling center (chi-squared)			+, non-significant	+, non-significant
White as a predictor of any services (chi-squared)			not tested	+, non-significant

Hispanic on use of counseling center (chi-squared)	+, non-significant	+, non-significant
Hispanic as a predictor of any services (chi-squared)	not tested	+, non-significant
Multi-racial on use of counseling center (chi-squared)	=, non-significant	=, non-significant
Multi-racial as a predictor of any services (chi-squared)	not tested	=, non-significant
Being female as a predictor of clinically elevated, non-service using students (chi-squared)	=, non-significant	-, non-significant
Being female as a predictor of clinically elevated, non-service using students (chi-squared)	not tested	-, non-significant
Being Black as a predictor of clinically elevated, non-service using students (chi-squared)	=, non-significant	=, non-significant
Being Black as a predictor of clinically elevated, non-service using students (chi-squared)	not tested	=, non-significant
Being Hispanic as a predictor of clinically elevated, non-service using students (chi-squared)	=, non-significant	-, non-significant
Being Hispanic as a predictor of clinically elevated, non-service using students (chi-squared)	not tested	-, non-significant

Being White as a predictor of clinically elevated, non-service using students (chi-squared)			-, non-significant	**-, significant**
Being White as a predictor of clinically elevated, non-service using students (chi-squared)			not tested	-, non-significant
Being Asian as a predictor of clinically elevated, non-service using students (chi-squared)			**+, significant**	**+, significant**
Being Asian as a predictor of clinically elevated, non-service using students (chi-squared)			not tested	**+, significant**
Multiracial as a predictor of clinically elevated, non-service using students (chi-squared)			-, non-significant	=, non-significant
Multiracial as a predictor of clinically elevated, non-service using students (chi-squared)			not tested	=, non-significant
% participants choosing HC	**59.90%**	**59.90%**	**above 50%**	**above 50%**
% participants choosing RC	**40.10%**	**40.10%**	**below 50%**	**below 50%**
Gender (male) on program selection (ReBoot Camp)			**+, significant**	**+, significant**
Being Black on program selection (ReBoot Camp)			=, non-significant	**+, significant**

Being Asian on program selection	-, non-significant	=, non-significant
Being White on program selection	=, non-significant	-, non-significant
Being Multiracial on program selection	=, non-significant	=, non-significant
Being Hispanic on program selection	+, non-significant	=, non-significant
Academic discipline (physical sciences + engineering) on program selection (ReBoot Camp) for males	**+, significant**	**+, significant**
Academic discipline (physical sciences + engineering) on program selection for females	+, non-significant	=, non-significant
Academic discipline (humanities) on program selection for males	-, non-significant	=, non-significant
Academic discipline (humanities) on program selection for females	+, non-significant	=, non-significant
Academic discipline (social sciences) on program selection for males	=, non-significant	+, non-significant
Academic discipline (social sciences) on program selection for females	-, non-significant	=, non-significant

Academic discipline (life sciences) on program selection for males	=, non-significant	=, non-significant
Academic discipline (life sciences) on program selection for females	**-, significant**	=, non-significant
Academic discipline (professional student) on program selection for males	=, non-significant	-, non-significant
Academic discipline (professional student) on program selection for females	=, non-significant	+, non-significant

**Note:* Key for Study 1 and 2 results direction:

= Mean or percentage of the referenced group is equal or nearly equal to that of the other group(s)

+ Mean or percentage of the referenced group is larger than that of the other group(s)

- Mean or percentage of the referenced group is smaller than that of the other group(s)

Appendix 3A

Module Design

Selection of Practices for Modules

A systematic review of practice and instructional elements in prevention programs for anxiety, depression, and stress in university students was used to inform our selection of skills and teaching strategies we should include in our program. The results were analyzed in 2015 after the initial round of article coding, so findings below are based on analyses conducted in fall 2015.

First, skills practices were identified from those common to 10% or more of universal prevention programs (regardless of delivery type) for anxiety/depression/stress in university students: relaxation (55%), cognitive monitoring/restructuring (55%), physical exercise (35%), insight building (30%), mindfulness (20%), problem solving (20%), stress management (20%), self-monitoring (15%), communication skills (15%), bio/neurofeedback (10%), and time management (10%).

Second, skills practices were identified from those common to 10% or more of online prevention programs (regardless of prevention level) for anxiety/depression/stress in university students: relaxation (64%), cognitive monitoring/restructuring (64%), insight building (43%), communication skills (43%), problem solving (29%), sleep hygiene (29%), self-monitoring (21%), mindfulness (21%), stress management (21%), activity scheduling (14%), and physical exercise (14%).

Third, skills practices were excluded based on impracticality and low inter-rater reliability and then the remaining skills practices were compiled into one list. We eliminated bio/neurofeedback, as the necessary devices would not have been easily access for all students in our online program. We eliminated stress management, as it was the only practice element with inter-rater reliability below the moderate kappa magnitude range. The final list of skills practices was: relaxation, cognitive monitoring/restructuring, insight building, self-monitoring, physical exercise, mindfulness, problem solving, communication skills, activity scheduling, time management, and sleep hygiene.

Fourth, instructional elements were identified from those common to 10% or more of universal prevention programs (regardless of delivery type) for anxiety/depression/stress in university students: psychoeducation (70%), peer engagement (15%), goal setting (15%), maintenance/relapse prevention, and services/resources awareness (10%),

Fifth, instructional elements were identified from those common to 10% or more of online prevention programs (regardless of prevention level) for anxiety/depression/stress in university students: psycheducation (93%), services/resources awareness (21%), peer engagement (14%), goal-setting (14%), motivational enhancement (14%), and performance feedback (14%).

Sixth, instructional elements were excluded based on impracticality and then the remaining instructional elements were compiled into one list. We modeling, motivational enhancement, and

performance feedback, as effective implementation requires either a live instructor or technology that can provide personalized feedback, which was beyond our means for the feasibility trial's online platform. The final list of instructional elements was: psychoeducation, peer engagement, goal-setting, maintenance/relapse prevention, and services/resource awareness.

Building the Modules

Practice skills and instructional elements were combined into modules, based on conceptual similarities and on standalone practice element frequency across existing interventions. This produced 10 modules:

1) Self-monitoring + Peer Engagement + Goal Setting → "Getting Set Up"
2) Relaxation → "De-Stress"
3) Cognitive monitoring/restructuring → "Challenge Your Thinking"
4) Insight building → "Find Your Challenge Zone"
5) Physical exercise → "Physical Exercise"
6) Mindfulness → "Pause"
7) Problem solving + Time Management + Sleep Hygiene → "Life Troubleshooting"
8) Communication Skills → "Decide How to Say It"
9) Activity Scheduling → "Make It Happen"
10) Maintenance/Relapse Prevention + Goal Setting (revisited) → "Looking Ahead"

* Weeks 2-10 all included Psychoeducation and Services/Resource Awareness

The principal investigator worked with a team of five other advanced graduate students to write the instructional content for all 10 modules. The principal investigator served as the primary author on five modules, and the five other graduate students each served as the primary author for one module. All modules were reviewed and edited by a second author.

Review of Modules Through a Focus Group

In order to select the 8 primary modules and the 2 optional modules, a 2-week version of the program was run with a 20-person focus group. The individuals in the focus group were undergraduate research assistants and graduate students from UCLA, none of whom were involved in the development of the original online platform. In the 2-week version, everyone in the focus group was instructed to practice one module daily (excluding weekends). At the end of the 2-week period, the focus group submitted anonymous feedback through an online survey and rated all modules from most favorite to least favorite. "Finding You Challenge Zone" and "De-Stress" emerged as the least popular modules overall, and therefore were selected as the optional modules, to compliment the full 8-week online mental health program.

Appendix 3B

Multiple Imputation Procedure

First, randomness of missing data was confirmed. Independent sample t-tests were run, and there were no pre-program symptom differences between post-survey respondents and those students that did not complete the post-program survey (PSS, p = .77, STAI-S, p = .27, PHQ-2, p = .28). There were, however, significantly more program completers who completed the post-survey though, with 114/119 program completers providing responses compared with 138/532 program non-initiators or partial completers. Therefore, t-tests were run only on the subsample of program non-completers: students who did not complete the program but did complete the post-program survey show no pre-program symptom differences than those students who did not complete the program and dropped out of the research study (PSS, p = .82, STAI-S, p = .43, PHQ-2, p = .54). Thus the retained students in the program non-completer group seemed to be representative.

Second, normality of data distribution was assessed and confirmed, as all skewness and kurtosis variables were below 1, as seen in this table:

						Skewness		Kurtosis	
	N	Minimum	Maximum	Mean	Std. Dev	Statistic	Std. Err	Statistic	Std. Err
Pre.STAI	563	20.00	77.00	42.3861	12.76114	.315	.103	-.642	.206
Post.STAI	226	20.00	78.00	41.4530	12.16806	.436	.162	-.112	.322
Pre.PSS	565	9	33	21.26	3.349	-.001	.103	.625	.205
Post.PSS	224	15.00	33.00	21.7391	3.24786	.527	.163	.576	.324
PHQ.2.Total	557	.00	6.00	1.6266	1.55263	.985	.104	.528	.207
PHQ.2.Total	225	.00	6.00	1.5156	1.37288	.944	.162	.809	.323
Valid N	187								

Third, individual scale item scores (instead of total scale scores) were automatically imputed with SPSS 23 using the constraints of each item's respective Likert minimum and maximum. Imputing item scores instead of total scores is preferred because, while it has no more or less "influence on the bias of scale-level parameter estimates, it had a substantial impact on efficiency, such that item-level imputation consistently produced a meaningful power advantage" (Gottschall, West, & Enders, 2012). Five imputations were run, with 150 maximum draws and 2 maximum parameter draws. Using the resulting multiply imputed item scores, total measure scores were re-totaled for each participant.

Unfortunately, the means of total scores for the measures were significantly different from the original data, so item-level imputation was abandoned. Using the same parameters, missing total scores were multiply imputed at the measure-level. The minimum and maximum for each measure total score was defined based on the original sample's real minimum and maximum total scores on the PSS, PHQ2, and STAI, rather than using the minimum and maximum possible for the measures. This time, the means of total scores for the measures were not significantly different from the original data, as seen below. Therefore, the multiply imputed data sample was based on measure-level total score imputation.

