A Multiple-Baseline Study of a Mobile Cognitive Behavioral Therapy for the Treatment of Eating Disorders in College Students

By

Brittany Kay Bohrer

M.A., University of Kansas, 2015

B.S., University of North Dakota, 2011

Submitted to the graduate degree program in Psychology and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Chair: Kelsie Forbush, Ph.D., L.P.

Ann Davis, Ph.D., M.P.H., ABPP

Rick Ingram, Ph.D.

David Jarmolowicz, Ph.D.

Sarah Kirk, Ph.D., ABPP

Date Defended: July 22, 2020

The dissertation committee for Brittany Kay Bohrer certifies that this is the approved version of the following dissertation:

A Multiple-Baseline Study of a Mobile Cognitive Behavioral Therapy for the Treatment of Eating Disorders in College Students

Chair: Kelsie Forbush, Ph.D., L.P.

Date Approved: July 22, 2020

Abstract

Introduction: Eating disorders are serious psychiatric disorders associated with substantial morbidity and the highest rate of mortality of all psychiatric disorders and they are especially prevalent among college students. Despite the severity and prevalence of eating disorders, fewer than 20% of college students with eating disorders receive help for their eating problems. Given the lack of eating-disorder treatment on many college campuses, mobile adaptations of evidence-based treatments may represent an innovative method of disseminating treatment to a larger number of college students with eating disorders. Thus, the purpose of this study was to administer a mobile, self-guided cognitive-behavioral therapy (CBT-gsh) for reducing eating-disorder psychopathology in college students with eating disorders. **Method:** A multiple-baseline design was used to examine intervention effects on eating-disorder psychopathology, clinical impairment, and internalizing psychopathology in three college students. Data were examined using visual analysis and Tau-*U* effect-size calculations. Observed means were compared to normed means. **Results:** Participants demonstrated significant decreases in Eating Pathology Clinical Outcomes Tracking (EPCOT) Total Score and Binge Eating. Results were mixed for EPCOT Restricting, Excessive Exercise, and Body Dissatisfaction. No participants reported purging. **Conclusions:** The current study was one of the first to examine mobile CBT-gsh for the treatment of eating disorders in college students and one of few applications of a multiple-baseline design to examine treatment effects in the field of eating disorders. The current findings encourage further testing of the intervention to replicate and extend the observed treatment effects. Mobile CBT-gsh may represent an innovative tool that could be scaled to reach a larger number of persons with eating disorders, especially those who are under-served. (Word Count: 265 words)

Acknowledgements

First and foremost, I want to express my deepest gratitude to the participants in the current study. None of this would have been possible without you and I am so grateful to have had the opportunity to watch you make changes that are truly inspiring. Furthermore, this project would not have been possible without funding support from the University of Kansas Research Excellence Initiative (Principal Investigator: Dr. Kelsie Forbush), the generous donors who awarded me an RSH Scholarship to conduct this work, and the Academy for Eating Disorders. I would like to thank my dissertation committee members for their patience. I must also thank Dr. Kelsie Forbush for her continued support, as I would not be where I am today without her mentorship. Thank you, Kelsie, for believing in me.

Enthusiastic thank-you's are also in order for several colleagues and friends (Anthony Rohr, Ashley Schlotfeldt, Erika Sorensen, Jordan Bonstrom, Mari Thomeczek, Natalie Babich, CARE Lab members, and Drs. Amber Jarnecke, Kara Christensen, Kate Esterline, and Kelsey Hagan) and for my clinical supervisors, particularly Dr. Cyd Schnacke, for providing me with overwhelming support and many laughs. Thank you for refusing to let me give up. I would also like to express gratitude to my family and friends for deciding a very long time ago to refrain from asking me when I will graduate (it finally happened!). I truly have some of the most supportive parents, friends, and a sister who care about me deeply, and for that I am forever grateful. Lastly, I want to express pride in myself for surviving graduate school, an experience that proved to be the most challenging and rewarding I have had to-date. The learning that took place during my graduate-school career will forever impact my personal- and professional endeavors. In the words of the late Freddie Mercury, "Don't Stop Me Now…"

Table of Contents

A Multiple-Baseline Study of a Mobile Cognitive Behavioral Therapy for the Treatment of Eating Disorders in College Students

Eating disorders are serious psychiatric disorders associated with substantial psychiatric and medical morbidity and mortality (Arcelus et al., 2011; Harris & Barraclough, 1998; Klump et al., 2009; Preti et al., 2011). Eating disorders are prevalent in college students, with prevalence estimates ranging from 3.6% in college men to 15.4% in college women (Eisenberg et al., 2011; Fitzsimmons-Craft et al., 2019). Despite the severity and prevalence of eating disorders in college students, fewer than 20% of college students with eating disorders receive treatment for their eating problems (Eisenberg et al., 2011). Mobile adaptations of evidence-based treatments may represent an innovative method of disseminating quality treatment to a larger number of college students with eating disorders. Thus, the purpose of this study was to administer self-guided cognitive-behavioral therapy delivered through a mobile-phone application for reducing eating-disorder psychopathology in college students with eating disorders. First I will describe the epidemiology of eating disorders in college students, then treatment barriers, and finally, cognitive-behavioral therapy for eating disorders and its innovative adaptations, including guided self-help and online/mobile delivery formats.

Epidemiology of Eating Disorders in College Students

The median age of onset for eating disorders is 18-21 years of age (Hudson et al., 2007), which coincides with the typical ages of most college students. Rates of eating disorders in college students are similar to population-based prevalence estimates (Galmiche et al., 2019; Hay et al., 2015; Stice et al., 2013); however, a greater number of college students may be at risk for eating disorders. One study of college women (*N*=186) found that disordered-eating prevalence estimates ranged from 40-49% (Berg et al., 2009). In a state-wide screen of Missourian college students, 38.9% of students who completed the screening survey were deemed to be “at high

risk" for an eating disorder (Fitzsimmons-Craft et al., 2019); of note, this estimate was in addition to the 15.4% of students who screened positive for an eating disorder in the same sample. Another large study of U.S. college students found that 29.7% screened positive for an eating disorder (Lipson et al., 2017). Furthermore, large-scale screening at the University of Kansas (KU) found that 20% of freshmen screened positive for an eating disorder (Forbush et al., unpublished data).

Unique challenges to college students that may increase eating-disorder psychopathology include increased independence over eating behaviors, given that students are often transitioning from meals prepared by their parents at home to unlimited food choices at all-you-can-eat style cafeterias (Smith-Jackson & Reel, 2012). Fears of weight gain (e.g., the dreaded "Freshman 15") may trigger unhealthy approaches to weight loss (Smith-Jackson & Reel, 2012). In addition to changes in the food environment, students experience isolation, increased stress related to the transition to college and/or academic demands, and potentially poor wellness and self-management habits (Howard et al., 2020; Wilson et al., 2015). The development of unhealthy eating habits is problematic because eating disorders are associated with increased risk for additional negative outcomes in college students. For example, academic- and social functioning may be impaired in college students with eating disorders (Eisenberg et al., 2009), which has been demonstrated to negatively impact students' future economic potential (Ashwood et al., 2015; Eisenberg et al.,2009). Taken together, the prevalence estimates and evidenced risks associated with eating disorders suggest that college students are particularly at risk for eating disorders—which may be due to social, developmental, and academic challenges that serve as risk factors for eating-disorder development—thus, eating-disorder treatments designed to address the needs of this unique population are critically needed.

Eating-Disorder Treatment Barriers in College Students

Unfortunately, fewer than 20% of college students with eating disorders receive focused care for their eating disorder (Eisenberg et al., 2011). In a large study of U.S. college students, only 13.5% of persons who screened positive for an eating disorder had received eating-disorder treatment in the past year (Lipson et al., 2017). College students reported perceived lack of time to address their eating problems, perceived lack of need for eating-disorder treatment, wanting to address the eating problems "on my own," and being unsure of where to access eating-disorder treatment on campus as barriers to seeking eating-disorder treatment (Eisenberg et al., 2011; Lipson et al., 2017; Sonneville & Lipson, 2018). College students may also have limited time to spend outside the demands of coursework and other meaningful college experiences (e.g., extracurricular sporting activities) and eating-disorder treatment is often time-consuming (e.g., Fairburn et al., 2009).

Additionally, some college students may experience financial hardship associated with the costs of higher education (Lipson et al., 2017). Indeed, an important contributing factor to low eating-disorder service-utilization rates among college students may be the high costs associated with eating-disorder treatment. Outpatient care was estimated to cost approximately $1,500 per year and estimates ranged from $9,000-$17,000 per year for inpatient treatment (Striegel-Moore et al., 2000). Given the psychiatric and medical complications related to eating disorders, healthcare costs are higher for persons with eating disorders than those without eating disorders and similar to healthcare costs for persons with depression (Mitchell et al., 2009). In sum, there are numerous eating-disorder treatment barriers for college students. Intensive treatments for eating disorders are costly and time-consuming, and access to quality eating-disorder-focused treatment may be limited on some college campuses. Even if students have the

option to access to eating-disorder-focused treatment on campus, college mental-health centers are often faced with the problem of having limited resources to meet the needs of an increasing number of students attempting to access mental-health treatment (Mowbray et al., 2006; Watkins et al., 2012; Xiao et al., 2017). Therefore, innovative and evidence-based treatment options for college students with eating disorders that are easily accessible and cost-effective are imperative.

Cognitive-Behavioral Therapy for Eating Disorders

One of the most researched treatments for eating disorders is enhanced cognitive behavioral therapy for eating disorders (CBT-E; Fairburn et al., 2009), which is a form of CBT that was developed specifically for the treatment of eating disorders. CBT-E was designed to be a transdiagnostic treatment modality that is suitable for any type of an eating-disorder diagnosis (Fairburn et al., 2009). CBT-E has the strongest evidence base for the treatment of bulimia nervosa and binge-eating disorder (Fairburn et al., 2009; Hay, 2013; National Collaborating Center for Mental Health, 2004; Yager et al., 2005). Thus, CBT-E is recommended as a "first-line" treatment for persons with bulimia nervosa and binge-eating disorder (National Collaborating Center for Mental Health, 2004; Yager et al., 2005).

CBT-E features psychoeducation and behavioral- and cognitive techniques to reduce eating-disorder behaviors and body dissatisfaction. Traditional CBT-E was designed to take place in-person over the course of 20 weeks. It is comprised of four stages. Patients are seen twice per week during Stage One, once per week during Stages Two and Three, and every-other-week during Stage Four. Stage One represents the most intense stage, as it is the stage where the most behavioral change takes place. Stage One of CBT-E focuses on encouraging the patient to engage in regular eating patterns, which is posited to interrupt episodes of fasting or energy restriction that, over time, lead to increased hunger and over-eating episodes. During this time,

self-monitoring and weekly weighing are introduced. Stage Two is a brief stage where the clinician and patient "take stock" of progress the patient has made thus far in treatment. Changes to the treatment plan can be made during Stage Two if the patient has not responded well to treatment thus far.

Stage Three is where the majority of cognitive work targeting the eating-disorder mechanisms is done. For example, Stage Three in CBT-E involves a focus on cognitive strategies to reduce eating-disorder thoughts and beliefs, such as shape and weight concerns. CBT-E assumes that the core eating-disorder psychopathology ("core psychopathology") is over-evaluation of weight, shape, and one's control over these factors ("shape and weight concerns;" see **Figure 1**). This core psychopathology is typically exemplified by persons with eating disorders using their shape/weight/control as primary ways to evaluate themselves (e.g., "Am I a good or bad person?"), versus using other metrics (e.g., work performance, relationships with loved ones) to evaluate themselves. In addition to cognitive change, Stage Three also focuses on behavioral-change techniques such as teaching clients to recognize and avoid binge-eating triggers and to monitor and reduce "body checking" behaviors, which involve body scrutiny (e.g., pinching or measuring one's body, frequently examining one's body in mirrors or reflective surfaces, etc.). Other treatment "modules" that focus on perfectionism, mood intolerance, and interpersonal concerns can also be added to the intervention, if desired, depending on the patient's presenting concerns. Finally, Stage Four is focused on maintenance of recovery and relapse prevention (Fairburn et al., 2009).

Guided Self-Help

CBT-based approaches have been adapted for use as a guided self-help (CBT-gsh) treatment for eating disorders. CBT-gsh was originally created to include eight sessions over the

course of 12 weeks (Striegel-Moore et al., 2010). Patients cover the therapeutic material between sessions on their own using a research-based self-help book. Therapist involvement during the sessions is brief (i.e., approximately 25-30 minutes) and the therapist (or "health coach") role is to provide support/guidance (i.e., rather than to lead the therapy). As such, CBT-gsh is a treatment modality that requires less time, resources, and training.

CBT-gsh treatment materials target eating-disorder behaviors (e.g., binge eating, body checking, restricting) and cognitions (e.g., negative body image, automatic thoughts). Patients meet with a health coach weekly. Consistent with CBT approaches, daily food logs are implemented and participants are asked to weigh themselves once per week. A primary goal of CBT-gsh is implementation of regular eating to avoid hunger that results from self-starvation, which can lead to binge eating. Another goal related to regular eating and food logs is for participants to learn that changes in eating (i.e., eating on a regular schedule) do not cause inevitable, out-of-control weight gain.

CBT-gsh has demonstrated efficacy in treating bulimia nervosa, binge-eating disorder, and 'other specified feeding or eating disorder' (OSFED) in adults (statistically significant effect sizes ranged from small to large; d = 0.08-1.01; average d = 0.37; Bailer et al., 2004; Grenon et al., 2017; Striegel-Moore et al., 2010; Traviss et al., 2011). National Institute for Health and Clinical Excellence (NICE) guidelines recommend CBT-gsh as a first step for treatment of bulimia nervosa and binge-eating disorder (National Collaborating Center for Mental Health, 2004). In addition to being a cost-effective first-line treatment for bulimia nervosa, binge-eating disorder, and OSFED (Lynch et al., 2010; Crow et al., 2013), CBT-gsh has been shown to be suitable for the treatment of subclinical disordered eating in adults (Striegel-Moore et al., 2010; Traviss et al., 2011), and may represent a useful secondary-prevention program. In other words,

CBT-gsh, when used to treat subclinical disordered eating, may serve to treat unhealthy eating behaviors *and* prevent the onset of eating disorders. This makes CBT-gsh's use for college students recommended because it provides a cost-effective and evidence-based secondary-prevention program, which in turn may reduce the number of students who would go on to develop eating disorders, ultimately reducing the number of students who need a higher level of care.

Online Adaptations. CBT-gsh for eating disorders has been adapted for online administration and has demonstrated preliminary efficacy in treating bulimia nervosa, binge-eating disorder, and OSFED in adults when delivered in an online format, as exhibited by significant improvements in eating-disorder psychopathology and general psychopathology, and abstinence from binge-eating and purging (statistically significant effect sizes ranged from small to large; d = 0.3-1.4; Carrard et al., 2011; Fernández-Aranda et al., 2009; Jacobi et al., 2012; Ljotsson et al., 2007; Robinson & Serfaty, 2008; Ruwaard et al., 2013; Sánchez-Ortiz et al., 2011; Zerwas et al., 2017). Internet-delivered CBT-gsh resulted in lower scores on common self-report eating-disorder measures, as demonstrated by reductions in eating-disorder self-report scales, such as scales that measure binge eating, purging, and distorted cognitions related to weight and shape (Carrard et al., 2011; Fernández-Aranda et al., 2009; Jacobi et al., 2012; Ljotsson et al., 2007; Ruwaard et al., 2013; Sánchez-Ortiz et al., 2011). Other research showed that Internet-delivered CBT-gsh led to greater rates of abstinence from eating-disorder behaviors at the end of treatment and maintenance of improvements at three-to-eighteen-month follow-up for adults with eating disorders, as compared to a waitlist-control group or an unsupported, self-directed-writing control group (Ljotsson et al., 2007; Robinson & Serfaty, 2008; Sánchez-Ortiz et al., 2011; Wagner et al., 2015). Furthermore, Internet-based therapies were found to be viewed

favorably in two large samples of adults who self-reported eating-disorder symptoms, and this was especially true for persons with a probable diagnosis of bulimia nervosa (Linardon et al., 2020; McClay et al., 2016).

A stepped-care examination of digital guided self-help interventions for eating disorders in college students revealed that treatment costs can be decreased and that fewer individuals need in-person treatment (Kass et al., 2017). Despite the potential to greatly reduce cost associated with the need for in-person treatments, few studies have evaluated digital interventions for eating disorders in college students. An online, secondary-prevention CBT-gsh for eating disorders demonstrated significant improvements in weight- and shape concerns in a sample of college women (Taylor et al., 2006). An adaptation of the same online, secondary-prevention CBT-gsh for eating disorders that also targeted depressive symptomatology demonstrated efficacy in improving eating-disorder symptoms and comorbid depressive symptoms in a sample of college women (i.e., significant reductions in scores on a self-report measure of depressive symptomatology; $d = 0.96$; Taylor et al., 2016). In sum, online adaptations of CBT-gsh have demonstrated efficacy in treating eating disorders and represent a promising secondary-prevention program for targeting eating- and mood symptoms in college students.

Mobile Applications. Mobile applications represent a logical next-step for CBT-gsh delivery that may overcome some of the service-utilization barriers reported by college students by improving accessibility and efficiency of treatment delivery. With this format, students could overcome the treatment barriers of cost, accessibility, and time by having access to a low-cost or free treatment that is accessible virtually anywhere and demands less of students' time relative to other treatments. Mobile health (m-health) interventions are also highly feasible in college students, given that 96% of people age 18-29 own a Smartphone (Pew Research, 2019).

Only a few studies to-date have examined mobile-assisted CBT-gsh for eating disorders; however, results are promising. For example, Hildebrandt and colleagues (2017) first examined the use of Noom Monitor, a smartphone application designed to streamline the self-monitoring and food logs components of CBT-gsh, as an adjunct to traditional, in-person CBT-gsh for eating disorders. They found that adults with bulimia nervosa or binge-eating disorder who received CBT-gsh + Noom Monitor (*n*=33) had greater reductions in objective binge eating and purging during treatment, and greater mealtime adherence compared to their counterparts who only received CBT-gsh (*n*=33; Hildebrandt et al., 2017). At six-month follow-up, both the CBT-gsh + Noom Monitor and CBT-gsh groups demonstrated similar remission and abstinence rates, suggesting comparable outcomes between the two approaches over time.

Hildebrandt and colleagues (2020) extended the 2017 study by comparing adults who received standard care (*n*=111) to adults who received CBT-gsh delivered via telemedicine and paired with the Noom Monitor for self-monitoring and food logs (*n*=114) in a randomized controlled trial (RCT) of adults with binge-type eating disorders.[1] The CBT-gsh + Noom Monitor group demonstrated superior treatment outcomes (i.e., statistically significant reductions in binge-eating days [*d* = -1.46] and clinical impairment [*d* = -2.75], as well as higher remission rates [56.7%]) compared to the adults who received standard care (remission rate = 30%; Hildebrandt et al., 2020). Taken together, the two Hildebrandt and colleagues (2017 & 2020) CBT-gsh + Noom Monitor studies represent important first-steps in demonstrating the efficacy of CBT-gsh for eating disorders assisted by mobile application. Fully mobile applications of

[1] Grilo (2020) raised concerns regarding the interpretation of the Hildebrandt and colleagues (2020) findings, considering the comparison group received "standard care" (i.e., persons in the comparison group received no treatment, which represents one of the the weakest types of comparison conditions for an RCT). As such, Grilo cautioned interpretation of the findings because comparison to a no-treatment group can magnify the effects demonstrated by the experimental group.

CBT-gsh—where participants complete all treatment materials via mobile application—represent a timely next-step in eating-disorder treatment development, given their potential to increase access and reduce costs and in-person resources associated with in-person eating-disorder treatment (Kazdin et al., 2017).

Only one study to-date has implemented a mobile CBT-gsh intervention for eating disorders (Fitzsimmons-Craft et al., 2019), and the sample used in this study consisted of college students with eating disorders at several large universities in Missouri. The mobile CBT-gsh implemented in the Fitzsimmons-Craft and colleagues (2019) study consisted of 63 online sessions that each lasted 5-10 minutes over the course of eight months. Participants were able to communicate with their "health coach" via the mobile application. Following a mass eating-disorder screening initiative at universities across the state of Missouri, n=145 students with eating disorders engaged with the mobile CBT-gsh (47.8% of the total number of students eligible; average number of sessions completed = 7.48). Students with eating disorders who engaged with the mobile CBT-gsh demonstrated significant decreases in binge eating and restricting over their time in treatment (Cohen's d = 0.39 [binge eating]; Cohen's d = 0.27 [restricting]). Students who screened positive for an eating disorder attended more sessions than is typical in college counseling centers (college counseling center average = 4.7 sessions; mode = 1 session; Center for Collegiate Mental Health, 2016). The findings from the Fitzsimmons-Craft and colleagues (2019) study provided preliminary support for the effectiveness of mobile CBT-gsh for treating eating disorders in college students. Furthermore, the results indicated that mobile interventions may serve to engage college students more so than traditional face-to-face approaches. In sum, mobile CBT-gsh has great potential to be an evidence-based, cost-effective,

and easily accessible treatment for eating disorders in college students that warrants further study.

Overview of Current Study

Eating disorders are accompanied by myriad negative outcomes and treatments specifically for eating disorders can be difficult to access, particularly on college campuses that often do not have specialty care. Mobile CBT-gsh approaches represent promising treatment frameworks that warrant further research and that, ultimately, could lead to improved longevity and quality-of-life for persons with these serious mental illness. Thus, the purpose of this study was to administer a mobile CBT-gsh to college students with eating disorders using a multiple-baseline design. Specific aims of the study included:

Study Aim 1 (Primary Outcome): Test the efficacy of a novel, mobile CBT-gsh intervention for reducing overall eating-disorder psychopathology within-person over time using a multiple-baseline design

Hypothesis 1: Participants will demonstrate reductions in eating-disorder psychopathology, as documented by significant within-person reductions of Total Composite scores on a self-report eating-disorder measure.

Study Aim 2 (Secondary Outcome): Test the efficacy of a novel, mobile CBT-gsh intervention for reducing specific eating-disorder behaviors and cognitions within-person over time using a multiple-baseline design

Hypothesis 2: Participants will demonstrate reductions in eating-disorder behaviors, as documented by significant within-person reductions of Binge Eating, Restricting, Purging, Excessive Exercise, and Body Dissatisfaction scores on a self-report eating-disorder measure.

Study Aim 3 (Secondary Outcome): Test the efficacy of a mobile CBT-gsh intervention for reducing psychiatric impairment due to an eating disorder

Hypothesis 3: Participants will have decreased within-person scores on a self-report measure of psychiatric impairment secondary to an eating disorder from pre- to end-of-treatment.

Study Aim 4 (Secondary Outcome): Test the efficacy of a mobile CBT-gsh intervention for reducing internalizing psychopathology

Hypothesis 4: Participants will have improved within-person scores on measures of comorbid psychopathology (i.e., self-reported general depression, panic, social anxiety symptoms, negative affect, and positive affect) at Week 8 of the intervention phase compared to pre-treatment.

Study Aim 5 (Secondary Outcome): Collect qualitative data on user experiences

Information gathered about user experience was collected to characterize participant satisfaction with the treatment and mobile application. No *a priori* hypotheses were made for Study Aim 5.

Methods

Dataset and Participants

Participants

The present study's inclusion/exclusion criteria were informed by Medical Standards published by the Academy for Eating Disorders (Academy for Eating Disorders, 2016) to ensure that participants were medically and psychological suitable to participate in outpatient, guided self-help therapy.

Inclusion/Exclusion Criteria. Inclusion criteria were: 1) Age ≥ 18 years; 2) Enrolled as a KU student; 3) *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (*DSM-5*;

American Psychiatric Association, 2013) eating-disorder diagnosis; 4) owned a smartphone; and 5) no uncorrected vision problems that would interfere with their ability to participate in the study via a mobile application.

Exclusion criteria were: 1) Age < 18 years; 2) not enrolled as a KU student; 3) no *DSM-5* eating-disorder diagnosis; 4) positive screen for anorexia nervosa and/or low body weight (i.e., Body Mass Index < 19.0); 5) presence of current moderate/severe suicidal ideation; 6) other significant psychopathology that could have interfered with treatment (e.g., severe obsessive-compulsive disorder or current substance-use disorder); 7) medical conditions that could have interfered with treatment (e.g., Type 1 diabetes mellitus, current pregnancy); and 8) medical instability that required a higher level of care (e.g., low heart rate, electrolyte disturbance, hypothermia, orthostasis, hypoglycemia, acute medical complications of malnutrition [cardiac issues, pancreatitis, seizures, syncope]).

Persons with anorexia nervosa or who were underweight were excluded because CBT is not the recommended first-line treatment for anorexia nervosa in adults (Hay, 2013; National Collaborating Center for Mental Health, 2004) and because of the medical risks associated with low body weight that require more careful, in-person monitoring (Mitchell & Crow, 2006; Academy for Eating Disorders, 2016). Thus, persons who screened positively for anorexia nervosa and/or had a low body weight were excluded from the present study and provided with a referral for a higher level-of-care (e.g., in-person outpatient therapy, partial-hospitalization program, etc.). Persons who were unwilling to complete a pre-intervention medical examination (to ensure appropriateness of outpatient CBT-gsh) and/or were unwilling to allow sessions to be videotaped (to ensure therapist adherence and the ability to provide supervision by Licensed Psychologists) were excluded.