Variable	Set	*N*	Mean	*F*	Sig.
Pre-Program STAI-S	Imputed	593	42.5224	0.001201	0.972366
	Original	563	42.3861		
Post-Program STAI-S	Imputed	593	42.4147	2.003512	0.157316
	Original	226	41.4530		
Pre-Program PSS	Imputed	593	21.28	0.014783	0.903249
	Original	560	21.28		
Post-Program PSS	Imputed	593	21.7037	0.048246	0.826199
	Original	224	21.7391		
Pre-Program PHQ-2	Imputed	593	1.6547	0.047424	0.827647
	Original	555	1.6306		
Post-Program PHQ.2	Imputed	593	1.5457	1.90672	0.167707
	Original	225	1.5156		

Appendix 3C

Qualitative Themes from Post-program Feedback Survey and Corresponding Examples

Theme	Frequency mentioned		Example responses
Reminders	79	69.6% had positive feedback, while the remaining 30.4% suggested changes (e.g., increased frequency, decreased frequency, different days of the week).	I like how we got reminder emails and specific suggestions on how to complete the challenges. I had a hard time remembering to do each challenge throughout the week. Text reminders would have been very helpful Sometimes I would forget to do the challenge and the second email later on in the week would remind me to either check in or to do the challenge.
Self-Motivation/ Accountability	66	25.8% felt the intervention design was still sufficient to maintain program motivation, while 74.2% expressed difficulty (e.g., no program features that emphasize accountability, easy to forget the program when busy).	I liked the consistency of the weekly emails and specific, achievable challenges. It can be overwhelming to self-motivate, and the structure and organization of the program helped me meet my goals. There's not much motivation to actually adhere to each week's goals, which was really disappointing to me. I feel like, at least for me, external motivation helps a lot when trying to get someone to change habits that are so internally ingrained.
Tips/Resources/ Suggestions/Extras	56	89.2% found these extras to be helpful, while the remaining	The tips and resources were helpful because it laid out some suggestions for how to complete the week and provided

		10.8% generally requested more volume or variety (as opposed to disliking the "extras" section).	some of the research behind it. I liked the extras worksheets and readings. The more I understand about what I can improve on, the more I feel in control of my life and my mood.
Challenge Difficulty	52	Of the 52 participants commenting on "challenge difficulty," 48.1% expressed finding the challenges to be "doable," "realistic," and to have "clear" explanations, while 38.4% expressed finding the challenges to be too "difficult," "time-consuming," "unspecific," as well as "too specific,", whereas the remaining 13.5% expressed finding the challenges too easy.	I liked it because it wasn't too demanding but gave nice helpful tips. I liked the weekly email reminders and calendar settings After a couple of weeks, I felt like the program was trying to change too many things, and I couldn't keep up. Eventually, I ended up ignoring the emails since I had a lot of other responsibilities.
Goal-Oriented/Goal-Setting	39	All responses mentioned the "goal-oriented" nature of the program (i.e., goal-setting was the first activity offered in the programming sequence) as a strength of the design, but some wanted more guidance in goal selection at the beginning or help remembering their goal later in the program.	I believe the program worked from be because by setting the goal at the beginning and having weekly reminders I felt obligated complete my goal because I had already in a way formally committed myself to complete the goal by stating it on the first survey
Challenge Flexibility/	37	56.8% discussed it as a strength (e.g., variety allowed	Each challenge was like a boost of attention into an important aspect of my life and I thought that was really

Customization		for personalization, multiple options kept it interesting), whereas 43.2% reported feeling unable to relate the presented skills options to their needs.	great. Also, I really liked the flexibility of the program. I managed to think about or do something for the challenges every week, even though I was really busy, and that made me feel good. giving a single more specific goal for the week while also allowing for subsequent secondary goals to be chosen freely might be advantageous. I find it helps to be told what to do on one front, and be able to choose what to do around that main goal.
Delivery Platform	37	48.6% emphasized the simplicity of getting content through emails and a website, whereas the remaining 51.4% had specific suggestions (e.g., developing an app, delivering content through short videos, incorporating social media).	A simple acted out video between two students or by a single student would work. Visual example would help ingrain the idea even better in my opinion. I think that following the instructions would be easier if there were an interactive calendar component incorporated into an app.
Website design (Interactivity, Design/Language/ technical issues)	34	20.6% reported that website had a "friendly tone," was "visually appealing," and had "user friendly graphics," whereas 41.2% wanted the website to be "more interactive," and the remaining 38.2% students had specific design or technical issues (e.g., "it was not clear where to check in," "email colors were not appealing," issues with	I loved the user friendly graphics on the site, it made it appealing to read. It'd be nice if the emails and the online interface was smoother and had a nicer UI.

		email delivery to inboxes)	
Accessibility, Self-Guided, Pacing, Convenience	29	96.7% found it to be a strength (e.g., “it was online, which was very convenient,” “it worked for me because it was based off of my own time and convenience”), while one student found the self-guided nature to be too demanding for planning purposes.	The program allows me to challenge myself to organize my time to do other activities that are beneficial to me meanwhile focusing on finishing my school works. It worked for me because it was based off of my own time and convenience. I liked that it was online, which was very convenient, and that it was very goal-oriented and reminded to keep working toward weekly goals to get to my main goals.
8-Week Sequential Module Format	26	73.1% mentioned it only in the context of being a strength of the program’s design, whereas 26.9% recommended changes to the ordering or number of modules. There were no complaints about the module design itself.	I like the reminders and the structure of having a weekly challenge! I also like that I had to check in. I liked how the program sent me weekly emails. I looked forward to receiving them and discovering new ways to improve my life. The readings were all really interesting. The reminder emails really helped as well.
Peer Facilitation	26	All of such responses indicated a desire for peer interaction to be facilitated through the program somehow (e.g., meetups, Instagram hashtags, indicators on the platform of other participants’ progress).	I guess I sort of wished to have another person to talk about the experience (a buddy during the program, if possible). Most of my friends weren't really interested. It'd be nice if the program could pair up students, but maybe that would be too complicated to do. I didn't like that I didn't know other people who were doing it as well. A social community of sorts to accompany this challenge would've been lovely. And to be honest people love to compete against each other and it often pushes us harder and further so that would've been cool.

Personalized Progress Tracking	22	All of such responses requested that the platform somehow track and indicate their ongoing program activity (e.g., “visually monitor progress,” “revisit previous check-in responses,” “a checklist interface for weekly skills”)	The one important thing that I think would help would be to have a page somewhere where I could see my goal and each of the check-ins I had completed, just to visually monitor my progress. I think I would have done better with the self-reporting if there were a platform involved with the Happiness Challenge for me to do so, maybe as part of the end-of-week check-ins (though maybe this was available and I missed it)
Incentives/rewards	11	64% identified the prize drawing as a main reason for motivation, and the remaining 36% said they wanted bigger and/or guaranteed rewards.	I don't really know how the motivation aspect of the program could be improved: maybe with a greater number of weekly prizes, so people believe they might actually have a shot of winning something? Too hard to keep up; not enough incentives

Appendix 3D

All Eight Versions of RM-MANOVAs Testing Outcome Change Detectability

RM-MANOVA Version	Outcome	*F*	*p*
Including students with additional service use			
Completer Sample			
	Time		
	Multivariate Test	3.46	.02
	STAI-S	1.35	.25
	PHQ-2	5.20	.02
	PSS	4.45	.04
	Time X Completer Status		
	Multivariate Test	3.44	.02
	STAI-S	7.61	.006
	PHQ-2	6.30	.01
	PSS	0.28	.60
Intent-to-Treat Sample			
	Time		
	Multivariate Test	10.30	< .001
	STAI-S	11.77	.001
	PHQ-2	19.24	< .001
	PSS	5.06	.03
	Time X Completer Status		
	Multivariate Test	10.38	< .001
	STAI-S	17.61	< .001
	PHQ-2	21.76	< .001
	PSS	0.54	.46
Multiple Imputation of Post-Survey Data			
	Time		
	Multivariate Test	7.62	< .001
	STAI-S	4.89	.03
	PHQ-2	13.68	< .001
	PSS	5.94	.02
	Time X Completer Status		
	Multivariate Test	5.84	.001
	STAI-S	9.51	.002
	PHQ-2	11.28	.002
	PSS	0.13	.72
Multiple Imputation of Pre- and Post-Survey Data			
	Time		

	Multivariate Test	5.64	.001
	STAI-S	3.05	.08
	PHQ-2	11.97	.001
	PSS	3.33	.07
Time X Completer Status			
	Multivariate Test	4.52	.004
	STAI-S	6.97	.009
	PHQ-2	9.29	.002
	PSS	0.03	.86
Excluding students with additional service use			
Completer Sample			
Time			
	Multivariate Test	2.94	.04
	STAI-S	1.12	.29
	PHQ-2	6.02	.02
	PSS	2.32	.13
Time X Completer Status			
	Multivariate Test	2.52	.06
	STAI-S	4.62	.03
	PHQ-2	5.18	.02
	PSS	0.000	.99
Intent-to-Treat Sample			
Time			
	Multivariate Test	10.24	< .001
	STAI-S	9.48	.002
	PHQ-2	21.69	< .001
	PSS	3.94	.05
Time X Completer Status			
	Multivariate Test	10.01	< .001
	STAI-S	12.59	< .001
	PHQ-2	22.51	< .001
	PSS	1.18	.28
Multiple Imputation of Post-Survey Data			
Time			
	Multivariate Test	7.62	< .001
	STAI-S	4.89	.03
	PHQ-2	13.68	< .001
	PSS	5.94	.02
Time X Completer Status			
	Multivariate Test	5.84	.001
	STAI-S	9.51	.002
	PHQ-2	11.28	.002

	PSS	0.13	.72
Multiple Imputation of Pre- and Post-Survey Data			
	Time		
	Multivariate Test	5.64	.001
	STAI-S	3.05	.08
	PHQ-2	11.97	.001
	PSS	3.33	.07
	Time X Completer Status		
	Multivariate Test	4.52	.004
	STAI-S	6.99	.009
	PHQ-2	9.29	.002
	PSS	0.03	.86

Appendix 4A

Intervention Adaption Between the Open Trial and Randomized Controlled Trial

Adaptions made to the online intervention platform were based off of common themes from the qualitative feedback collected in the open trial. Primary adaptions based on feedback related to the following common qualitative themes:

1. **More interactivity**. One in ten feedback providers indicated desire for more interactivity through the online platform. Moreover, many other feedback themes necessitated more interactivity (e.g., customized progress tracking).
2. **Reminders.** Participant feedback suggested that in general email reminders are great, but there were variable requests for more/less frequency and specific timing.
3. **Motivation. Although** accessibility benefits were mentioned, a major negative trade-off noted were the difficulties with self-motivation and self-accountability.
4. **Progress tracking.** Participants requested more goal-oriented and progress-monitoring features for duration of program.
5. **Practice difficulty.** Participant feedback ranged from saying that module skills practice was: too easy, just right, or too hard.
6. **Social connection.** Some participants described a lacking sense of community and a desire for more peer support. Many did not comment on this theme, so too much adaption would likely be poorly received, but those that did comment on this theme felt strongly.

	Previous Design	**Considerations**	**Design Solution**
More interactivity	Static website + email list		Need for platform to now support individual user accounts
Reminders	Pre-programed single reminder on Friday	Variable preferences according to participant feedback	Customizable account settings
Motivation	Motivation not assessed		Measure at baseline to assess moderation
Progress tracking	Goal-setting activity at beginning of program, encouraged self-monitoring off-line	What features are possible now that there are individual user accounts	1. Goal-setting included as part of account set-up process; 2. Goal plan is clickable on dashboard for duration of intervention; 3. All self-monitoring

			features that are now incorporated into dashboard (i.e., log history, stress and mood rating graph)
Practice difficulty	Provided list of activities and instructed to practice three. One-size-fits-all	Variable preferences according to participant feedback	Instructions allow participants to choose how many activities to practice each week. Introduction of the Medals system (1-2 activities = bronze; 3-4 activities = silver; 5+ activity = gold)
Social connection	In-person weekly events that were poorly attended	How to facilitate sense of social connection online while (a) protecting between-user anonymity and (b) preventing iatrogenic user-generated content	Campus-specific livestream with auto-pushed anonymous user activity

Appendix 4B

Adapted Treatment Motivation Questionnaire

Instructions: *Right now you are planning to sign-up for either The Happiness Challenge or ReBoot Camp online program. This questionnaire concerns student's reasons for signing-up and their feelings about the online programming. Different people have different reasons, and we want to know how true each of these reasons is for you.*

Please indicate how true each reason is for you, using the following scale:

1	*2*	*3*	*4*	*5*	*6*	*7*
not at all true			*somewhat true*			*very true*

I'm signing up for this program because:

1. I really want to make some changes in my life
2. I won't feel good about myself if I don't get some help.
3. I was referred by an advisor, boss, or counselor.
4. I feel so guilty about my problems that I have to do something about it.
5. It is important to me personally to work on my problems.

If I remain in the online program it will probably be because:

6. I'll get in trouble if I don't.
7. I'll feel very bad about myself if I don't.
8. I'll feel like a failure if I don't.
9. I feel like it's the best way to help myself.
10. I don't really feel like I have a choice about staying in the program.
11. I feel it is in my best interests to complete the program.

Rate each of the following in terms of how true each statement is for you.

12. I signed up for this program now because I was under pressure to get help.
13. I am not sure this program will work for me.
14. I am confident this program will work for me.
15. I decided to sign up for this program because I was interested in getting help.
16. I'm not convinced that this program will help me change my habits.
17. I am responsible for choosing to sign-up.
18. I doubt that this program will solve my problems.
19. I chose this program because I think it is an opportunity for change.
20. I am not very confident that I will get results from this program.

Subscales

External Reasons: 3, 6, 10, 12
Internal Reasons: 1, 2, 4, 5, 7, 8, 9, 11, 15, 17, 19
Confidence in Treatment: 13(R), 14, 16(R), 18(R), 20(R)

Appendix 4C

Program Satisfaction Questionnaire

Take a moment to consider your responses for each of the following questions. The more informative and honest your responses are, the better able we will be to improve the program in the future.

Q39. How would you rate the quality of the StriveWeekly platform and program content?

4 – excellent
3 – good
2 – fair
1 – poor
0 – really bad

Q40. If a friend wanted to learn new skills or form better habits, would you recommend StriveWeekly to him or her?

0 – no, definitely not
1 – no, I don't think so
2 – I'm not sure
3 – yes, I think so
4 – yes, definitely

Q41. Did this program help you make progress towards your goal?

4 – yes, definitely
3 – yes, I think so
2 – I'm not sure
1 – no, I don't think so
0 – no, definitely not

Q42. Did this program help you cope more effectively with stress, anxiety, and/or depressed mood?

0 – no, definitely not
1 – no, I don't think so
2 – I'm not sure
3 – yes, I think so
4 – yes, definitely

Q43. Overall, how satisfied are you with The Happiness Challenge or ReBoot Camp?

4 – Very satisfied
3 – Mostly satisfied
2 – Indifferent
1 – Mildly dissatisfied
0 – Very dissatisfied

Appendix 4D

Rationale for Covariate Selection Approach

We used this covariate selection approach as the most important baseline covariate variables are those that are related to treatment outcome, regardless of if they are imbalanced in condition (Pocock, Assmann, Enos, & Kasten, 2002; Raab, Day, & Sales, 2000). The concern that covariate-adjusted models will bias treatment effect estimates is lower with large trials, whereas the advantages are that inclusion of prognostic covariates will produce a treatment effect estimate that has more accuracy in terms of: (a) p-value, (b) magnitude, and (c) precision – and thus statistical power (Kahan, Jairath, Dore, & Morris, 2014; Pocock et al., 2002). Although it has been recommended that such covariates are selected based on past trials rather than during the analysis phase of trails (e.g., Raab et al., 2000), others (e.g., Pocock et al., 2002) have acknowledged that this approach may not be practical, as predefined covariates may not end up related to treatment outcome in a new trial. Instead, a predefined statistical strategy for covariate selection is more feasible while still alleviating concerns of otherwise subjective post hoc covariate selection (Kahan et al., 2014; Pocock et al., 2002). Indeed, simulation studies based on real trial data have shown that power is increased by including prognostic covariates, and thus the benefit of their inclusion in large trials outweighs the risk of inflation of type I error, which is less likely given moderate or large samples (Kahan et al., 2014).

Still, the CONSORT statement recommended reporting both the unadjusted average and the adjusted treatment effects (Schulz, Altman, Moher, & CONSORT Group, 2010). Therefore, the reported findings in the results section are based on adjusted analyses, but unadjusted analyses are also provided below in Table 4D.1. The unadjusted models are all still significant, but the overall portion of variance explained is smaller, further justifying covariate inclusion.

Table 4D.1. Mixed linear effects model without covariate predictors.

	Model	**Interaction term**	
	R^2	***t***	***p***
Intervention Effects (TimeA*Group)			
Depression	.01	-2.56	.01
Anxiety	.02	-2.73	.007
Stress	.01	-2.68	.007
Replication Effects (TimeB*Group)			
Depression	.01	3.45	< .001
Anxiety	.02	3.28	.001
Stress	.01	2.70	.007

Appendix 4E

Mixed linear effects results for main model and robustness check models

	Model R^2	Interaction term t	p
Intervention Effects (TimeA*Group)			
Depression			
Main Model	.15	-3.05	.002
Robustness A	*.11*	*-2.38*	*.02*
Robustness B	*.10*	*-1.98*	*.05*
Robustness C	*.14*	*-3.51*	*< .001*
Anxiety			
Main Model	.08	-3.01	.003
Robustness A	*.06*	*-2.76*	*.006*
Robustness B	*.06*	*-2.42*	*.02*
Robustness C	*.07*	*-3.60*	*< .001*
Stress			
Main Model	.07	-2.92	.004
Robustness A	*.05*	*-2.22*	*.03*
Robustness B	*.05*	*-0.98*	*.33*
Robustness C	*.07*	*-3.29*	*.001*
Replication Effects (TimeB*Group)			
Depression			
Main Model	.15	3.48	< .001
Robustness A	*.11*	*2.92*	*.004*
Robustness B	*.10*	*2.20*	*.03*
Robustness C	*.14*	*3.27*	*.001*
Anxiety			
Main Model	.08	3.24	.001
Robustness A	*.06*	*2.20*	*.03*
Robustness B	*.06*	*1.96*	*.05*
Robustness C	*.07*	*3.20*	*.001*
Stress			
Main Model	.07	2.69	.007
Robustness A	*.05*	*2.15*	*.03*
Robustness B	*.05*	*1.62*	*.11*
Robustness C	*.07*	*3.07*	*.002*

TMQ Moderation Effects (TimeA*Group*TMQ)			
Depression			
Main Model	.20	-2.67	.008
Robustness A	.19	-2.75	.006
Robustness B	.17	-2.27	.02
Robustness C	.20	-2.43	.02
Anxiety			
Main Model	.12	-2.69	.007
Robustness A	.12	-2.76	.006
Robustness B	.11	-3.02	.003
Robustness C	.13	-3.28	.001
Stress			
Main Model	.12	-1.78	.07
Robustness A	.11	-1.78	.08
Robustness B	.11	-1.46	.15
Robustness C	.12	-1.05	.29
GRIT Moderation Effects (TimeA*Group*GRIT)			
Depression			
Main Model	.19	1.39	.17
Anxiety			
Main Model	.09	0.86	.39
Stress			
Main Model	.08	-0.64	.52

Sample Sizes Key

Main Model, n = 1607

Robustness A, n = 1367

Robustness B, n = 661

Robustness C, n = 9

Appendix 4F

Effect Sizes Calculated from Raw Means and from Estimate Marginal Means (EMMs)

Comparing results of effect sizes from raw means (Table 4F.1) versus from EMMs (Table 4F.1), bigger effect sizes are produced by EMMs. Given that EMMs are adjusted for differences by covariates, and that research dropout status differed by the gender covariate ($p < .001$), the effect sizes from the EMMs are more for interpretation of estimated intervention effects.