Recruitment

Participants were recruited from the student body at KU. Recruitment for the present study (titled, "*Multiple-baseline Intervention for College-student Eating [MICE]*") was paired with recruitment for a larger RCT (titled, "*Building healthy Eating and Self-Esteem Together for University students [BEST-U]*"); both studies used the same inclusion/exclusion criteria and treatment materials. All study procedures were approved by the KU Institutional Review Board and participants provided informed consent prior to engaging in any study-related procedures.

Screening and Recruitment. The total population of students enrolled at KU in Spring 2020 was *N*=23,160 (including undergraduate and graduate students at the Lawrence and Edwards campuses; University of Kansas, 2020). Our research team collaborated with the Office of Student Affairs to send recruitment materials to a random sample of KU students (*n*=2,000). Randomly selected KU students were sent an email that invited them to participate in a free, online screening to determine eligibility for a paid treatment study. In addition, the research team advertised using brochures, posters, clinician referral, word-of-mouth, and social media. All screening questions (see *Measures* below) were completed via REDCap, a Health Insurance Portability Accountability Act of 1996 (HIPAA)-compliant electronic survey and data management system.

Eligible students who screened positive for an eating disorder completed a psychological evaluation to confirm an eating-disorder diagnosis and study eligibility; they also completed a medical evaluation to ensure medical stability (see **Figure 2**). Students completed the psychological evaluation in-person or via telehealth. They had their height and weight measured by our research team if they completed the psychological evaluation in-person or during their medical evaluation if they completed the psychological evaluation via telehealth.

Procedure

Novel coronavirus

The present study was affected to some degree by the acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic. During the SARS-CoV-2 pandemic, federal- and state-mandates (e.g., stay-at-home orders) were effected to slow the spread of the disease. KU implemented safety precautions, including moving all courses online for the remainder of the semester (March 17) and suspending non-essential research activities on campus (March 23). As such, minor adjustments had to be made to the present study procedures (all modifications were approved, in advance of implementation, by the Institutional Review Board). The study team paused recruitment during the initial stages of the pandemic development in the United States, the psychological evaluation was shifted to telehealth, and participant weights were confirmed at the medical evaluation (rather than in the research laboratory during the psychological evaluation). The present study design was largely unaffected, however, because the questionnaires and treatment materials were originally designed to be delivered via mobile application and telehealth. The study team followed state guidelines as they evolved regarding the use of telemedicine across state lines.

Methodology

A non-concurrent multiple-baseline design was used to measure changes in symptomatology over time. The goal of single-subject designs is to demonstrate that a change in target behavior is caused by an intervention. Single-subject designs employ two or more *phases* to examine behavior; a phase is a series of observations that employ the same conditions. The first phase in a single-subject design is the *baseline phase*. The goal during the baseline phase is to establish stability in the measurements of the target behavior(s) before an intervention is

applied. Measurement stability can be established in the form of measurement *level* (i.e., scores are relatively similar with little variability from measurement to measurement) or measurement *trend* (e.g., scores demonstrate a consistent increasing or decreasing trend from measurement to measurement).

The next phase in a single-subject design is the *intervention phase*. During the intervention phase, an intervention (e.g., treatment) is applied in attempt to effect change in the target behavior(s). The goal during the intervention phase is to demonstrate a significant change in the target behavior(s) from the baseline phase (i.e., before the intervention was introduced). Researchers look for a large and/or immediate change in the target behavior(s) from the baseline to intervention phase to establish that an effect has occurred.

Multiple-baseline designs are an extension of the single-subject design and are used to demonstrate causal effects between independent and dependent variables within and across participants (Horner et al., 2005). In multiple-baseline designs, the independent variable (i.e., intervention or baseline phase) is actively manipulated to allow for experimental control of the intervention effect, and the dependent variables (i.e., eating-disorder and general psychopathology, clinical impairment) are measured repeatedly across time.

Multiple-baseline designs seek to replicate treatment effects across participants, behaviors, or settings (Kratochwill et al., 2010). In the current study, I was interested in testing treatment effects across participants. I accomplished this by employing baseline- and intervention phases for each participant individually. Specifically, in multiple-baseline designs (1) treatment was administered only after establishing a baseline for each participant independently and (2) treatment administration was staggered over time across participants. The baseline-phase length was different for each participant in an attempt to demonstrate the

treatment effect at varying time points; if treatment effects are demonstrated across time points, the argument to support the treatment having caused changes in the target behavior is strengthened.

In multiple-baseline designs, comparisons are made within- and between-subjects to control threats to internal validity (Kratochwill et al., 2010; Kratochwill et al., 2013). Within-subject comparisons demonstrate whether an effect has taken place from the baseline- to intervention phases. In other words, researchers look for a noticeable difference in the target behavior between the baseline phase (when the intervention is absent) and the intervention phase (when the intervention is applied). Treatment is only administered after within-person measurements of the desired behavior(s) are stable during the baseline phase; thus, each participant serves as their own control for confounding variables.

Between-subjects comparisons are made by testing whether the effect can be demonstrated in multiple, separate participants. Replication of treatment effects across participants is necessary to increase external validity (Horner et al., 2005). That is, if an effect can be demonstrated in two or more participants, the effect can be said to be generalized across participants and external validity is strengthened with each subsequent replication. The administration of treatment after establishing a baseline for each participant and staggering the treatment administration over time across participants represents active manipulation of the independent variable.

Data Points and Participant Grouping. Methodologists recommend a minimum of three data points per phase (i.e., baseline and intervention phases) to demonstrate treatment effects in a multiple-baseline design (Kratochwill et al., 2013); this recommendation was met and/or exceeded by administering a minimum of three and a maximum of five baseline

assessments and n=10 weekly assessments during the intervention phase for each participant (total $N_{assessments}$=13-15 per participant).

Participants were grouped in a triad such that the triad completed the treatment non-concurrently, with time-staggered treatment administration within the triad. For each participant (N=3), treatment was administered following a visual analysis of stability in baseline measurements of relevant eating-disorder psychopathology. I computed a Total Composite Score from the Binge Eating, Purging, Restricting, and Excessive Exercise scales from the Eating Pathology Clinical Outcomes Tracking (EPCOT; see *Measures* below) because these behaviors map onto CBT-gsh targets. Furthermore, we expected these scales to remain relatively stable during the baseline phase and to exhibit a decrease following introduction of the treatment.

Visual analysis is the recommended approach to data analysis for single-case designs (Kratochwill et al., 2013). Visual analysis calls for examination of six features to assess treatment effects: (1) level; (2) trend; (3) variability; (4) immediacy of the effect; (5) overlap; and (6) consistency across similar phases (Kratochwill et al., 2010; Vannest & Ninci, 2015). *Level* refers to mean values from each phase. *Trend* refers to potential decreasing/increasing patterns in the baseline data that may affect intervention effects. *Variability* refers to the observed stability and narrowness of the range of values in each phase. *Immediacy of the effect* refers to the comparison of the mean of the last three values of the baseline phase to the mean of the first three values of the intervention phase. *Overlap* refers to the number of overlapping values from each phase when compared to the other phase. *Consistency across similar phases* refers to comparisons of similar phases across participants (i.e., comparison of the baseline phases across participants, comparison of the intervention phases across participants). N=3 participants were chosen because interpretations of treatment effects would be stronger if they

could be replicated across participants (Kratochwill et al., 2010); *N*=3 meets recommended standards for multiple-baseline designs, which call for a minimum of three participants (Kratochwill et al., 2013).

Threats to Internal Validity. There are several threats to internal validity that were considered. *Maturation* refers to change within a participant that may or may not be attributable to the intervention (e.g., fatigue, acquiring new skills as one ages). *History* can be thought of as influences outside of the intervention that may influence participants (these may happen before or during their participation); history events are often unpredictable and difficult to control (e.g., stress related to SARS-CoV-2 and self-isolation). I attempted to address these threats to internal validity by staggering treatment administration over time across participants, and by extending the baseline phase for a participant if necessary (i.e., if indicated by visual analysis of the data). Staggering treatment administration and extending the baseline phase helped ensure that maturation or history effects could be ruled out as an explanation of a treatment effect because treatment effects were observed at different points in time across participants. *Instrumentation* refers to changes in the properties of the measurements used in the study, and *testing* is a participant's reaction to repeated assessment. Instrumentation threats were not a significant concern because consistency in the assessment questions and measurement delivery were consistent across measurement time points. Testing effects (i.e., learning effects) were not a significant concern in the present study because the format and purpose of the study were known to participants when they were enrolled in the study. Experimental fatigue was not a significant concern because of the relatively low demands of engagement with the study materials (i.e., less than one hour, once per week); additionally, participants could choose to interact with the study materials whenever they wanted within the week timeframe.

Finally, *statistical regression* ("regression toward the mean") refers to the movement of a participant's score on the dependent variable toward the population mean score for that variable. This is particularly likely for persons with extreme scores on the dependent variable. Statistical regression can also be thought of as participant improvement on any dependent variable that could be expected if the intervention had not been implemented. For example, in treatment-seeking samples, it is reasonable to expect improvement on part of the participant merely by enrolling in treatment (i.e., before the intervention has been implemented and completed; DeYoung & Bottera, 2018). To address this threat to internal validity, repeated measurements were taken throughout the baseline and intervention phases and treatment administration was staggered across participants.

Rationale for Choosing Multiple-Baseline Design. Single-case designs—and multiple-baseline designs in particular—are uncommon in the eating-disorders literature (DeYoung & Bottera, 2018). Research employing multiple-baseline designs for eating-disorder interventions represents a practical opportunity to expand scientific knowledge in the field and represents a fitting methodology for the study of eating-disorder interventions given common characteristics of eating-disorder research samples (e.g., high rates of psychiatric- and medical comorbidity, relatively small sample sizes, complexity of eating-disorder clinical presentations; DeYoung & Bottera, 2018).

Multiple-baseline designs are ideal for testing research questions in clinical settings (Kratochwill et al., 2010; DeYoung & Bottera, 2018). Multiple-baseline designs can fall anywhere along the continuum from effectiveness to efficacy. Thus, multiple-baseline designs can be useful in evaluating the effects of a novel treatment. Another advantage of multiple-baseline designs is that they can be used when there are fewer resources, such as limited time or

funding, because fewer participants are needed to demonstrate treatment effects. Given that fewer resources and participants are needed to demonstrate treatment effects, multiple-baseline designs are ideal designs for pilot studies and proof-of-concept treatments before moving forward with an RCT.

Multiple-baseline designs allow for interpretations of change while being a feasible way to test a pilot treatment. Multiple-baseline designs are adaptive and flexible and, in turn, involve a planned degree of uncertainty (e.g., treatment administration varies across time and across participants). Multiple-baseline designs are especially well-suited to identify treatment effects for each participant involved, effects which otherwise may be "washed out" or missed in group-based designs. Although I considered an RCT, given this was a pilot study, a multiple-baseline design provided estimates of treatment effects that can be used to inform a larger RCT in future research, while still maximizing internal validity. Indeed, single-case designs such as multiple-baseline designs represent a necessary first-step to demonstrate treatment effects that inform empirically supported treatments (Horner et al., 2005; Kratochwill et al., 2013; DeYoung & Bottera, 2018).

The National Institutes of Health Obesity-Related Behavioral Intervention Trials (ORBIT) model (Czajkowski et al., 2015) offers a useful framework for conceptualizing the current study. The ORBIT model was designed to inform early-stage treatment development for the management of chronic health conditions. The ORBIT model itself was informed by models of drug research and adapted to fit behavioral-intervention research. According to the ORBIT model (see **Figure 3**), *Phase I* involves identification and definition of a problem area (i.e., a research question; *Phase Ia*), and design of the essential features of the intervention for said problem (*Phase Ib*); *Phase II* involves proof-of-concept research (*Phase IIa*) and pilot studies

(*Phase IIb*), which are necessary before moving to *Phase III* (efficacy trials) and *Phase IV* (effectiveness studies; Czajkowski et al., 2015). The ORBIT model recommends systematic literature review as one method for satisfying *Phase Ia*; I fulfilled this recommendation by completing a comprehensive literature review, which directly informed the current study. *Phase Ib* of the ORBIT model requires refinement of the treatment materials, which the present study team completed over the course of six months with assistance from several eating-disorder-treatment specialists. The current study was best characterized as a *Phase IIa* study because the goal was to evaluate CBT-gsh in a sample of college students with eating-disorders. The goal of *Phase IIa* studies is to produce clinically significant effects in a small sample to inform whether it is appropriate to move on to testing the intervention in larger samples (Czajkowski et al., 2015).

Study Flow

The present study included several sequential steps (see *Measures* below and **Table 2** for description of assessment tools used for screening and intake purposes; see **Figure 4** for a visual description of the study flow):

1. ***Screening:*** Participants completed an initial online screening via REDCap. There was an informed-consent statement preceding the survey and participation in the survey indicated willingness to participate in the study and that the participant was over the age of 18 years. If the participant was eligible based on their online-screening responses, they were scheduled to attend an in-person or telehealth psychological evaluation.

2. ***Psychological Evaluation:*** Participants completed the psychological evaluation in-person or through a HIPAA-compliant video-conferencing application ("Healthie"). The psychological evaluation began with the informed consent. Participants who completed the

psychological evaluation through Healthie were sent consent forms prior to initiating the telehealth session and were required to provide signed copies to the research team prior to beginning study procedures. After providing consent to participate in the study, the severity of participants' eating-disorder behaviors, comorbid psychopathology, suicidal thoughts (if any), weight history, medical history, social and environmental influences on their eating behaviors, cultural barriers or resources, family history of mental illness, and goals were assessed. At the end of the psychological evaluation, if the participant continued to be interested and eligible for the study, they were asked to complete a medical evaluation to ensure physical health and appropriateness for outpatient, guided self-help therapy.

3. ***Medical Evaluation:*** The medical evaluation assessed for exclusion criteria (see *Inclusion/Exclusion Criteria* above; Academy for Eating Disorders. 2016).

4. ***Study Enrollment:*** Participants who completed and were deemed eligible based on the first three steps above were enrolled in the baseline phase of the present study.

a. *Baseline Phase:* Participants completed weekly assessments until stability in measurements of the target behaviors was observed by visual analysis, after which they were moved to the intervention phase. If no stability in measurements was observed during the baseline phase, participants were moved to the intervention phase after five weeks.

b. *Intervention Phase:* The CBT-gsh material was grouped into sequential themes ("modules"); all treatment materials were available via the mobile application (see **Appendix A** for screenshots of the mobile application). Participants completed 1-2 treatment modules per week over the course of 10 weeks. Each week during the intervention phase, the participant met with their "health coach" for 25-30 minutes; the

"health coach" was a graduate student or postdoctoral researcher who completed a semester-long course on the Treatment of Eating Disorders at KU or completed the 16-week University of Oxford online CBT-E training (Centre for Research on Eating Disorders at Oxford [CREDO], 2020). Coaching sessions occurred via telephone, in-person, or via the HIPAA-compliant telemedicine video chat function of Healthie. On a weekly basis, group supervision was provided by two licensed psychologists (Drs. Kelsie Forbush and Sara Gould [Clinical Psychologist, Children's Mercy Eating Disorders Center, and KU Faculty Affiliate]) to ensure that participants were progressing appropriately, answer questions, and ensure that therapists were maintaining treatment adherence. Dr. Forbush also provided individual supervision of health coaches on a regular basis.

c. *End-of-Treatment:* Participants completed a survey about their experiences in the study.

Cognitive-Behavioral Treatment Materials

Participants (*N*=3) completed a 10-week course of mobile CBT-based modules that targeted eating-disorder behaviors and related cognitions (see screenshots of a treatment module in **Appendix A**). A 10-week treatment course was consistent with recent research supporting the effectiveness of a 10-session CBT for non-underweight eating disorders (Rose & Waller, 2017; Waller et al., 2018). The information presented in the modules was based on evidence-based approaches and developed at KU (Principal Investigator: Dr. Kelsie Forbush).

Consistent with traditional CBT-E, the first part of the mobile CBT-gsh was intensive. Participants completed two treatment modules per week during the first four weeks of the intervention phase (and one module per week for the remaining six weeks). During the first

week, self-monitoring of eating-disorder behaviors and weighing was introduced. Participants were expected to self-monitor daily and weigh themselves weekly. During the second week, participants learned about the consequences of disordered eating and were asked to commit to refraining from "unhelpful behaviors" (e.g., purging). During this week they were also asked to engage in regular eating (i.e., three meals and three snacks per day). These behavioral components were continued for the rest of the mobile CBT-gsh.

The mobile CBT-gsh also: 1) Included experiments to challenge thoughts; 2) provided exposure to avoided foods; 3) taught skills to help analyze and disrupt unhelpful patterns (e.g., behavior-chain analysis for problem behaviors; "urge surfing" to resist urges to engage in eating-disorder behaviors); and 4) provided psychoeducation about other relevant topics, such as stressors related to college life, the importance of self-care, and relapse prevention. Participants were asked to challenge cognitive distortions related to their eating disorder (e.g., automatic thoughts that prompt eating-disorder behaviors). They also learned about and were asked to challenge cognitions related to over-evaluation of their body weight/shape and negative body image and to monitor and reduce body checking. In addition, the treatment modules asked for participant feedback about their therapy experiences and asked them to reflect on what they were learning.

Mobile Application

The online materials were available via a mobile application created through the MEI Research Platform® (https://pilrhealth.com/), a mobile-based tool for data collection (similar to REDCap; see **Appendix A** for sample screenshots of the mobile application). The MEI Research Platform® mobile application and server are compliant with HIPAA guidelines and participants were provided a researcher-generated, unique identifier and password for use in the study. No

participant data was stored within the mobile application; thus, if another person stole or "hacked" a participant's phone, they would not be able to access any study data, questions, or materials.

The MEI Research Platform® provided the opportunity to monitor study adherence. Frequency and duration of sign-ins to the treatment application and interaction with the materials by participants was monitored daily (e.g., I could see what materials they accessed and when). Furthermore, study payment was based on completeness of study components, so I was able to monitor adherence when evaluating whether participants received payment on a weekly basis.

Study Outcomes. See **Table 1** for study aims, hypotheses, and outcomes. The primary outcome was overall eating-disorder psychopathology, measured by a Total Composite score from the EPCOT (unpublished; see *Measures* below). Secondary outcomes included: (1) Eating-disorder behaviors and cognitions, measured by the Binge Eating, Restricting, Purging, Excessive Exercise, and Body Dissatisfaction scales of the EPCOT; (2) clinical impairment, measured by the Clinical Impairment Assessment (CIA; Bohn et al., 2008); and (3) Internalizing psychopathology, measured by the Positive and Negative Affect Schedule-Expanded Form (PANAS-X; Watson & Clark, 1999) and the Inventory of Depression and Anxiety Symptoms, Second Edition (IDAS-II; Watson et al., 2011).

Dependent Variables. Baseline- and intervention-phase measurements of the EPCOT Total Composite score (primary outcome) were gathered weekly (see *Measures* below and **Table 2**). Because the individual scales each contained different numbers of items within the scale, EPCOT scores for the targeted behaviors (i.e., EPCOT Binge Eating, Restricting, Purging, Excessive Exercise) were totaled and averaged to create a composite score for each scale. I also created a Total Composite Score by summing and averaging the composite scores for each of the

four scales. This Total Composite Score was used to establish stability in the baseline phase prior to moving participants to the intervention phase. EPCOT Body Dissatisfaction (secondary outcome) was examined separately from the behavioral EPCOT scales (i.e., Binge Eating, Restricting, Purging, Excessive Exercise) because it measures eating-disorder cognitions that are targeted by the CBT-gsh. Participant CIA (secondary outcome) was measured in the screening questionnaire and at end-of-treatment. Participant PANAS-X and IDAS-II (secondary outcomes) were measured during the psychological evaluation and during weeks four and eight of the intervention phase.

Participant Compensation. Participants earned $50.00 for the initial two-hour psychological evaluation and medical evaluation. They received a minimum of $50.00 for their participation (i.e., if they completed the psychological/medical evaluations and no other study components). The maximum amount participants earned was variable, as each participant's baseline phase length varied. The minimum total payment a participant could earn if they completed all study components was $164.00 ($50.00 for the psychological/medical evaluations, $114.00 for the minimum baseline phase length = three weeks, plus 10 weeks for the intervention phase; see example in **Table 3**). The maximum total payment a participant could earn if they completed all study components was $168.00 ($50.00 for the psychological/medical evaluations, $122.00 for the maximum baseline phase length = five weeks, plus 10 weeks for the intervention phase).

Payment was contingent upon participation. I used a contingency-management framework for participant payment. Contingency management is a widely used and well-validated approach for clinical and behavioral research (Petry et al., 2000; Roll & Shoptaw, 2006; Stanger et al., 2016). Contingency management has been shown to increase participant

adherence to study protocols, study retention, and participant benefit and to decrease study dropout (Petry et al., 2000; Roll & Shoptaw, 2006; Stanger, Lansing, & Budney, 2016). Contingency management pays participants for the study components they complete and earning potential increases over time. Within the contingency-management approach, participants were free to choose not to complete any study component.

During the baseline phase, participants earned $2.00 for each week they completed the weekly surveys. If participants chose not to complete a weekly survey in the baseline phase, they earned $0.00 for that week. Participants earned a bonus $10.00 for completing the baseline phase (i.e., once they moved to the intervention phase, they completed the baseline phase and were given the $10.00 bonus).

The intervention phase included a graded pay schedule. To earn payment each week, participants must have completed all study components for that week (i.e., read the treatment module within the app, completed the weekly surveys). The reinforcement magnitude started at $4.00 and increased by $0.50 for each completed week during the first five modules of the intervention phase, and by $1.00 for each completed week during the second five modules of the intervention phase. To encourage continued participation, the amount participants could earn during the second five weeks of the intervention phase was increased by $2.00 (i.e., the amount participants could earn in week five of the intervention phase was increased to $9.00 and then increased by $1.00 each week). In addition, participants earned $4.00 for completing the expanded surveys administered in weeks four and eight of the intervention phase. If participants chose not to complete a study component, they earned $0.00 for that week and their contingency was then reset to the original reinforcement magnitude ($4.00). After the participant demonstrated consistency in completing study components (i.e., completed two consecutive

study weeks), their reinforcement magnitude "caught up" to where it was had they not missed a week. Participants earned a $10.00 bonus at completion of the intervention phase (i.e., completion of all 10 modules) regardless of whether they chose not to complete a study component. In other words, participants earned the $10.00 bonus for completing all 10 treatment modules even if they missed a week (so long as they went on to complete all treatment modules). **Table 3** outlines the reinforcement schedule.

It is important to note that participant payment was based on completeness of the study components, not based on quality (i.e., participants were paid as long as they completed the study components with reasonable effort and their payment was not contingent upon assessment scores, response to treatment, etc.). We did not reimburse participants directly for medical costs associated with their medical evaluation (e.g., we did not pay the participant's co-pay or out-of-pocket costs for their health-care visit or labs).

Measures

See **Table 2** for the assessment schedule and **Appendix B** for copies of the questionnaires; participants were asked to complete self-report measures weekly for the duration of study enrollment, and daily during the intervention phase.

Demographics. Demographic information, including age, gender, sexuality, and race/ethnicity was gathered from participants in the screening questionnaire and psychological evaluation.

Body Mass Index (BMI). To characterize the sample and rule out anorexia nervosa, self-reported height and weight were assessed in the screening questionnaire. Participant height and weight measurements were taken in-person during either their psychological or medical

evaluation (after completing a signed Release of Information). BMI values were calculated using BMI = 703 x weight (lbs)/height (in)^2.

Eating Pathology. The Eating Pathology Clinical Outcomes Tracking (EPCOT; unpublished) is a 45-item self-report measure of eating-disorder psychopathology over the past week. The EPCOT was created based on the Eating Pathology Symptoms Inventory (EPSI; Forbush et al., 2013), which contains the same items as the EPCOT and measures symptoms over the past month. An unpublished study demonstrated that the EPCOT had a good fit to the EPSI factor structure (K. Forbush, personal communication, May 15, 2020). The EPSI demonstrated a robust eight-factor structure in clinical samples and seven-to-eight-factor structures in non-clinical samples (Forbush et al., 2013; Forbush et al., 2014; Coniglio et al., 2018). Additionally, the EPSI demonstrated evidence for excellent internal consistency (Cronbach's alpha = 0.84-0.89), two-to-four week test-retest reliability (mean r = .73), and convergent and discriminant validity in varying samples and across binary genders (Forbush et al., 2013; Forbush et al., 2014; Coniglio et al., 2018).