Table 4F.1. Effect Sizes (Raw Means) by Condition

		pre *M* (*se*)	post *M* (*se*)	FU *M* (*se*)	Pre-Post *d*	Post-FU *d*
Intervention, *n* = 804						
	Dep	5.45 (0.16)	4.58 (0.20)	4.64 (0.24)	0.19	-0.02
	Anx	4.72 (0.13)	3.82 (0.17)	3.35 (0.20)	0.25	0.14
	Str	7.28 (0.15)	6.58 (0.20)	6.02 (0.25)	0.17	0.13
Waitlist, *n* = 803						
	Dep	5.41 (0.16)	5.05 (0.20)	3.94 (0.22)	0.08	0.27
	Anx	4.63 (0.13)	4.28 (0.16)	2.95 (0.18)	0.09	0.38
	Str	7.27 (0.15)	7.30 (0.20)	5.73 (0.23)	-0.01	0.37
Intervention - Waitlist						
	Dep				0.11	-0.29
	Anx				0.16	-0.24
	Str				0.18	-0.24

Table 4F.2. Effect Sizes (Estimated Marginal Means) by Condition

		pre *M* (*se*)	post *M* (*se*)	FU *M* (*se*)	Post-Pre *d*	FU-post *d*
Intervention, *n* = 804						
	Dep	5.51 (0.14)	4.61 (0.18)	4.65 (0.21)	0.23	-0.01
	Anx	4.76 (0.13)	3.86 (0.16)	3.5 (0.17)	0.26	0.11
	Str	7.33 (0.14)	6.69 (0.19)	6.14 (0.22)	0.16	0.14
Waitlist, *n* = 803						
	Dep	5.34 (0.14)	5.16 (0.17)	4.14 (0.21)	0.05	0.27
	Anx	4.61 (0.13)	4.32 (0.15)	3.13 (0.17)	0.08	0.37
	Str	7.26 (0.14)	7.31 (0.18)	5.85 (0.23)	-0.01	0.37
Intervention - Waitlist						
	Dep				0.18	-0.28
	Anx				0.18	-0.26
	Str				0.17	-0.23

References

ACHA. (2016). *The American College Health Association National College Health Assessment (ACHA-NCHA), Spring 2016 Reference Group report.* Hanover, MD. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15813230

AL-Asadi, A. M., Klein, B., & Meyer, D. (2014). Pretreatment attrition and formal withdrawal during treatment and their predictors: An exploratory study of the anxiety online data. *Journal of Medical Internet Research, 16,* e152. https://doi.org/10.2196/jmir.2989

Alfonsson, S., Olsson, E., & Hursti, T. (2016). Motivation and treatment credibility predicts dropout, treatment adherence, and clinical outcomes in an internet-based cognitive behavioral relaxation program: A randomized controlled trial. *Journal of Medical Internet Research*, *18*. https://doi.org/10.2196/jmir.5352

Allart-Van Dam, E., Hosman, C. M. H., & Keijsers, G. P. J. (2004). A new instrument to assess participant motivation for involvement in preventive interventions. *Journal of Clinical Psychology*, *60,* 555-565. https://doi.org/10.1002/jclp.10236

American College Health Association. (2018). *American College Health Association-National College Health Assessment II: Reference Group Executive Summary Fall 2017*. Hanover, MD. Retrieved from http://www.acha-ncha.org/docs/NCHA-II_FALL_2017_REFERENCE_GROUP_EXECUTIVE_SUMMARY.pdf

Andreasen, A. R. (2004). A social marketing approach to changing mental health practices directed at youth and adolescents. *Health Marketing Quarterly*, *21*, 51–75. https://doi.org/10.1300/J026v21n04_04

Anshel, M. (1996). (1996). Effect of chronic aerobic exercise and progressive relaxation on motor performance and affect. *Behavioral Medicine*, *21*, 186–205.

https://doi.org/10.1080/08964289.1996.9933757

Arch, J. J., & Craske, M. G. (2008). Acceptance and commitment therapy and cognitive behavioral therapy for anxiety disorders: Different treatments, similar mechanisms? *Clinical Psychology: Science and Practice, 15*, 263–279. https://doi.org/10.1111/j.1468-2850.2008.00137.x

Arnett, J. (2000). Emerging adulthood: a theory of development from the later teens through the twenties. *American Psychologist*, *55*, 469–480. http://dx.doi.org/10.1037/0003-066X.55.5.469

Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J. A., Fiedler, K., … Wicherts, J. M. (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality*, *27*, 108–119. https://doi.org/10.1002/per.1919

Auerbach, R. P., Alonso, J., Axinn, W. G., Cuijpers, P., Ebert, D. D., Green, J. G., … Bruffaerts, R. (2016). Mental disorders among college students in the World Health Organization World Mental Health Surveys. *Psychological Medicine*, *46*, 2955–2970. https://doi.org/10.1017/S0033291716001665

Auerbach, R. P., Mortier, P., Bruffaerts, R., Alonso, J., Benjet, C., Cuijpers, P., … Kessler, R. C. (2018). WHO World Mental Health Surveys International College Student Project: Prevalence and Distribution of Mental Disorders. *Journal of Abnormal Psychology*, *127*, 623–638. https://doi.org/10.1037/abn0000362

Bahlmann, J., Angermeyer, M. C., & Schomerus, G. (2013). „Burnout' statt „Depression'—Eine Strategie zur Vermeidung von Stigma? / Calling it "burnout" instead of 'depression'—A strategy to avoid stigma? *Psychiatrische Praxis*, *40*, 78–82. https://doi.org/10.1055/s-0032-1332891

Bamber, M. D., & Kraenzle Schneider, J. (2016). Mindfulness-based meditation to decrease stress and anxiety in college students: A narrative synthesis of the research. *Educational Research Review*, *18*, 1–32. https://doi.org/10.1016/j.edurev.2015.12.004

Banerjee, M. (1999). Beyond kappa: A review of interrater agreement measures. *Canadian Journal of Statistics-Revue Canadienne De Statistique*, *27*, 3–23. https://doi.org/10.2307/3315487

Barlow, D. H., Allen, L. B., & Choate, M. L. (2004). Toward a unified treatment for emotional disorders. *Behavior Therapy*, *35*, 205-230. https://doi.org/10.1016/S0005-7894(04)80036-4

Becker, K., Boustani, M., Gellatly, R., & Chorpita, B. F. (2018). Forty years of engagement research in children's mental health services: Multidimensional measurement and practice elements. *Journal of Clinical Child and Adolescent Psychology*, *47*, 1–23. https://doi.org/10.1080/15374416.2017.1326121

Becker, K. D., Buckingham, S. L., Rith-Najarian, L., & Kline, E. R. (2016). The Common Elements of treatment engagement for clinically high-risk youth and youth with first-episode psychosis. *Early Intervention in Psychiatry*, *10*, 455–467. https://doi.org/10.1111/eip.12283

Becker, S. J. (2015). Direct-to-consumer marketing: A complementary approach to traditional dissemination and implementation efforts for mental health and substance abuse interventions. *Clinical Psychology: Science and Practice*, *22*, 85–100. https://doi.org/10.1111/cpsp.12086

Beiter, R., Nash, R., McCrady, M., Rhoades, D., Linscomb, M., Clarahan, M., & Sammut, S. (2015). The prevalence and correlates of depression, anxiety, and stress in a sample of college students. *Journal of Affective Disorders*, *173*, 90–96.

https://doi.org/10.1016/j.jad.2014.10.054

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B, 53,* 289-300. https://doi.org/10.2307/2346101

Bianchi, R., Verkuilen, J., Brisson, R., Schonfeld, I. S., & Laurent, E. (2016). Burnout and depression: Label-related stigma, help-seeking, and syndrome overlap. *Psychiatry Research*, *245*, 91–98. https://doi.org/10.1016/j.psychres.2016.08.025

Blanco, C., Okuda, M., Wright, C., Hasin, D. S., Grant, B. F., Liu, S. M., & Olfson, M. (2008). Mental health of college students and their non-college-attending peers: results from the National Epidemiologic Study on Alcohol and Related Conditions. *Archives of General Psychiatry*, *65*, 1429–1437. https://doi.org/10.1001/archpsyc.65.12.1429

Bliese, P. (2016). multilevel: Multilevel Functions. Retrieved from https://cran.r-project.org/web/packages/multilevel/index.html

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2013). Comprehensive Meta-Analysis Version 3. Englewood, NJ: Biostat.

Boustani, M. M., Frazier, S. L., Becker, K. D., Bechor, M., Dinizulu, S. M., Hedemann, E. R., … Pasalich, D. S. (2015). Common elements of adolescent prevention programs: minimizing burden while maximizing reach. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*, 209–219. https://doi.org/10.1007/s10488-014-0541-9

Braithwaite, S. R., & Fincham, F. D. (2007). ePREP: Computer Based Prevention of Relationship Dysfunction, Depression and Anxiety. *Journal of Social and Clinical Psychology*, *26*, 609–622. https://doi.org/10.1521/jscp.2007.26.5.609

Braithwaite, S. R., & Fincham, F. D. (2009). A randomized clinical trial of a computer based

preventive intervention: replication and extension of ePREP. *Journal of Family Psychology, 23*, 32–38. https://doi.org/10.1037/a0014061

Brown, M. T., & Chambers, M. (1986). Student and faculty perceptions of counseling centers. What's in a name? *Journal of Counseling Psychology, 33*, 155–158. https://doi.org/10.1037/0022-0167.33.2.155

Buchanan, J. L. (2012). Prevention of Depression in the College Student Population: A Review of the Literature. *Archives of Psychiatric Nursing, 26*, 21–42. https://doi.org/10.1016/j.apnu.2011.03.003

Burns, J. M., Davenport, T. A., Durkin, L. A., Luscombe, G. M., & Hickie, I. B. (2010). The internet as a setting for mental health service utilisation by young people. *The Medical Journal of Australia*, *192*, S22-26. https://doi.org/bur11029_fm [pii]

Calear, A. L., & Christensen, H. (2010). Review of internet-based prevention and treatment programs for anxiety and depression in children and adolescents. *Medical Journal of Australia*, *192*, S12-4. https://doi.org/cal10829_fm [pii]

Camacho, Á., Cordero, E. D., & Perkins, T. (2016). Psychometric Properties of the DASS-21 Among Latina/o College Students by the US-Mexico Border. *Journal of Immigrant and Minority Health, 18*, 1017–1023. https://doi.org/10.1007/s10903-016-0415-1

Cavanagh, K., Strauss, C., Cicconi, F., Griffiths, N., Wyper, A., & Jones, F. (2013). A randomised controlled trial of a brief online mindfulness-based intervention. *Behaviour Research and Therapy*, *51*, 573–578. https://doi.org/10.1016/j.brat.2013.06.003

CDS Advisory Board. (2018). Common Data Set Initiative. Retrieved March 10, 2018, from http://www.commondataset.org/

Chang, C. (2008). Increasing mental health literacy via narrative advertising. *Journal of Health*

Communication, *13*, 37–55. https://doi.org/10.1080/10810730701807027

Cheung, J. H., Burns, D. K., Sinclair, R. R., & Sliter, M. (2016). Amazon Mechanical Turk in organizational psychology: an evaluation and practical recommendations. *Journal of Business and Psychology*, *32*, 1–15. https://doi.org/10.1007/s10869-016-9458-5

Chiauzzi, E., Brevard, J., Thum, C., Decembrele, S., & Lord, S. (2008). MyStudentBody-Stress: an online stress management intervention for college students. *Journal of Health Communication*, *13*, 555–572. https://doi.org/10.1080/10810730802281668

Chorpita, B. F., & Daleiden, E. L. (2018). Coordinated strategic action: Aspiring to wisdom in mental health service systems. *Clinical Psychology: Science and Practice*, *25*, 1-14. https://doi.org/10.1111/cpsp.12264

Chorpita, B., & Daleiden, E. (2014). Structuring the collaboration of science and service in pursuit of a shared vision. *Journal of Clinical Child & Adolescent Psychology*, *43*, 323–338. http://dx.doi.org/10.1080/15374416.2013.82829.