Eight scales comprise the EPCOT (Binge Eating, Restricting, Excessive Exercise, Purging, Body Dissatisfaction, Cognitive Restraint, Muscle Building, Negative Attitudes Toward Obesity) and items are scored on a four-point scale (0 = never, 4 = very often). For the purpose of the present study, I converted scale scores for the behaviorally oriented EPCOT scales (Binge Eating, Restricting, Excessive Exercise, and Purging) to average composite scores by summing the items in the scale and dividing each sum by the number of items in the respective scale. I chose these four scales because they map onto CBT-gsh targets and because I expected scores to be affected significantly by the intervention. I also created a Total Composite score by summing the composite scores for each of the four scales and dividing the sum by four. I chose to create

the Total Composite score because I did not expect to observe much change in it during the baseline phase (i.e., I expected it to be stable) and I expected the Total Composite score to decrease during the intervention phase. I also examined EPCOT Body Dissatisfaction because of the scale's relevance to the CBT-gsh model and because I expected Body Dissatisfaction scores to be impacted by the intervention. The EPCOT was administered during the psychological evaluation and weekly during the baseline- and intervention phases of the present study.

The Eating Disorder Diagnostic Scale (EDDS; Stice et al., 2000) is a brief, self-report measure of *DSM* diagnostic criteria for eating disorders. The EDDS demonstrated good test-retest reliability ($r = 0.87$) over a one-week period, and good internal consistency (Cronbach's alpha = 0.89) in large samples of females; evidence for convergent and predictive validity were also established (Stice et al., 2004; Stice et al., 2000). The EDDS was administered to participants as part of the screening questionnaire to assess participants' eating-disorder behaviors and cognitions.

Clinical Impairment. The Clinical Impairment Assessment (CIA; Bohn et al., 2008) is a 16-item self-report measure of personal, cognitive, and social impairment associated with an eating disorder over the past 28 days. Items were scored on a four-point Likert scale (0 = no impairment; 3 = severe impairment) and summed to create a composite score (CIA range = 0-48). A cut-score of 16 (sensitivity = 76%; specificity = 86%) was recommended (Bohn et al., 2008). The CIA demonstrated good test-retest reliability over a 20-week period (intra-class correlation coefficient = 0.86). Significant correlations between CIA scores and clinician ratings of impairment ($r = 0.68$) provided evidence for construct validity; significant correlations between the CIA and Eating Disorder Examination Questionnaire global scores ($r = 0.89$) provided evidence for convergent validity. The CIA also demonstrated sensitivity to change from

pre- to post-treatment in the sample with which it was validated, as exhibited by a significant decrease in CIA global scores ($p < .0001$) and a statistically significant correlation between change in CIA global scores and clinician-rated impairment ($r = 0.86$; Bohn et al., 2008). The CIA was administered as part of the screening questionnaire, the surveys administered in weeks four and eight of the intervention phase, and the post-study survey to assess impairment associated with participants' eating-disorder behaviors and cognitions.

Internalizing Psychopathology. The Inventory of Depression and Anxiety Symptoms, Second Version (IDAS-II; Watson et al., 2011) is a 99-item self-report measure of symptoms of internalizing psychopathology (i.e., depression, bipolar, and anxiety symptoms) over the past two weeks. Items were rated on a five-point Likert scale (0 = never, 4 = very often) and summed to create a higher-order General Depression scale and 18 subscales. The General Depression, Panic, and Social Anxiety scales were tested in the present study. The IDAS-II demonstrated strong evidence for internal consistency (Cronbach's alpha = 0.72-0.90), and criterion-related, convergent, and discriminant validity compared to other self-report and interview-based assessments of internalizing psychopathology in clinical and non-clinical samples (Watson et al., 2011). The IDAS-II was administered during the psychological evaluation and in weeks four and eight of the intervention phase.

The Positive and Negative Affect Schedule—Expanded Form (PANAS-X; Watson & Clark, 1999) is a 60-item self-report measure of positive- and negative-affect symptoms during the "past few weeks." Items were rated on a five-point Likert scale (1 = very slightly or not at all, 5 = extremely) and comprised two higher-order scales (General Positive Affect, General Negative Affect) and 11 specific affects. The two higher-order scales were tested in the current study. The PANAS-X demonstrated excellent internal consistency (Cronbach's alpha = 0.82-

0.90) and convergent and discriminant validity in varying clinical- and non-clinical samples (Watson & Clark, 1999). The PANAS-X was administered during the psychological evaluation and in weeks four and eight of the intervention phase.

Participant Satisfaction. The study research team created an *ad-hoc* participant satisfaction survey that participants completed at the end of treatment. The participant satisfaction survey consisted of 17 questions about their satisfaction with the treatment and mobile application that were rated on a seven-point Likert scale (1 = Strongly Disagree, 7 = Strongly Agree). Participants were also asked two open-ended questions about what they liked best and what they would change about the treatment and/or mobile application.

Visual and Statistical Analyses

Researchers have historically relied upon visual analysis and strong internal validity of single-subject-designed studies to interpret intervention effects (Olive & Smith, 2005; Smith, 2012). They argued that applied clinical importance and applicability were more important to single-subject research than statistical significance (Olive & Smith, 2005). Furthermore, data from single-case designs often violate assumptions on which many statistical tests are based (Olive & Smith, 2005). Recently, researchers have recommended supplementing visual analysis with statistical analyses to estimate effect sizes for the treatment effects (Smith, 2012). Although there are several statistical tests that can be used to supplement visual analysis, no consensus exists regarding which approach is best (Smith, 2012), as each has advantages and disadvantages relative to other statistical tests that can estimate treatment effect sizes. Tau-*U* represents a non-parametric statistic that provides an estimate of effect size and statistical significance (Parker et al., 2011; Tarlow, 2017). Advantages of Tau-*U* include its ability to control for statistically significant baseline trends that would affect treatment effects and its suitability for single-case

data (Parker et al., 2011; Tarlow, 2017; Vannest & Ninci, 2015). Tau-*U* effects are bound between +/-1.00 and can be interpreted as small (Tau-*U* = 0.00-0.20), moderate (Tau-*U* = 0.21–0.60), large (Tau-*U* = 0.61–0.80), or very large (Tau-*U* > 0.80).

To address the primary and secondary outcomes, I visually analyzed the baseline- and intervention-phase data for each dependent variable (Study Aims 1-2). To supplement the visual analysis, I used a web-based application to calculate Tau-*U* (Vannest et al., 2016). Means for each phase were calculated using the last three EPCOT scores for each scale and these means were compared to normed means (Study Aim 2). For Participants 1 and 2, I compared their means to normed means for college women. I compared Participant 3's means to normed means for general psychiatric outpatients because he was in graduate school and was older than the college sample in which the EPSI scores were normed. I examined changes in self-reported clinical impairment by comparing each participant's CIA score from the screening questionnaire to their end-of-treatment CIA score (Study Aim 3). I examined changes in self-reported internalizing psychopathology by comparing each participant's IDAS-II and PANAS-X scores from the psychological evaluation to their intervention-phase Week 8 scores (Study Aim 4). Qualitative results from the participant satisfaction survey are included below (Study Aim 5).

Results

No participants dropped out of the study following enrollment. Participant 1 had missing data for Week 1 of the intervention phase and Participant 3 had missing data for Week 4 of the intervention phase (both due to technological difficulties within the mobile application). Participant 2 had no missing data. None of the participants reported purging behaviors; as such, results related to EPCOT Purging are not discussed. Demographic data are presented in **Table 4**.

Primary Outcome

Visual Analysis

For all participants, visual inspection of EPCOT Total Score baseline data revealed stability (Study Aim 1; see **Figure 5**). Participant 1 demonstrated a decrease in EPCOT Total Scores from the baseline- to intervention phase and the trend stayed relatively stable during the intervention phase. Participant 2 exhibited relatively steady decreases in her EPCOT Total Scores during the baseline and intervention phases. Participant 3 demonstrated initial decreases in his EPCOT Total Scores during the baseline phase and there was considerable overlap between his baseline-and intervention-phase EPCOT Total Scores.

Tau-U Effect Sizes

Large and significant effects were observed for EPCOT Total Scores for Participant 1 (Tau-*U* = -0.82; 90% Confidence Interval [CI] = -1.00, -0.16; $p = 0.04$) and Participant 2 (Tau-*U* = -0.90; 90% CI = -1.00, -0.32; $p = 0.01$). Changes in Participant 3's EPCOT Total Scores were non-significant (Tau-*U* = -0.20; 90% CI = -0.75, 0.35; $p = 0.55$). The weighted average for EPCOT Total Score across participants was significant and indicated the presence of a large effect (Tau-*U* = -0.62; 90% CI = -0.97, -0.28; $p = 0.003$).

Secondary Outcomes

Binge Eating

Visual Analysis. For all participants, visual analysis revealed decreases in EPCOT Binge Eating across time (Study Aim 2; see **Figure 6**). Participant 1's EPCOT Binge Eating score was relatively high at the start of the intervention phase and demonstrated a noticeable decrease that stayed relatively consistently low for Weeks 4-10 of the intervention phase. Participant 2's EPCOT Binge Eating scores demonstrated a decreasing trend from the beginning of her baseline

phase to end-of-treatment. Participant 3's EPCOT Binge Eating scores were relatively low across both phases and exhibited a decrease from the baseline to intervention phase.

Tau-*U* Effect Sizes. Large and significant effects were observed for EPCOT Binge Eating scores for Participant 1 (Tau-*U* = -0.82; 90% CI = -1.00, -0.16; $p = 0.04$), Participant 2 (Tau-*U* = -0.85; 90% CI = -1.00, -0.27; $p = 0.02$), and Participant 3 (Tau-*U* = -0.99; 90% CI = -1.00, -0.43; $p = 0.003$). The weighted average for EPCOT Binge Eating across participants was significant and indicated the presence of a large effect (Tau-*U* = -0.89; 90% CI = -1.00, -0.54; $p < 0.001$).

Normed-Mean Comparisons. Comparisons to the age-matched, gender-normed means revealed that Participants 1 and 2's EPCOT Binge Eating scores decreased from >1 standard deviation above normed means during the baseline phase to >1 standard deviation below the normed mean during the intervention phase for Participant 1, and to within one standard deviation below the normed mean for Participant 2. Participant 3's scores were within one standard deviation below the normed mean in the baseline phase and >1 standard deviation below normed means in the intervention phase.

Restricting

Visual Analysis. For Participants 1, visual analysis revealed increases for EPCOT Restricting over time (Study Aim 2; see **Figure 7**). Visual inspection of the data revealed a decrease for EPCOT Restricting for Participant 2. Participant 3's EPCOT Restricting scores were consistently low across phases.

Tau-*U* Effect Sizes. The effects were non-significant for Participant 1 (Tau-*U* = 0.04; 90% CI = -0.62, 0.70; $p = 0.93$) and Participant 3 (Tau-*U* = 0.04; 90% CI = -0.50, 0.59; $p = 0.89$). The observed effect for Participant 2 was large and significant (Tau-*U* = -0.73; 90% CI = -

1.00, -0.14; $p = 0.04$). The weighted average across participants was non-significant (Tau-U = -0.22; 90% CI = -0.57, 0.13; $p = 0.30$).

Normed-Mean Comparisons. Comparisons to normed means for EPCOT Restricting revealed that the increased means for Participants 1 and 3 still fell below normed means. Participant 2's EPCOT Restricting scores were below the normed mean.

Excessive Exercise

Visual Analysis. For Participants 1 and 2, visual analysis revealed increases for EPCOT Excessive Exercise scores (Study Aim 2; see **Figure 8**). Visual inspection of the data revealed variable EPCOT Excessive Exercise scores for Participant 1. A noticeable decrease in EPCOT Excessive Exercise scores was observed for Participant 2 from Week 3 of the intervention phase that stayed near zero for the remainder of treatment. Participant 3 had variable EPCOT Excessive Exercise scores across both phases.

Tau-*U* Effect Sizes. The observed effects were non-significant for Participant 1 (Tau-U = 0.30; 90% CI = -0.36, 0.96; $p = 0.46$), Participant 2 (Tau-U = -0.40; 90% CI = -0.98, 0.18; p = 0.26), and Participant 3 (Tau-U = 0.33; 90% CI = -0.22, 0.88; $p = 0.32$). The weighted average across participants was non-significant (Tau-U = 0.07; 90% CI = -0.27, 0.42; $p = 0.73$).

Normed-Mean Comparisons. Participant 1 reported an increase in EPCOT Excessive Exercise that surpassed the age-matched, gender-normed mean; however, her intervention-phase EPCOT Excessive Exercise mean was within one standard deviation of the normed mean. Similarly, despite increased EPCOT Excessive Exercise scores across phases for Participant 3, the increased mean still fell below the normed mean (i.e., within one standard deviation). Participant 2's EPCOT Excessive Exercise means were below the normed means across phases.