Chorpita, B. F., Becker, K. D., & Daleiden, E. L. (2007). Understanding the common elements of evidence-based practice: Misconceptions and clinical examples. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*, 647–652. https://doi.org/10.1097/chi.0b013e318033ff71

Chorpita, B. F., & Daleiden, E. L. (2009). Mapping evidence-based treatments for children and adolescents: application of the distillation and matching model to 615 treatments from 322 randomized trials. *Journal of Consulting and Clinical Psychology*, *77*, 566–579. http://dx.doi.org/10.1037/a0014565

Chorpita, B. F., Daleiden, E. L., & Weisz, J. R. (2005a). Identifying and selecting the common elements of evidence based interventions: A distillation and matching model. *Mental Health*

Services Research, 7, 5–20. https://doi.org/10.1007/s11020-005-1962-6

Chorpita, B. F., Daleiden, E. L., & Weisz, J. R. (2005b). Modularity in the design and application of therapeutic interventions. Applied and Preventive Psychiatry: Current scientic perspectives. *Applied and Preventive Psychology*, *11*, 141–156. https://doi.org/http://dx.doi.org/10.1016/j.appsy.2005.05.002

Chorpita, B. F., Rotheram-Borus, M. J., Daleiden, E. L., Bernstein, A., Cromley, T., Swendeman, D., & Regan, J. (2011). The old solutions are the new problem: How do we better use what we already know about reducing the burden of mental illness? *Perspectives on Psychological Science*, *6*, 493–497. https://doi.org/10.1177/1745691611418240

Christensen, H., Griffiths, K. M., & Farrer, L. (2009). Adherence in internet interventions for anxiety and depression. *Journal of Medical Internet Research*, *11*. https://doi.org/10.2196/jmir.1194

Christensen, H., Pallister, E., Smale, S., Hickie, I. B., & Calear, A. L. (2010). Community-based prevention programs for anxiety and depression in youth: a systematic review. *The Journal of Primary Prevention*, *31*, 139–170. https://doi.org/10.1007/s10935-010-0214-8

Clarke, A. M., Kuosmanen, T., & Barry, M. M. (2014). A Systematic review of online youth mental health promotion and prevention interventions. *Journal of Youth and Adolescence, 44*, 90–113. https://doi.org/10.1007/s10964-014-0165-0

Clement, S., Schauman, O., Graham, T., Maggioni, F., Evans-Lacko, S., Bezborodovs, N., … Thornicroft, G. (2015). What is the impact of mental health-related stigma on help-seeking? A systematic review of quantitative and qualitative studies. *Psychological Medicine, 45*, 11–27. https://doi.org/10.1017/S0033291714000129

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress.

Journal of Health and Social Behavior, *24*, 385–396. https://doi.org/10.2307/2136404

Conley, C. S., Durlak, J. A., & Kirsch, A. C. (2015). A Meta-analysis of universal mental health prevention programs for higher education students. *Prevention Science, 16*, 487–507. https://doi.org/10.1007/s11121-015-0543-1

Conley, C. S., Kirsch, A. C., Dickson, D. A., & Bryant, F. B. (2014). Negotiating the Transition to College. *Emerging Adulthood*, *2*, 195–210. https://doi.org/10.1177/2167696814521808

Copeland, V. C., Koeske, G., & Greeno, C. G. (2004). Child and mother client satisfaction questionnaire scores regarding mental health services: Race, age, and gender correlates. *Research on Social Work Practice*, *14*, 434-442. https://doi.org/10.1177/1049731504265839

Craske, M. G. (2012). Transdiagnostic treatment for anxiety and depression. *Depression and Anxiety*, *29*, 749–753. https://doi.org/10.1002/da.21992

Cuijpers, P. (2018). The Dodo Bird in the Digital Age: How E-Mental Health Can Improve Prevention and Treatment of Mental Health Problems. In *52nd Annual ABCT Convention*. Washington, D.C.

Cuijpers, P., Straten, A. Van, Smit, F., Mihalopoulos, C., & Beekman, A. (2008). Preventing the onset of depressive disorders : A Meta-analytic review of psychological interventions. *American Journal of Psychiatry*, *165*, 1272–1280. https://doi.org/10.1176/appi.ajp.2008.07091422

Cukrowicz, K. C., & Joiner, T. E. (2007). Computer-based intervention for anxious and depressive symptoms in a non-clinical population. *Cognitive Therapy and Research*, *31*, 677–693. https://doi.org/10.1007/s10608-006-9094-x

Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data.

Journal of Experimental Social Psychology, 66, 4–19. https://doi.org/10.1016/j.jesp.2015.07.006

Danitz, S. B., & Orsillo, S. M. (2014). The mindful way through the semester: An investigation of the effectiveness of an acceptance-based behavioral therapy program on psychological wellness in first-year students. *Behavior Modification, 38,* 549–566. https://doi.org/10.1177/0145445513520218

Davies, E. B., Morriss, R., & Glazebrook, C. (2014). Computer-delivered and web-based interventions to improve depression, anxiety, and psychological well-being of university students: A systematic review and meta-analysis. *Journal of Medical Internet Research, 16*. https://doi.org/10.2196/jmir.3142

Deci, E., & Ryan, R.M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. Plenum Press: New York, NY.

De Wilde, E. F., & Hendriks, V. M. (2005). The client satisfaction questionnaire: Psychometric properties in a Dutch addict population. *European Addiction Research, 11*,157–162. https://doi.org/10.1159/000086396

Disabato, D. J., Short, J. L., Lameira, D. M., Bagley, K. D., & Wong, S. J. (2018). Predicting help-seeking behavior: The impact of knowing someone close who has sought help. *Journal of American College Health, 66,* 731-738. https://doi.org/10.1080/07448481.2018.1440568

Donkin, L., Hickie, I. B., Christensen, H., Naismith, S. L., Neal, B., Cockayne, N. L., & Glozier, N. (2013). Rethinking the dose-response relationship between usage and outcome in an online intervention for depression: Randomized controlled trial. *Journal of Medical Internet Research, 15*. https://doi.org/10.2196/jmir.2771

Drake, E. D., Sladek, M. D., & Doane, L. D. (2016). Daily cortisol activity, loneliness, and

coping efficacy in late adolescence: A longitudinal study of the transition to college. *International Journal of Behavioral Development, 40,* 334–345. https://doi.org/10.1177/0165025415581914

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the short grit scale (Grit-S). *Journal of Personality Assessment, 91*, 166–174. https://doi.org/10.1080/00223890802634290

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple , graphical test measures of funnel plot asymmetry. *BMJ, 315*, 629–634. https://doi.org/10.1016/S0895-4356(99)00204-8

Eisenberg, D., Downs, M. F., Golberstein, E., & Zivin, K. (2009). Stigma and help seeking for mental health among college students. *Medical Care Research and Review, 66,* 522–541.

Eisenberg, D., Golberstein, E., & Gollust, S. E. (2007). Help-seeking and access to mental health care in a university student population. *Medical Care, 45,* 594–601. http://dx.doi.org/10.1097/MLR.0b013e31803bb4c1

Eisenberg, D., Golberstein, E., & Hunt, J. (2009). Mental health and academic success in college. *The B . E . Journal of Economic Analysis & Policy Contributions College College, 9,* 1–35. https://doi.org/10.2202/1935-1682.2191

Eisenberg, D., Gollust, S. E., Golberstein, E., & Hefner, J. L. (2007). Prevalence and correlates of depression, anxiety, and suicidality among university students. *The American Journal of Orthopsychiatry, 77,* 534–542. https://doi.org/10.1037/0002-9432.77.4.534

Eisenberg, D., Hunt, J., & Speer, N. (2013). Mental health in American colleges and universities: variation across student subgroups and across campuses. *The Journal of Nervous and Mental Disease, 201,* 60–67. http://dx.doi.org/10.1097/NMD.0b013e31827ab077

Eisenberg, D., Hunt, J., Speer, N., & Zivin, K. (2011). Mental health service utilization among college students in the United States. *The Journal of Nervous and Mental Disease, 199,* 301–308. https://doi.org/10.1097/NMD.0b013e3182175123

Ellis, L. A., Collin, P., Davenport, T. A., Hurley, P. J., Burns, J. M., & Hickie, I. B. (2012). Young men, mental health, and technology: implications for service design and delivery in the digital age. *Journal of Medical Internet Research*, *14,* e160-169. https://doi.org/10.2196/jmir.2291\r10.2196/jmir.2291.