Body Dissatisfaction

Visual Analysis. Visual analysis revealed relatively stable EPCOT Body Dissatisfaction scores across both phases for Participants 1 and 2 and Participant 3's scores increased from the baseline phase to the intervention phase (Study Aim 2; see **Figure 9**).

Tau-*U* Effect Sizes. The observed effect for Participant 1's EPCOT Body Dissatisfaction scores was non-significant (Tau-*U* = -0.07; 90% CI = -0.73, 0.59; *p* = 0.85). Participant 2 demonstrated a large and significant decrease in EPCOT Body Dissatisfaction scores across phases (Tau-*U* = -0.73; 90% CI = -1.00, -0.14; *p* = 0.04). Participant 3 demonstrated a small yet significant increase in EPCOT Body Dissatisfaction scores (Tau-*U* = 0.78; 90% CI = 0.23, 1.00; *p* = 0.02). This observed increase in his EPCOT Body Dissatisfaction score contrasted with his end-of-treatment participant satisfaction survey, which indicated that he "Strongly Agreed" that he improved his body image, and he wrote, "I would never have made as much progress with respect to improving my eating habits and my body image if it weren't for the [MICE] program." The weighted average effect across participants for EPCOT Body Dissatisfaction was non-significant (Tau-*U* = 0.01; 90% CI = -0.33, 0.36; *p* = 0.96).

Normed-Mean Comparisons. Participant 1's EPCOT Body Dissatisfaction means were within one standard deviation above the age-matched, gender-normed mean across phases. Participant 2's EPCOT Body Dissatisfaction scores in the baseline- and intervention phases were within one standard deviation above and below the normed mean, respectively. Although Participant 3's EPCOT Body Dissatisfaction scores increased across phases, both scores were below the normed mean.

Clinical Impairment

All participants reported improvements in psychiatric impairment secondary to their eating disorder (Study Aim 3). Participant 1's CIA score decreased from 24 to 16. Participant 2's

CIA score decreased from 18 to 9. Participant 3's CIA score decreased from 28 to 10. These changes are notable given that a score of 16 has been shown to represent an optimal cut-point for distinguishing cases of eating disorders from non-cases (Bohn et al., 2008).

Internalizing Psychopathology

Findings were mixed regarding participants' internalizing psychopathology (Study Aim 4). Participant 1's IDAS-II General Depression score increased from 19 to 29. Participant 2's IDAS-II General Depression score increased from 36 to 39. Participant 3's IDAS-II General Depression score decreased from 33 to 29. Participant 1's IDAS-II Panic score was eight at both time points. Participant 2's IDAS-II Panic score increased from 6 to 12. Participant 3's IDAS-II Panic score increased from 3 to 4. Participants 1 and 3's IDAS-II Social Anxiety scores exhibited no change (Participant 1's score = 0 at both time points; Participant 3's score = 6 at both time points). Participants 2's IDAS-II Social Anxiety score decreased from 13 to 10.

Participant 1's PANAS-X Negative Affect score increased from 9 to 15. Participant 2's PANAS-X Negative Affect score decreased from 14 to 12. Participant 3's PANAS-X Negative Affect score decreased from 23 to 14. Participant 1's PANAS-X Positive Affect score decreased from 16 to 14. Participant 2's PANAS-X Positive Affect score decreased 20 to 10. Participant 3's PANAS-X Positive Affect score decreased from 13 to 8.

Participant Satisfaction

Participants rated their satisfaction with the treatment and mobile application highly (Study Aim 5). They endorsed "Agree" or "Strongly Agree" for questions regarding their satisfaction with the treatment modules, coaching, and time commitment. Two of the three participants rated their satisfaction with the mobile application's appearance and user-friendliness highly ("Agree" or "Strongly Agree") and Participant 2 rated her satisfaction with

the mobile application slightly lower (her answers to the questions about the mobile application ranged from “Somewhat Disagree” to “Somewhat Agree”). All participants reported that they noticed improvements in their eating behaviors and body image as a result of their participation and that they felt proud of their progress (answers ranged from “Somewhat Agree” to “Strongly Agree”).

For changes they would make to the program, participants’ feedback mostly centered on the mobile application’s user interface. All participants reported the treatment modules and coaching as the best parts of their experiences. Participant 2 noted that the modules that encouraged her to take more notes (e.g., behavior-chain analysis) were most effective for her. Participant 3 described the information presented in the modules and weekly coaching sessions as “paced very well.” He stated, “Each week there was a new topic for analysis and this weekly pattern helped me tackle one problem at a time, instead of diving into the deep end directly.” All participants reported being satisfied with the length of the program.

Discussion

The present study administered a mobile CBT-gsh to college students with eating disorders using a multiple-baseline design. Visual analysis revealed decreases in EPCOT Total Scores for each participant, and the weighted average effect across participants was large and significant (primary outcome; Study Aim 1). The results for the primary outcome were encouraging and support future work to test CBT-gsh in a larger cohort of students. These results are also consistent with the work of Fitzsimmons-Craft and colleagues (2019), who also found that CBT-gsh delivered via mobile-phone application resulted in reductions in total eating-disorder psychopathology over the course of treatment.

Visual analysis revealed decreases in EPCOT Binge Eating for all participants and the Tau-*U* effect sizes were large and significant (secondary outcome; Study Aim 2). Taken together, the decreases in EPCOT Binge Eating scores across time and across participants suggest that the CBT-gsh reduced binge eating, which is consistent with the CBT-gsh target of implementing a regular eating pattern. The decreases in EPCOT Binge Eating scores are also consistent with other CBT-gsh studies that demonstrated decreases in binge eating (Carrard et al., 2011; Jacobi et al., 2012; Ljotsson et al., 2007; Ruwaard et al., 2013).

Visual analysis revealed increases in EPCOT Restricting and Excessive Exercise scores for several participants (secondary outcomes; Study Aim 2). Despite these relative increases, the effects were non-significant and the increased EPCOT Restricting and Excessive Exercise scores still fell below normed means and/or were within one standard deviation of the normed means. In other words, although several EPCOT Restricting and Excessive Exercise scores increased visually across time, these increases did not represent pathological increases when compared to age-matched, gender-normed means. Furthermore, these increases were not found to be statistically significant, indicating that CBT-gsh had no effect on restricting or excessive exercise.

Findings regarding body dissatisfaction were mixed (secondary outcome; Study Aim 2). Two participants' EPCOT Body Dissatisfaction scores were higher than normed means during the baseline phase; one of these participant's scores remained high during the intervention phase (and the Tau-*U* effect size was non-significant) and the other participant's scores decreased significantly during the intervention phase, as exhibited by a large and significant Tau-*U* effect size. Although the third participant's EPCOT Body Dissatisfaction score increased significantly across phases (as exhibited by a small yet significant Tau-*U* effect size), both scores were below

the normed mean and he self-reported experiencing improvements in his body image as a result of participation in the current study. In sum, body dissatisfaction was common in this sample and findings regarding body dissatisfaction were inconclusive. Previous research has shown that body dissatisfaction is slower to change in persons with bulimia nervosa (Chapa et al., 2020), which may partially explain the relative lack of improvement seen for body dissatisfaction in the current sample, particularly given the short duration of treatment.

All participants demonstrated improvements in self-reported clinical impairment secondary to their eating disorder (secondary outcome; Study Aim 3). All participants' scores at end-of-treatment were equal to or less than the recommended cut score that is used to distinguish eating-disorder cases from non-cases (CIA = 16; Bohn et al., 2008). This is a notable finding, given that it shows that participants in the treatment experienced fewer problems due to their eating disorder in the areas of academic and work life, family and friendship, and concentration.

Findings for internalizing psychopathology were mixed (secondary outcome; Study Aim 4). Several participants exhibited increases in self-reported internalizing symptoms across time. Reasons for such increases are unclear. One hypothesis is that stressors unrelated to the current study (e.g., stress related to the pandemic, stress related to semester finals) increased participants' internalizing psychopathology. Alternatively, the treatment material covered near Week 8 features food exposures, which may have been distressing to some participants and/or exacerbated body dissatisfaction. Thus, the timing of when we evaluated internalizing psychopathology may not be optimal given the work going on in the program at that time. Future research could re-design the assessment schedule so that additional time points are added to evaluate changes in internalizing psychopathology.

The majority of the qualitative information from the participant satisfaction survey was positive (secondary outcome; Study Aim 5). Participants rated their satisfaction with the treatment modules, coaching, and time commitment highly. Participants also indicated that they agreed that they made meaningful changes to their eating habits and body image as a result of their participation. This qualitative feedback, coupled with the results discussed above, give preliminary evidence that that this mobile CBT-gsh is an effective treatment for reducing binge eating and overall eating-disorder psychopathology, as well as reducing eating-disorder-related clinical impairment.

The current study was consistent with and extended previous research into the efficacy of mobile-assisted CBT-gsh for eating disorders. For example, the current study used a mobile CBT-gsh that closely resembled traditional CBT-gsh for eating disorders in many ways, with some exceptions. The overall duration of the treatment in the Fitzsimmons-Craft and colleagues (2019) study spanned eight months. In contrast, the current study's active treatment duration spanned 10 weeks (plus an additional three-to-five weeks for the baseline phase), which is consistent with new recommendations for treatment length of in-person CBT for eating disorders (Rose & Waller, 2017; Waller et al., 2018) and consistent with the timeframe of the Hildebrandt and colleagues (2017 and 2020) studies that each spanned 12 weeks. The current study extended previous mobile CBT-gsh research by delivering coaching sessions via telehealth, which eliminated the need for students to be on campus for in-person visits. In contrast, coaching sessions were delivered in-person for the Hildebrandt and colleagues (2017 and 2020) studies and via a text-messaging feature within the mobile application in the Fitzsimmons-Craft and colleagues (2019) study. Of course, the students in the Fitzsimmons-Craft and colleagues (2019) study benefited by being able to text message their coach remotely, but the study design did not

grant them “face time” with their coach (i.e., the opportunity to see and speak with their coach in-person or virtually). Text messaging may be preferable for some students but the effectiveness of text messaging in a therapeutic context—especially CBT-gsh for eating disorders—has yet to be evaluated. In the current study, qualitative participant feedback indicated that coaching sessions were one of the most beneficial aspects of the treatment, so it is likely that text messaging may be less useful or contribute to greater drop out than face-to-face or telehealth coaching. Finally, the current study extended the work done by Hildebrandt and colleagues (2017 and 2020) by eliminating the need for additional treatment materials (e.g., a hard-copy of the CBT-gsh material). The fully mobile nature of the current study and the Fitzsimmons-Craft and colleagues (2019) study eliminated additional materials students needed to be able to engage with treatment. Given the ubiquity of technology, mobile CBT-gsh has the potential to reach a greater number of college students with eating disorders because mobile CBT-gsh requires only a smartphone for implementation, rather than additional materials and resources.

Mobile CBT-gsh represents a cost-effective alternative to face-to-face eating-disorder treatment modalities, which are often expensive and difficult to access (Striegel-Moore et al., 2000). Indeed, the incremental worth of mobile-assisted interventions may lie in their cost-effectiveness. The cost-effectiveness of an intervention is related to direct costs (e.g., cost of treatment materials) and indirect costs (e.g., cost of transportation to the healthcare facility) associated with the intervention. The mobile CBT-gsh used in the current study has the potential to reduce direct- and indirect costs associated with eating-disorder treatment in college students, thus increasing cost-effectiveness, by eliminating additional resources needed to engage with treatment (beyond a smartphone) and allowing participants to complete the coaching sessions remotely. Taken together, the findings of the current study and the intervention’s potential to be

highly cost-effective suggest that future testing of the intervention in a larger sample is warranted.

The current study allowed for testing of the mobile CBT-gsh for eating disorders in a smaller sample of college students prior to larger-scale implementation. As previously discussed, the current study represented a *Phase IIa* study within the ORBIT model (Czajkowski et al., 2015) because it evaluated mobile CBT-gsh in a small sample of students with eating disorders. The goal of *Phase IIa* studies is to produce clinically significant effects in a small sample, which was accomplished for several of the outcomes in the present study. *Phase IIa* studies directly inform whether it is appropriate to move on to testing the intervention in larger samples. Indeed, the multiple-baseline design of the current study helped me to understand the individual needs of participants, which in turn will allow for the intervention to be tailored to fit the hypothesized needs of a larger population. In other words, I was able to administer the intervention using a study design that highlighted individual differences and outcomes in a small sample, which helped me better understand, prior to implementing the intervention in a larger college sample, "what works" and "what works *for whom*."