Ellis, L. a, Campbell, A. J., Sethi, S., & O'Dea, B. M. (2011). Comparative randomized trial of an online cognitive-behavioral therapy program and an online support group for depression and anxiety. *Journal of CyberTherapy and Rehabilitation*, *4,* 461–467. https://doi.org/10.1037/t00735-000

Emshoff, J. G. (2008). Researchers, practitioners, and funders: Using the framework to get us on the same page. *American Journal of Community Psychology, 41,* 393-403. https://doi.org/10.1007/s10464-008-9168-x

Escoffery, C., Miner, K. R., Adame, D. D., Butler, S., McCormick, L., & Mendell, E. (2005). Internet use for health information among college students. *Journal of American College Health*, *53,* 183–188. https://doi.org/10.3200/JACH.53.4.183-188

Eskreis-Winkler, L., Shulman, E. P., Beal, S. A., & Duckworth, A. L. (2014). The grit effect: Predicting retention in the military, the workplace, school and marriage. *Frontiers in Psychology*, *5,* 36. https://doi.org/10.3389/fpsyg.2014.00036

Evans, T. M., Bira, L., Gastelum, J. B., Weiss, L. T., & Vanderford, N. L. (2018). Evidence for a mental health crisis in graduate education. *Nature Biotechnology, 36,* 282. https://doi.org/10.1038/nbt.4089

Fan, X., Miller, B. C., Park, K.-E., Winward, B. W., Christensen, M., Grotevant, H. D., & Tai, R. H. (2006). An Exploratory study about inaccuracy and invalidity in adolescent self-report surveys. *Field Methods, 18,* 223–244. https://doi.org/10.1177/152822X06289161

Farrer, L., Griffiths, K. M., Christensen, H., Mackinnon, A. J., & Batterham, P. J. (2014). Predictors of adherence and outcome in internet-based cognitive behavior therapy delivered in a telephone counseling setting. *Cognitive Therapy and Research, 38,* 358–367. https://doi.org/10.1007/s10608-013-9589-1

Farrer, L., Gulliver, A., Chan, J. K. Y., Batterham, P. J., Reynolds, J., Calear, A., … Griffiths, K. M. (2013). Technology-based interventions for mental health in tertiary students: Systematic review. *Journal of Medical Internet Research, 15,* e101. https://doi.org/10.2196/jmir.2639

Fleming, T., Bavin, L., Lucassen, M., Stasiak, K., Hopkins, S., & Merry, S. (2018). Beyond the trial: Systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. *Journal of Medical Internet Research, 20,* e99. https://doi.org/10.2196/jmir.9275

Flett, G. L., Blankstein, K. R., & Martin, T. R. (1995). Procrastination, negative self-evaluation, and stress in depression and anxiety. In *Procrastination and Task Avoidance* (pp. 137–167). https://doi.org/10.1007/978-1-4899-0227-6_7

Frazier, P., Meredith, L., Greer, C., Paulsen, J. A., Howard, K., Dietz, L. R., & Qin, K. (2015). Randomized controlled trial evaluating the effectiveness of a web-based stress management program among community college students. *Anxiety, Stress and Coping, 28,* 576–586. https://doi.org/10.1080/10615806.2014.987666

Galdas, P. M., Cheater, F., & Marshall, P. (2005). Men and health help-seeking behaviour:

Literature review. *Journal of Advanced Nursing, 49,* 616-623. https://doi.org/10.1111/j.1365-2648.2004.03331.x

Gallo, K. P., Comer, J. S., & Barlow, D. H. (2013). Direct-to-consumer marketing of psychological treatments for anxiety disorders. *Journal of Anxiety Disorders*, *27*, 793–801. https://doi.org/10.1016/j.janxdis.2013.03.005

Gallo, K. P., Comer, J. S., Barlow, D. H., Clarke, R. N., & Antony, M. M. (2015). Direct-to-consumer marketing of psychological treatments: A Randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *83*, 994–998. https://doi.org/10.1037/a0039470

Garland, A. F., Hawley, K. M., Brookman-Frazee, L., & Hurlburt, M. S. (2008). Identifying common elements of evidence-based psychosocial treatments for children's disruptive behavior problems. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47,* 505–514. https://doi.org/10.1097/CHI.0b013e31816765c2

Geraghty, A. W. A., Wood, A. M., & Hyland, M. E. (2010). Attrition from self-directed interventions: Investigating the relationship between psychological predictors, intervention content and dropout from a body dissatisfaction intervention. *Social Science and Medicine*, *71,* 30–37. https://doi.org/10.1016/j.socscimed.2010.03.007

Gilbody, S., Richards, D., Brealey, S., & Hewitt, C. (2007). Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): a diagnostic meta-analysis. *Journal of General Internal Medicine*, *22*, 1596–1602. https://doi.org/10.1007/s11606-007-0333-y

Glasgow, R. E., Vogt, T. M., & Boles, S. M. (1999). Evaluating the public health impact of health promotion interventions: The RE-AIM framework. *American Journal of Public Health, 89*, 1322-1327. https://doi.org/10.2105/AJPH.89.9.1322

Gordon, R. S. (1983). An operational classification of disease prevention. *Public Health Reports*,

98, 107–109. https://doi.org/10.2307/4627374

Gottschall, A. C., West, S. G., & Enders, C. K. (2012). A Comparison of Item-Level and Scale-Level Multiple Imputation for Questionnaire Batteries. *Multivariate Behavioral Research*, *47*(1), 1–25. https://doi.org/10.1080/00273171.2012.640589

Grassi, A., Gaggioli, A., & Riva, G. (2009). The green valley: the use of mobile narratives for reducing stress in commuters. *Cyberpsychology & Behavior*, *12*, 155–161. https://doi.org/10.1089/cpb.2008.0156

Gulliver, A., Griffiths, K. M., & Christensen, H. (2010). Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. *BMC Psychiatry*, *10*, 113-121. https://doi.org/10.1186/1471-244X-10-113

Heath, P. J., Seidman, A. J., Vogel, D. L., Cornish, M. A., & Wade, N. G. (2017). Help-seeking stigma among men in the military: The interaction of restrictive emotionality and distress. *Psychology of Men and Masculinity*, *18*, 193–197. https://doi.org/10.1037/men0000111

Hebert, E. a, Vincent, N., Lewycky, S., & Walsh, K. (2010). Attrition and adherence in the online treatment of chronic insomnia. *Behavioral Sleep Medicine*, *8*, 141–150. https://doi.org/10.1080/15402002.2010.487457

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational and Behavioral Statistics*, *6*, 107–128. https://doi.org/10.3102/10769986006002107

Higa-McMillan, C. K., Francis, S. E., Rith-Najarian, L., & Chorpita, B. F. (2016). Evidence base update: 50 years of research on treatment for child and adolescent anxiety. *Journal of Clinical Child and Adolescent Psychology*, *45*, 91–113. https://doi.org/10.1080/15374416.2015.1046177

Hughes, J. (2018). reghelper: Helper Functions for Regression Analysis. Retrieved from https://cran.r-project.org/web/packages/reghelper/index.html

Hunt, J., & Eisenberg, D. (2010). Mental health problems and help-seeking behavior among college students. *Journal of Adolescent Health*, *46*, 3–10. https://doi.org/10.1016/j.jadohealth.2009.08.008

Hyun, J. K., Quinn, B. C., Madon, T., & Lustig, S. (2006). Graduate Student Mental Health: Needs Assessment and Utilization of Counseling Services. *Journal of College Student Development*, *47*, 247–266. https://doi.org/10.1353/csd.2006.0030

IBM Corp. (2016). IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.

Jacobson, N. C., & Newman, M. G. (2017). Anxiety and depression as bidirectional risk factors for one another: A meta-analysis of longitudinal studies. *Psychological Bulletin, 143,* 1155-1200. https://doi.org/10.1037/bul0000111

Jaeger, B. (2017). r2glmm: Computes R Squared for Mixed (Multilevel) Models. Retrieved from https://cran.r-project.org/web/packages/r2glmm/index.html

Kahan, B., Jairath, V., Dore, C., & Morris, T. (2014). The risks and rewards of covariate adjustment in randomized trials: an assessment of 12 outcomes from 8 studies. *Trials, 15,* 139. https://doi.org/10.1186/1745-6215-15-139

Karyotaki, E., Kleiboer, A., Smit, F., Turner, D. T., Pastor, A. M., Andersson, G., … Cuijpers, P. (2015). Predictors of treatment dropout in self-guided web-based interventions for depression: An "individual patient data" meta-analysis. *Psychological Medicine, 45*, 2717-2726. https://doi.org/10.1017/S0033291715000665

Kauer, S. D., Mangan, C., & Sanci, L. (2014). Do online mental health services improve help-seeking for young people? a systematic review. *Journal of Medical Internet Research, 16,*

e66-84. https://doi.org/10.2196/jmir.3103

Kaye, M. P., Faber, A., Davenport, K. E., & Perkins, D. F. (2018). Common components of evidence-informed home visitation programs for the prevention of child maltreatment. *Children and Youth Services Review, 90,* 94–105. https://doi.org/10.1016/j.childyouth.2018.05.009

Keenan-Miller, D., Hammen, C. L., & Brennan, P. A. (2007). Health outcomes related to early adolescent depression. *Journal of Adolescent Health, 41*, 256-262. https://doi.org/10.1016/j.jadohealth.2007.03.015

Kelders, S. M., Kok, R. N., Ossebaard, H. C., & Van Gemert-Pijnen, J. E. W. C. (2012). Persuasive system design does matter: A systematic review of adherence to web-based interventions. *Journal of Medical Internet Research*. https://doi.org/10.2196/jmir.2104

Kelly, P. J., Kyngdon, F., Ingram, I., Deane, F. P., Baker, A. L., & Osborne, B. A. (2018). The Client Satisfaction Questionnaire-8: Psychometric properties in a cross-sectional survey of people attending residential substance abuse treatment. *Drug and Alcohol Review, 37*, 79-86. https://doi.org/10.1111/dar.12522

Ketchen Lipson, S., Gaddis, S. M., Heinze, J., Beck, K., & Eisenberg, D. (2015). Variations in student mental health and treatment utilization across US colleges and universities. *Journal of American College Health, 63,* 388–396. https://doi.org/10.1080/07448481.2015.1040411

Keyes, C. L., Eisenberg, D., Perry, G. S., Dube, S. R., Kroenke, K., & Dhingra, S. S. (2012). The Relationship of Level of Positive Mental Health with Current Mental Disorders in Predicting Suicidal Behavior and Academic Impairment in College Students. *Journal of American College Health, 60,* 126–133. https://doi.org/10.1080/07448481.2011.608393

Klink, R. R. (2000). Creating brand names with meaning: The use of sound symbolism.

Marketing Letters, *11*, 5–20. https://doi.org/10.2307/40216555

Knight, R. G., Waal-Manning, H. J., & Spears, G. F. (1983). Some norms and reliability data for the State--Trait Anxiety Inventory and the Zung Self-Rating Depression scale. *The British Journal of Clinical Psychology / the British Psychological Society*, *22*, 245–249. https://doi.org/10.1111/j.2044-8260.1983.tb00610.x

Knudsen, K. S., Boustani, M. M., Chu, W., Wesley, K., & Chorpita, B. F. (2018, July). Do current publication standards reliably describe content of manualized universal prevention programs? *Poster Presented at 2018 Association for Psychological Science*.