Of note, all participants in the current study reported their sexual orientation as bisexual or gay (i.e., sexual minority). The sexual-minority status of participants in the current study was consistent with research demonstrating that: 1) sexual-minority college students were more likely than heterosexual college students to seek mental-health treatment (Eisenberg et al., 2011a); and 2) sexual-minority college students reported higher rates of eating-disorder symptoms than their heterosexual peers (Diemer et al., 2015; Hazzard et al., 2020; Simone et al., 2020). Eating-disorder-related academic impairment was higher for sexual-minority students compared to their heterosexual peers as well, highlighting the differential impact eating disorders may have on

sexual-minority students (Simone et al., 2020). Although the present study's mobile application and treatment materials were not developed or tailored specifically for people in sexual minorities, future research to develop tailored interventions for sub-populations of college students may improve outcomes.

Specifically, a minority-stress framework may be useful in understanding the differential impact of eating disorders on sexual-minority students. Within a minority-stress framework, persons of marginalized identities are posited to be exposed to excess stress as a result of discrimination and victimization in relation to their social position, which is often in the minority (hence "minority stress"). For persons who identify as a member of a sexual-minority, it is theorized that excess stress related to their often-stigmatized sexual orientation may put them at increased risk for adverse outcomes (i.e., "sexual-minority stress;" Meyer, 2003). Myriad negative mental-health outcomes have been documented and theorized to develop as a result of sexual-minority stress, including mood- and anxiety disorders (Bränström, 2017; Cochran et al., 2003; Meyer, 2003). Given the high rates of comorbidity among mood-, anxiety-, and eating disorders, it is reasonable to expect that for some, sexual-minority stress could result in the development of eating disorders.

Limitations

The current study findings must be considered in light of several limitations. First, there were a few minor technological difficulties that resulted in missing one week's EPCOT scores for two participants because the surveys were not administered to the participants. However, the technical difficulties did not prevent participants from accessing the treatment modules. The technological difficulties happened during weeks within the participants' overall study enrollment that were not primary time points (i.e., they happened during weeks when the surveys

administered were not to be used to calculate phase means or for decisions to move participants between phases). Although the missing data may have affected the visual analyses, it likely did not impact the majority of the study findings. Moreover, missing data was extremely minimal given the daily nature of some logs and the fact that the majority of time points were completed.

Another set of limitations that must be considered are possible threats to internal validity. History events are often unpredictable and difficult to control and a non-concurrent multiple-baseline design can be susceptible to these types of effects. I attempted to address this threat to internal validity by staggering treatment administration over time across participants, and by extending the baseline phase for participants when necessary. Staggering treatment administration and extending the baseline phase helped ensure that maturation or history effects could be ruled out as an explanation of any treatment effects because treatment effects were observed at different points in time across participants. That said, the global pandemic was an unprecedented event and it is unclear how much the findings were impacted by the pandemic.

Statistical regression (i.e., participant improvement on any dependent variable that could be expected if the intervention had not been implemented) was another possible threat to internal validity. It is reasonable to expect improvement on part of the participant merely by enrolling in treatment (i.e., before the intervention has been implemented; DeYoung & Bottera, 2018) and this may have partially explained the low variability in Participant 3's EPCOT Total Composite scores. Participant 3's baseline measurements decreased steadily during Weeks 1-3 of the baseline phase; the measurements then stabilized from Week 3 to Week 5 and remained consistently low for the remainder of his study enrollment (his scores were low across scales, thus limiting the amount of improvement possible to observe during the intervention phase).

Although it is possible that statistical regression affected Participant 3's scores during the baseline phase, it does not explain the treatment effects observed across time and participants.

Strengths

In addition to the limitations, several strengths ought to be noted for the current study. First, there were diverse identities represented in the sample that are not typically well-represented in eating-disorder treatment research. Second, the participants gave positive feedback regarding their study participation and the treatment in general, suggesting that the current study's treatment and design were acceptable. Indeed, the treatment effects observed, coupled with the positive feedback expressed, suggest that the participants in the current study noticed a benefit from their participation. Finally, the intervention tested in the current study filled a major service gap by providing eating-disorder treatment resources at a university that previously did not provide access to affordable CBT-gsh for eating disorders.

Conclusion and Future Directions

The current study was novel as it was one of the first to examine mobile CBT-gsh for the treatment of eating disorders in college students. The current study represented one of few applications of a multiple-baseline design to examine treatment effects in the field of eating disorders. Although results of the current study are promising, future research is needed to extend the current study findings. For example, although the current study sample met recommended standards for multiple-baseline designs, which call for a minimum of three participants (Kratochwill et al., 2013), interpretations of treatment effects are stronger if they can be replicated across participants multiple times (Kratochwill et al., 2010). Subsequent successful replications are needed to increase external validity. Specifically, if an effect can be demonstrated in two or more participants—as it was for certain outcomes in the present study—

the effect can be said to be generalized across participants and external validity is strengthened with each subsequent replication. Strengthened external validity, in turn, will bolster the argument that the observed effects are related to the intervention, which would indicate readiness to move to *Phase IIb* or *Phase III* of the ORBIT model to conduct additional pilot tests and tests of treatment efficacy (Czajkowski et al., 2015). In addition to subsequent successful replications of treatment effects, additional supplementations to visual analysis may be worth exploring. The current study followed recommended guidelines for visual- and statistical analysis of the data and future research may wish to implement additional cutting-edge analytic techniques. Two such techniques are simulation modeling analysis, which uses bootstrapping techniques to estimate statistical significance for treatment effects, and the conservative dual-criterion method, which may increase inter-rater reliability of visual analytic results (Swoboda et al., 2010; Wolfe et al., 2018). Use of additional statistical techniques to analyze the observed treatment effects and strengthening external validity would increase the study findings' innovation and impact.

Additional future directions include implementation of personalized, in-the-moment feedback based on real-time assessment of user behavior. For example, the mobile CBT-gsh could be augmented to provide real-time suggestions for cognitive- and behavioral skills to use based on ecological momentary assessment of users' mood and eating-disorder symptoms. For example, the mobile application could survey users at random intervals and provide feedback based on the information the user provides. Alternatively, the mobile application could be modified to include a feature to provide tailored recommendations for skills to use based on the user's current mood and eating symptoms (e.g., "I'm having an urge to binge") and based on skills that have been helpful in the user's past. Such a feature would provide therapeutic support between the weekly coaching sessions, when users are practicing the skills they have been

learning in therapy. The personalized, automated feedback provided between coaching sessions could be especially useful for college students who may need ancillary support to manage eating-disorder treatment in addition to their academic course load.

The current study tested a novel version of CBT-gsh delivered via mobile phone application and individual brief coaching sessions in a sample of college students. Although several modules included information relevant to college students (e.g., time management, social pressures, and fear of missing out or "FOMO"), feedback from the health coaches indicated that additional modules focused on interpersonal- and academic challenges unique to college students would be helpful. The health coaches also indicated that splitting the behavior-chain-analysis material from the "mindtraps" material would be beneficial, as each topic represents a sizeable amount of information to cover in one coaching session. Modifying the existing treatment package according to the feedback provided by the health coaches would allow the treatment to be even more tailored to college students' needs and ensure that participants and coaches had adequate time to cover important information presented in each module. Lastly, once the intervention has been tailored, tested, and advanced along the ORBIT model framework, reasonable future studies include testing the intervention on more college campuses and evaluating the best approach to stepped-care and triage for eating disorders in college students.

In summary, the current study showed promise for mobile CBT-gsh in the treatment of eating disorders in college students. Participants demonstrated improvements in self-reported overall eating-disorder psychopathology and binge eating, specifically. The findings for other eating-disorder symptoms were mixed, although the participants uniformly provided positive feedback on the end-of-treatment survey, indicating that they noticed improvements in their eating- and body-image problems as a result of their participation. The current findings add to

the small but encouraging body of literature investigating mobile interventions for college students with eating disorders. Indeed, mobile CBT-gsh may represent an innovative tool that could be used on a larger scale to reach college students with eating disorders who are under-served.

References

Academy for Eating Disorders. (2016). *Medical Care Standards Guide: Critical Points for Early Recognition and Medical Risk Management in the Care of Individuals with Eating Disorders*. https://www.aedweb.org/resources/online-library/publications/medical-care-standards

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Arcelus, J., Mitchell, A. J., Wales, J., & Nielsen, S. (2011). Mortality rates in patients with anorexia nervosa and other eating disorders: A meta-analysis of 36 studies. *Archives of General Psychiatry, 68*, 724-731.

Ashwood, J. S., Stein, B. D., Briscombe, B., Sontag-Padilla, L. M., Woodbridge, M. W., May, E., & Burnam, M. (2015). Payoffs for California college students and taxpayers from investing in student mental health. Santa Monica, CA: Rand Corporation.

Bailer, U., de Zwaan, M., Leisch, F., Strnad, A., Lennkh-Wolfsberg, C., El-Giamal, N., . . . Kasper, S. (2004). Guided self-help versus cognitive-behavioral group therapy in the treatment of bulimia nervosa. *International Journal of Eating Disorders, 35*, 522-537.

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. doi:http://dx.doi.org.www2.lib.ku.edu/10.1016/S0005-7894(03)80001-1

Berg, K. C., Frazier, P., & Sherr, L. (2009). Change in eating disorder attitudes and behavior in college women: Prevalence and predictors. *Eating Behaviors, 10*, 137-142. doi: https://doi.org/10.1016/j.eatbeh.2009.03.003

Bohn, K., Doll, H. A., Cooper, Z., O'Connor, M., Palmer, R. L., & Fairburn, C. G. (2008). The measurement of impairment due to eating disorder psychopathology. *Behaviour Research and Therapy, 46*, 1105-1110.

Bränström, R. (2017). Minority stress factors as mediators of sexual orientation disparities in mental health treatment: A longitudinal population-based study. *Journal of Epidemiology and Community Health*, 71, 446-452. doi:http://dx.doi.org/10.1136/jech-2016-207943

Carrard, I., Fernandez-Aranda, F., Lam, T., Nevonen, L., Liwowsky, I., Volkart, A., . . . Norring, C. (2011). Evaluation of a guided internet self-treatment programme for bulimia nervosa in several European countries. *European Eating Disorders Review, 19*, 138-149.

Center for Collegiate Mental Health. (2016). Center for collegiate mental health: 2016 annual report. Retrieved from https://sites.psu.edu/ccmh/ files/2017/01/2016-Annual-Report-FINAL_2016_01_09-1gc2hj6.pdf

Centre for Research on Eating Disorders at Oxford (CREDO). (2020, June 8). *Web-centred training in CBT-E or behavioral activation*. https://www.credo-oxford.com/9.2.html

Chapa, D. A., Hagan, K. E., Forbush, K. T., Clark, K. E., Tregarten, J. P., & Argue, S. (2020). Longitudinal trajectories of behavior change in a national sample of patients seeking eating-disorder treatment. *International Journal of Eating Disorders*. doi: 10.1002/eat.23272

Cochran, S. D., Sullivan, J. G., & Mays, V. M. (2003). Prevalence of mental disorders, psychological distress, and mental health services use among lesbian, gay, and bisexual adults in the United States. *Journal of Consulting and Clinical Psychology, 71*, 53–61. doi:https://doi.org/10.1037/0022-006X.71.1.53

Coniglio, K. A., Becker, K. R., Tabri, N., Keshishian, A. C., Miller, J. D., Eddy, K. T., & Thomas, J. J. (2018). Factorial integrity and validation of the Eating Pathology Symptoms Inventory (EPSI). *Eating Behaviors, 31*, 1-7. doi: https://doi.org/10.1016/j.eatbeh.2018.07.004

Crow, S. J., Agras, W. S., Halmi, K. A., Fairburn, C. G., Mitchell, J. E., & Nyman, J. A. (2013). A cost effectiveness analysis of stepped care treatment for bulimia nervosa. *International Journal of Eating Disorders, 46*, 302-307.

Czajkowski, S. M., Powell, L. H., Adler, N., Naar-King, S., Reynolds, K. D., Hunter, C. M., ... & Epel, E. (2015). From ideas to efficacy: The ORBIT model for developing behavioral treatments for chronic diseases. *Health Psychology*, 34, 971.

De Young, K. P., & Bottera, A. R. (2018). A summary of reporting guidelines and evaluation domains for using single-case experimental designs and recommendations for the study of eating disorders. *International Journal of Eating Disorders*. doi: https://doi.org/10.1002/eat.22887

Diemer, E. W., Grant, J. D., Munn-Chernoff, M. A., Patterson, D. A., & Duncan, A. E. (2015). Gender identity, sexual orientation, and eating-related pathology in a national sample of college students. *Journal of Adolescent Health, 57*, 144-149. doi:https://doi.org/10.1016/j.jadohealth.2015.03.003

Eisenberg, D., Golberstein, E., & Hunt, J. B. (2009). Mental health and academic success in college. *The BE Journal of Economic Analysis & Policy, 9*.