Kohl, L. F. M., Crutzen, R., & De Vries, N. K. (2013). Online prevention aimed at lifestyle behaviors: A systematic review of reviews. *Journal of Medical Internet Research*, *15*, e146. https://doi.org/10.2196/jmir.2665

Kornfield, R., Donohue, J., Berndt, E. R., & Alexander, G. C. (2013). Promotion of prescription drugs to consumers and providers, 2001-2010. *PLoS ONE*, *8*, e55504. https://doi.org/10.1371/journal.pone.0055504

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*, 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174. https://doi.org/10.2307/2529310

Lannin, D. G., Vogel, D. L., Brenner, R. E., Abraham, W. T., & Heath, P. J. (2015). Does Self-Stigma Reduce the Probability of Seeking Mental Health Information? *Journal of Counseling Psychology*, *63*, 351–358. https://doi.org/10.1037/cou0000108

Larsen, D. L., Attkisson, C. C., Hargreaves, W. A., & Nguyen, T. D. (1979). Assessment of

client/patient satisfaction: Development of a general scale. *Evaluation and Program Planning, 2,* 197-207. https://doi.org/10.1016/0149-7189(79)90094-6

Larun, L., Nordheim, L. V, Ekeland, E., Hagen, K. B., & Heian, F. (2006). Exercise in prevention and treatment of anxiety and depression among children and young people. *Cochrane Database of Systematic Reviews*, *3*, CD004691. https://doi.org/10.1002/14651858.CD004691.pub2

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). emmeans: Estimated Marginal Means, aka Least-Squares Means. Retrieved from https://cran.r-project.org/web/packages/emmeans/index.html

Leon, A. C., Davis, L. L., & Kraemer, H. C. (2011). The role and interpretation of pilot studies in clinical research. *Journal of Psychiatric Research*, *45*, 626–629. https://doi.org/10.1016/j.jpsychires.2010.10.008

Leong, F. T., & Lau, A. S. (2001). Barriers to providing effective mental health services to Asian Americans. *Mental Health Services Research, 3,* 201-214. DOI: 10.1023/A:1013177014788

Lieven, T., Grohmann, B., Herrmann, A., Landwehr, J. R., & van Tilburg, M. (2014). The Effect of brand gender on brand equity. *Psychology and Marketing*, *31*, 371–385. https://doi.org/10.1002/mar.20701

Lindsey, M. A., Brandt, N. E., Becker, K. D., Lee, B. R., Barth, R. P., Daleiden, E. L., & Chorpita, B. F. (2014). Identifying the Common Elements of Treatment Engagement Interventions in Children's Mental Health Services. *Clinical Child and Family Psychology Review*, *17,* 283–298. https://doi.org/10.1007/s10567-013-0163-x

Lintvedt, O. K., Griffiths, K. M., Sørensen, K., Østvik, A. R., Wang, C. E. a, Eisemann, M., & Waterloo, K. (2013). Evaluating the effectiveness and efficacy of unguided internet-based

self-help intervention for the prevention of depression: a randomized controlled trial. *Clinical Psychology & Psychotherapy*, *20*, 10–27. https://doi.org/10.1002/cpp.770

Löwe, B., Kroenke, K., & Gräfe, K. (2005). Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *Journal of Psychosomatic Research*, *58*(2), 163–171. https://doi.org/10.1016/j.jpsychores.2004.09.006

Lyubomirsky, S., Dickerhoof, R., Boehm, J. K., & Sheldon, K. M. (2011). Becoming happier takes both a will and a proper way: an experimental longitudinal intervention to boost well-being. *Emotion*, *11*, 391. http://dx.doi.org/10.1037/a0022575

Mackey, T. K., Cuomo, R. E., & Liang, B. A. (2015). The rise of digital direct-to-consumer advertising?: Comparison of direct-to-consumer advertising expenditure trends from publicly available data sources and global policy implications. *BMC Health Services Research*, *15*, 236. https://doi.org/10.1186/s12913-015-0885-1

Mahmoud, J., Staten, R., Hall, L., & Lennie, T. (2012). The Relationship among Young Adult College Students' Depression, Anxiety, Stress, Demographics, Life Satisfaction, and Coping Styles. *Issues in Mental Health Nursing*, *33*, 149–156. https://doi.org/10.3109/01612840.2011.632708

Moreno-Peral, P., Conejo-Cerón, S., Rubio-Valera, M., Fernández, A., Navas-Campaña, D., Rodríguez-Morejón, A., … Bellón, J. Á. (2017). Effectiveness of psychological and/or educational interventions in the prevention of anxiety: A systematic review, meta-analysis, and meta-regression. *JAMA Psychiatry*, *74*, 1021-1029. https://doi.org/10.1001/jamapsychiatry.2017.2509

Murray, L. K., Dorsey, S., Haroz, E., Lee, C., Alsiary, M. M., Haydary, A., … Bolton, P. (2014). A common elements treatment approach for adult mental health problems in low- and

middle-income countries. *Cognitive and Behavioral Practice, 21*, 111–123. https://doi.org/10.1016/j.cbpra.2013.06.005

Musiat, P., Conrod, P., Treasure, J., Tylee, A., Williams, C., & Schmidt, U. (2014). Targeted prevention of common mental health disorders in university students: Randomised controlled trial of a transdiagnostic trait-focused web-based intervention. *PLoS ONE, 9*, e93621. https://doi.org/10.1371/journal.pone.0093621

NAMI. (2012). *College students speak: Survey report on mental health*. Retrieved from www.nami.org/collegereport

National Institutes of Health. (2013). PAR-10-038: Dissemination and Implementation Research in Health (R01). Retrieved December 23, 2018, from https://grants.nih.gov/grants/guide/pa-files/PAR-10-038.html

Norton, P. J. (2007). Depression Anxiety and Stress Scales (DASS-21): Psychometric analysis across four racial groups. *Anxiety, Stress and Coping, 20*, 253–265. https://doi.org/10.1080/10615800701309279

Oei, T. P. S., Sawang, S., Goh, Y. W., & Mukhtar, F. (2013). Using the Depression Anxiety Stress Scale 21 (DASS-21) across cultures. *International Journal of Psychology, 48*, 1018-1029. https://doi.org/10.1080/00207594.2012.755535

Olivier, J., & Bell, M. L. (2013). Effect sizes for 2×2 contingency tables. *PLoS ONE, 8*, e58777. https://doi.org/10.1371/journal.pone.0058777

Orsmond, G. I., & Cohn, E. S. (2015). The distinctive features of a feasibility study: Objectives and guiding questions. *OTJR Occupation, Participation and Health, 35*, 169-177. https://doi.org/10.1177/1539449215578649

Ottenbreit, N. D., & Dobson, K. S. (2004). Avoidance and depression: The construction of the

Cognitive-Behavioral Avoidance Scale. *Behaviour Research and Therapy, 42,* 293–313. https://doi.org/10.1016/S0005-7967(03)00140-2

Owen, N., Goode, A., Sugiyama, T., Javad Koohsari, M., Healy, G., Fjeldsoe, B., & Eakin, E. (2017). Designing for Dissemination in Chronic Disease Prevention and Management. In R. Brownson, G. Colditz, & E. Proctor (Eds.), *Dissemination and Implementation Research in Health: Translating Science to Practice* (2nd ed., pp. 107–120). New York, NY: Oxford University Press.

Pedrelli, P., Nyer, M., Yeung, A., Zulauf, C., & Wilens, T. (2015). College students: Mental health problems and treatment considerations. *Academic Psychiatry, 39,* 503-511. https://doi.org/10.1007/s40596-014-0205-9

Pocock, S., Assmann, S., Enos, L., & Kasten, L. (2002). Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. *Statistics in Medicine, 21,* 2917–2930. https://doi.org/10.1002/sim.1296

Powell, B. J., Beidas, R. S., Lewis, C. C., Aarons, G. A., McMillen, J. C., Proctor, E. K., & Mandell, D. S. (2017). Methods to Improve the Selection and Tailoring of Implementation Strategies. *Journal of Behavioral Health Services and Research, 44,* 177-194. https://doi.org/10.1007/s11414-015-9475-6

PracticeWise. (2012). *Psychosocial and Combined Treatments Coding Manual*. Satellite Beach, FL: PracticeWise, LLC.

R Core Team. (2013). R: A Language and Environment for Computing. *R Foundation for Statistical Computing, Vienna, Austria*. Vienna, Austria: R Foundation for Statistical Computing. https://doi.org/http://www.R-project.org/

Raab, G., Day, S., & Sales, J. (2000). How to select covariates to include in the analysis of a

clinical trial. *Controlled Clinical Trials*, *21*, 330–342.

Rabin, B., & Brownson, R. (2017). Terminology for Dissemination and Implementation Research. In R. Brownson, G. Colditz, & E. Proctor (Eds.), *Dissemination and Implementation Research in Health: Translating Science to Practice* (2nd ed., pp. 19–46). New York, NY: Oxford University Press.

Ranøyen, I., Lydersen, S., Larose, T., Weidle, B., Skokauskas, N., Thomsen, P., & Indredavik, M. (2018). Developmental course of anxiety and depression from adolescence to young adulthood in a prospective Norwegian clinical cohort. *European Child & Adolescent Psychiatry*, *27,* 1413–1423. https://doi.org/10.1007/s00787-018-1139-7

Rapee, R. M., Wignall, A., Sheffield, J., Kowalenko, N., Davis, A., McLoone, J., & Spence, S. H. (2006). Adolescents' reactions to universal and indicated prevention programs for depression: Perceived stigma and consumer satisfaction. *Prevention Science*, *7*, 167–177. https://doi.org/10.1007/s11121-006-0035-4

Ratanasiripong, P., Ratanasiripong, N., & Kathalae, D. (2012). Biofeedback intervention for stress and anxiety among nursing students: A Randomized controlled trial. *International Scholarly Research Network Nursing*, 1–5. https://doi.org/10.5402/2012/827972

Reavley, N., & Jorm, A. F. (2010). Prevention and early intervention to improve mental health in higher education students: A review. *Early Intervention in Psychiatry, 4,* 132-142. https://doi.org/10.1111/j.1751-7893.2010.00167.x

Reetz, D., Bershad, C., LeViness, P., & Whitlock, M. (2016). *The association for university and college counseling center directors annual survey (Reporting period: September 1, 2015 through August 31, 2016).*

Regehr, C., Glancy, D., & Pitts, A. (2013). Interventions to reduce stress in university students:

A review and meta-analysis. *Journal of Affective Disorders*, *148*, 1–11. https://doi.org/10.1016/j.jad.2012.11.026

Rennis, L., McNamara, G., Seidel, E., & Shneyderman, Y. (2015). Google it!: urban community college students' use of the internet to obtain self-care and personal health information. *College Student Journal*, *49*, 414–426.

Richards, D. (2009). Features and benefits of online counselling: Trinity college online mental health community. *British Journal of Guidance and Counselling*, *37*, 231–242. https://doi.org/10.1080/03069880902956975

Rickwood, D., Deane, F. P., Wilson, C. J., & Ciarrochi, J. V. (2005). Young people's help-seeking for mental health problems . *Australian E-Journal for the Advancement of Mental Health*, *4*, 218–251. https://doi.org/10.5172/jamh.4.3.218

Rith-Najarian, L. R. (2015). The Happiness Challenge at UCLA: Lessons learned [Unpublished presentation] (pp. 1–16). Los Angeles, CA: MindWell Committee Monthly Meeting (March, 2015).

Rith-Najarian, L. R., Daleiden, E. L., & Chorpita, B. F. (2016). Evidence-based decision making in youth mental health prevention. *American Journal of Preventive Medicine*, *51*, S132–S139. https://doi.org/10.1016/j.amepre.2016.05.018

Rochlen, A. B., Blazina, C., & Raghunathan, R. (2002). Gender role conflict, attitudes toward career counseling, career decision-making, and perceptions of career counseling advertising brochures. *Psychology of Men and Masculinity*, *3*, 127–137. https://doi.org/10.1037/1524-9220.3.2.127

Rose, R. D., Buckey, J. C., Zbozinek, T. D., Motivala, S. J., Glenn, D. E., Cartreine, J. a., & Craske, M. G. (2013). A randomized controlled trial of a self-guided, multimedia, stress

management and resilience training program. *Behaviour Research and Therapy, 51,* 106–112. https://doi.org/10.1016/j.brat.2012.11.003

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86,* 638–641. http://dx.doi.org/10.1037/0033-2909.86.3.638

Rotenstein, L. S., Ramos, M. A., Torre, M., Bradley Segal, J., Peluso, M. J., Guille, C., … Mata, D. A. (2016). Prevalence of depression, depressive symptoms, and suicidal ideation among medical students a systematic review and meta-analysis. *JAMA - Journal of the American Medical Association, 316,* 2214-2236. https://doi.org/10.1001/jama.2016.17324

Rotheram-Borus, M. J., Swendeman, D., & Chorpita, B. F. (2012). Disruptive innovations for designing and diffusing evidence-based interventions. *American Psychologist, 67,* 463–476. https://doi.org/10.1037/a0028180

Ryan, M. L., Shochet, I. M., & Stallman, H. M. (2010). Universal online interventions might engage psychologically distressed university students who are unlikely to seek formal help. *Advances in Mental Health, 9,* 73–83. https://doi.org/10.5172/jamh.9.1.73

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55,* 68-78. doi: 10.1037110003-066X.55.1.68

Ryan, R. M., & Deci, E. L. (2008). A self-determination theory approach to psychotherapy: The motivational basis for effective change. *Canadian Psychology, 49,* 186–193. DOI: 10.1037/a0012753

Ryan, R. M., Plant, R. W., & O'Malley, S. (1995). Initial motivations for alcohol treatment: Relations with patient characteristics, treatment involvement, and dropout. *Addictive Behaviors, 20,* 279–297. https://doi.org/10.1016/0306-4603(94)00072-7

Salzer, M. (2012). A comparative study of campus experiences of college students with mental illnesses versus a general college sample. *Journal of American College Health, 60*, 1–7. https://doi.org/10.1080/07448481.2011.552537

Sandler, I., Ostrom, A., Bitner, M. J., Ayers, T. S., Wolchik, S., & Daniels, V. S. (2005). Developing effective prevention services for the real world: A prevention service development model. *American Journal of Community Psychology, 35*, 127–142. https://doi.org/10.1007/s10464-005-3389-z

Saravanan, C., & Wilks, R. (2014). Medical students' experience of and reaction to stress: The role of depression and anxiety. *The Scientific World Journal, 2014*, 1-8. https://doi.org/10.1155/2014/737382

Schulenberg, J. E., Sameroff, A. J., & Cicchetti, D. (2004). The transition to adulthood as a critical juncture in the course of psychopathology and mental health. *Development and Psychopathology, 16*, 799–806. https://doi.org/10.1017/S0954579404040015

Schulz, K., Altman, D., Moher, D., & CONSORT Group. (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC Medicine, 8*, 18-27. doi: 10.1136/bmj.c332.

Schwartz, S. J., Cote, J. E., & Arnett, J. J. (2005). Identity and agency in emerging adulthood: Two development routes in the individualization process. *Youth & Society, 37*, 201–229. https://doi.org/10.1177/0044118x05275965

Schwarz, N. (2007). Cognitive aspects of survey methodology. *Applied Cognitive Psychology, 21*, 277-287. https://doi.org/10.1002/acp.1340

Solomon, D. A., Keller, M. B., Leon, A. C., Mueller, T. I., Lavori, P. W., Shea, M. T., … Endicott, J. (2000). Multiple recurrences of major depressive disorder. *American Journal of*

Psychiatry, *157*, 229-233. https://doi.org/10.1176/appi.ajp.157.2.229

Sontag-Padilla, L., Woodbridge, M. W., Mendelsohn, J., D'Amico, E. J., Osilla, K. C., Jaycox, L. H., … Stein, B. D. (2016). Factors affecting mental health service utilization among California public college and university students. *Psychiatric Services*, *67*, 890–897. https://doi.org/10.1176/appi.ps.201500307

Southam-Gerow, M. a, Daleiden, E. L., Chorpita, B. F., Bae, C., Mitchell, C., Faye, M., & Alba, M. (2014). MAPping Los Angeles County: Taking an Evidence-Informed Model of Mental Health Care to Scale. *Journal of Clinical Child and Adolescent Psychology*, *43*, 190–200. https://doi.org/10.1080/15374416.2013.833098

Spielberger, C. D. (2010). State-Trait Anxiety Inventory. In *The Corsini Encyclopedia of Psychology*. https://doi.org/10.1002/9780470479216.corpsy0943

Stevens, J. (1992). *Applied multivariate statistics for the social sciences* (2nd ed.). Hillsdale, NJ: Erlbaum. https://doi.org/10.1198/tas.2003.s213

Sun, S., Hoyt, W. T., Brockberg, D., Lam, J., & Tiwari, D. (2016). Acculturation and enculturation as predictors of psychological help-seeking attitudes (HSAs) among racial and ethnic minorities: A meta-analytic investigation. *Journal of Counseling Psychology*, *63*, 617–632. https://doi.org/10.1037/cou0000172

Torgerson, C. J. (2006). Publication bias: The achilles' heel of systematic reviews? *British Journal of Educational Studies*, *54*, 89–102. https://doi.org/10.1111/j.1467-8527.2006.00332.x

Torous, J., Nicholas, J., Larsen, M. E., Firth, J., & Christensen, H. (2018). Clinical review of user engagement with mental health smartphone apps: Evidence, theory and improvements. *Evidence-Based Mental Health*. https://doi.org/10.1136/eb-2018-102891

Tourangeau, R., & Yan, T. (2007). Sensitive Questions in Surveys. *Psychological Bulletin, 133,* 859–883. https://doi.org/10.1037/0033-2909.133.5.859

Van Ballegooijen, W., Cuijpers, P., Van Straten, A., Karyotaki, E., Andersson, G., Smit, J. H., & Riper, H. (2014). Adherence to internet-based and face-to-face cognitive behavioural therapy for depression: A meta-analysis. *PLoS ONE, 9,* e100674. https://doi.org/10.1371/journal.pone.0100674

Vogel, D. L., Wade, N. G., Wester, S. R., Larson, L., & Hackler, A. H. (2007). Seeking help from a mental health professional: The influence of one's social network. *Journal of Clinical Psychology, 63,* 233-245. https://doi.org/10.1002/jclp.20345

Webb, T., Joseph, J., Yardley, L., & Michie, S. (2010). Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *Journal of Medical Internet Research, 12,* e4. https://doi.org/10.2196/jmir.1376

Weisz, J. R., Sandler, I. N., Durlak, J. A., & Anton, B. S. (2005). Promoting and protecting youth mental health through evidence-based prevention and treatment. *American Psychologist, 60,* 628–648. https://doi.org/10.1037/0003-066X.60.6.628

Weisz, J. R., Ugueto, A. M., Herren, J., Afienko, S. R., & Rutt, C. (2011). Kernels vs. Ears, and Other Questions for a Science of Treatment Dissemination. *Clinical Psychology: Science and Practice, 18,* 41–46. https://doi.org/10.1111/j.1468-2850.2010.01233.x

Werner-Seidler, A., Perry, Y., Calear, A. L., Newby, J. M., & Christensen, H. (2017). School-based depression and anxiety prevention programs for young people: A systematic review and meta-analysis. *Clinical Psychology Review, 51,* 30-47. https://doi.org/10.1016/j.cpr.2016.10.005

Wojtowicz, M., Day, V., & McGrath, P. J. (2013). Predictors of participant retention in a guided online self-help program for university students: Prospective cohort study. *Journal of Medical Internet Research, 15,* e96. https://doi.org/10.2196/jmir.2323

Yorgason, J. B., Linville, D., & Zitzman, B. (2008). Mental health among college students: Do those who need services know about and use them. *Journal of American College Health, 57*, 173–182.

Yusoff, M. S. B., & Esa, A. R. (2015). A DEAL-based intervention for the reduction of depression, denial, self-blame and academic stress: A randomized controlled trial. *Journal of Taibah University Medical Sciences, 10,* 82–92. https://doi.org/10.1016/j.jtumed.2014.08.003

Zivin, K., Eisenberg, D., Gollust, S. E., & Golberstein, E. (2009). Persistence of mental health problems and needs in a college student population. *Journal of Affective Disorders, 117,* 180–185.