Eisenberg, D., Nicklett, E. J., Roeder, K., & Kirz, N. E. (2011). Eating disorder symptoms among college students: Prevalence, persistence, correlates, and treatment-seeking.

Journal of American College Health, 59, 700-707. doi:https://doi.org/10.1080/07448481.2010.546461

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

Fairburn, C., Cooper, Z., Doll, H., O'Connor, M., Bohn, K., Hawker, D., . . . Palmer, R. (2009). Transdiagnostic cognitive-behavioral therapy for patients with eating disorders: A two-site trial with 60-week follow-up. *American Journal of Psychiatry, 166*, 311-319.

Fernández-Aranda, F., Núñez, A., Martínez, C., Krug, I., Cappozzo, M., Carrard, I., . . . Penelo, E. (2009). Internet-based cognitive-behavioral therapy for bulimia nervosa: A controlled study. *Cyberpsychology & Behavior, 12*, 37-41.

Fitzsimmons-Craft, E. E., Firebaugh, M. L., Graham, A. K., Eichen, D. M., Monterubio, G. E., Balantekin, K. N., ... & Wilfley, D. E. (2019). State-wide university implementation of an online platform for eating disorders screening and intervention. *Psychological Services, 16*, 239. doi: https://doi.org/10.1037/ser0000264

Forbush, K. T., Wildes, J. E., & Hunt, T. K. (2014). Gender norms, psychometric properties, and validity for the Eating Pathology Symptoms Inventory. *International Journal of Eating Disorders, 47,* 85-91. doi:https://doi.org/10.1002/eat.22180

Forbush, K. T., Wildes, J. E., Pollack, L. O., Dunbar, D., Luo, J., Patterson, K., Petruzzi, L., Pollpeter, M., Miller, H., Stone, A., Bright, A., & Watson, D. (2013). Development and validation of the Eating Pathology Symptoms Inventory (EPSI). *Psychological Assessment, 25*, 859–878. doi:https://doi.org/10.1037/a0032639

Galmiche, M., Déchelotte, P., Lambert, G., & Tavolacci, M. P. (2019). Prevalence of eating disorders over the 2000–2018 period: A systematic literature review. *The American Journal of Clinical Nutrition, 109*, 1402-1413. doi: https://doi.org/10.1093/ajcn/nqy342

Grenon, R., Schwartze, D., Hammond, N., Ivanova, I., McQuaid, N., Proulx, G., & Tasca, G. A. (2017). Group psychotherapy for eating disorders: A meta-analysis. *International Journal of Eating Disorders, 50*, 997-1013. doi:10.1002/eat.22744

Grilo, C. M. (2020). Smartphone-assisted delivery of cognitive-behavioral guided self-help for binge eating: Cautionary musings of implications given the importance of comparison groups [Editorial]. *The American Journal of Psychiatry, 177*, 110-112. doi: https://doi.org/10.1176/appi.ajp.2019.19121256

Harris, E. C., & Barraclough, B. (1998). Excess mortality of mental disorder. *The British Journal of Psychiatry, 173*, 11-53.

Hay, P. (2013). A systematic review of evidence for psychological treatments in eating disorders: 2005–2012. *International Journal of Eating Disorders, 46*, 462-469.

Hay, P., Girosi, F., & Mond, J. (2015). Prevalence and sociodemographic correlates of DSM-5 eating disorders in the Australian population. *Journal of Eating Disorders, 3*, 19.

Hazzard, V. M., Simone, M., Borg, S. L., Borton, K. A., Sonneville, K. R., Calzo, J. P., & Lipson, S. K. (2020). Disparities in eating disorder risk and diagnosis among sexual minority college students: Findings from the national Healthy Minds Study. *International Journal of Eating Disorders*. doi:10.1002/eat.23304

Hildebrandt, T., Michaelides, A., Mackinnon, D., Greif, R., DeBar, L., & Sysko, R. (2017). Randomized controlled trial comparing smartphone assisted versus traditional guided

self-help for adults with binge eating. *International Journal of Eating Disorders, 50*, 1313-1322. doi:10.1002/eat.22781

Hildebrandt., T., Michaeledes, A., Mayhew, M., Greif, R., Sysko, R., Toro-Ramos, T., & DeBar, L. (2020). Randomized controlled trial comparing health coach-delivered smartphone-guided self-help with standard care for adults with binge eating. *American Journal of Psychiatry, 177*, 134-142. doi:10.1176/appi.ajp.2019.19020184

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, 71, 165-179.

Howard, L. M., Romano, K. A., & Heron, K. E. (2020). Prospective changes in disordered eating and body dissatisfaction across women's first year of college: The relative contributions of sociocultural and college adjustment risk factors. *Eating Behaviors, 36*, 5. doi:http://dx.doi.org.www2.lib.ku.edu/10.1016/j.eatbeh.2019.101357

Hudson, J. I., Hiripi, E., Pope, H. G., & Kessler, R. C. (2007). The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. *Biological Psychiatry, 61*, 348-358.

Jacobi, C., Völker, U., Trockel, M. T., & Taylor, C. B. (2012). Effects of an Internet-based intervention for subthreshold eating disorders: A randomized controlled trial. *Behaviour Research and Therapy, 50*, 93-99.

Kass, A. E., Balantekin, K. N., Fitzsimmons-Craft, E. E., Jacobi, C., Wilfley, D. E., & Taylor, C. B. (2017). The economic case for digital interventions for eating disorders among United States college students. *International Journal of Eating Disorders, 50*, 250-258. doi: 10.1002/eat.22680

Kazdin, A. E., Fitzsimmons-Craft, E. E., & Wilfley, D. E. (2017). Addressing critical gaps in the treatment of eating disorders. *International Journal of Eating Disorders, 50*, 170-189. doi: 10.1002/eat.22670

Klump, K. L., Bulik, C. M., Kaye, W. H., Treasure, J., & Tyson, E. (2009). Academy for eating disorders position paper: Eating disorders are serious mental illnesses. *International Journal of Eating Disorders, 42*, 97-103.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M & Shadish, W. R. (2010). *Single-case designs technical documentation*. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education, 34*, 26-38.

Linardon, J., Shatte, A., Tepper, H., & Fuller-Tyszkiewicz, M. (2020). A survey study of attitudes toward, and preferences for, e-therapy interventions for eating disorder psychopathology. *International Journal of Eating Disorders*. doi: 10.1002/eat.23268

Lipson, S. K., Jones, J. M., Taylor, C. B., Wilfley, D. E., Eichen, D. M., Fitzsimmons-Craft, E. E., & Eisenberg, D. (2017). Understanding and promoting treatment-seeking for eating disorders and body image concerns on college campuses through online screening, prevention and intervention. *Eating Behaviors, 25,* 68-73. doi:10.1016/j.eatbeh.2016.03.020

Ljotsson, B., Lundin, C., Mitsell, K., Carlbring, P., Ramklint, M., & Ghaderi, A. (2007). Remote treatment of bulimia nervosa and binge eating disorder: A randomized trial of Internet-assisted cognitive behavioural therapy. *Behaviour Research and Therapy, 45*, 649-661.

Lynch, F. L., Striegel-Moore, R. H., Dickerson, J. F., Perrin, N., DeBar, L., Wilson, G. T., & Kraemer, H. C. (2010). Cost-effectiveness of guided self-help treatment for recurrent binge eating. *Journal of Consulting and Clinical Psychology, 78*, 322-333. doi:10.1037/a0018982

McClay, C. A., Waters, L., Schmidt, U., & Williams, C. (2016). A survey of attitudes towards computerized self-help for eating disorders within a community-based sample. *Behavioural and Cognitive Psychotherapy, 44,* 65. doi:10.1017/S1352465814000484

Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: Conceptual issues and research evidence. *Psychological Bulletin, 129*, 674. doi:10.1037/0033-2909.129.5.674

Mitchell, J. E., & Crow, S. (2006). Medical complications of anorexia nervosa and bulimia nervosa. *Current Opinion in Psychiatry, 19*, 438-443.

Mitchell, J. E., Myers, T., Crosby, R., O'Neill, G., Carlisle, J., & Gerlach, S., (2009). Healthcare utilization in patients with eating disorders. *International Journal of Eating Disorders, 42*, 571-574. doi:10.1002/eat.20651

Mowbray, C. T., Megivern, D., Mandiberg, J. M., Strauss, S., Stein, C. H., Collins, K., . . . Lett, R. (2006). *Campus mental health services: Recommendations for change*. American Journal of Orthopsychiatry, 76, 226-237. doi:http://dx.doi.org.www2.lib.ku.edu/10.1037/0002-9432.76.2.226

National Collaborating Center for Mental Health. (2004). *Eating disorders: Core interventions in the treatment and management of anorexia nervosa, bulimia nervosa and related eating disorders*: British Psychological Society (UK).

Olive, M. L., & Smith, B. W. (2005). Effect size calculations and single subject designs. *Educational Psychology, 25*, 313-324. doi: https://doi.org/10.1080/0144341042000301238

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy, 42*, 284-299. doi: https://doi.org/10.1016/j.beth.2010.08.006

Petry, N. M., Martin, B., Cooney, J. L., & Kranzler, H. R. (2000). Give them prizes and they will come: Contingency management for treatment of alcohol dependence. *Journal of Consulting and Clinical Psychology, 68*, 250. doi:10.1037//C022-006X.68.2.250

Pew Research. (2019, June 12). *Mobile Fact Sheet*. https://www.pewresearch.org/internet/fact-sheet/mobile/

Preti, A., Rocchi, M., Sisti, D., Camboni, M., & Miotto, P. (2011). A comprehensive meta-analysis of the risk of suicide in eating disorders. *Acta Psychiatrica Scandinavica, 124*, 6-17.

Robinson, P., & Serfaty, M. (2008). Getting better byte by byte: A pilot randomised controlled trial of email therapy for bulimia nervosa and binge eating disorder. *European Eating Disorders Review, 16*, 84-93.

Roll, J. M., & Shoptaw, S. (2006). Contingency management: Schedule effects. *Psychiatry Research, 144*, 91-93. doi:https://doi.org/10.1016/j.psychres.2005.12.003

Rose, C., & Waller, G. (2017). Cognitive–behavioral therapy for eating disorders in primary care settings: Does it work, and does a greater dose make it more effective? *International Journal of Eating Disorders, 50*, 1350-1355.

Ruwaard, J., Lange, A., Broeksteeg, J., Renteria-Agirre, A., Schrieken, B., Dolan, C. V., & Emmelkamp, P. (2013). Online cognitive–behavioural treatment of bulimic symptoms: A randomized controlled trial. *Clinical Psychology & Psychotherapy, 20*, 308-318.

Sánchez-Ortiz, V., Munro, C., Stahl, D., House, J., Startup, H., Treasure, J., . . . Schmidt, U. (2011). A randomized controlled trial of internet-based cognitive-behavioural therapy for bulimia nervosa or related disorders in a student population. *Psychological Medicine, 41*, 407-417.

Simone, M., Askew, A., Lust, K., Eisenberg, M. E., & Pisetsky, E. M. (2020). Disparities in self-reported eating disorders and academic impairment in sexual and gender minority college students relative to their heterosexual and cisgender peers. *International Journal of Eating Disorders, 53*, 513-524. doi:10.1002/eat.23226

Smith, J. D. (2012). Single-case experimental designs: A systematic review of published research and current standards. *Psychological Methods, 17*, 510-550. doi:http://dx.doi.org.www2.lib.ku.edu/10.1037/a0029312

Smith-Jackson, T., & Reel, J. J. (2012). Freshmen women and the "freshman 15": Perspectives on prevalence and causes of college weight gain. *Journal of American College Health, 60*, 14-20. doi:http://dx.doi.org.www2.lib.ku.edu/10.1080/07448481.2011.555931

Sonneville, K. R., & Lipson, S. K. (2018). Disparities in eating disorder diagnosis and treatment according to weight status, race/ethnicity, socioeconomic background, and sex among college students. *International Journal of Eating Disorders, 51*, 518-526. doi:10.1002/eat.22846

Stanger, C., Lansing, A. H., & Budney, A. J. (2016). Contingency management approaches for adolescent substance use disorders. *Child and Adolescent Psychiatric Clinics of North America, 25*, 645. doi:10.1016/j.chc.2016.05.002

Stice, E., Fisher, M., & Martinez, E. (2004). Eating Disorder Diagnostic Scale: Additional evidence of reliability and validity. *Psychological Assessment, 16*, 60.

Stice, E., Marti, C. N., & Rohde, P. (2013). Prevalence, incidence, impairment, and course of the proposed DSM-5 eating disorder diagnoses in an 8-year prospective community study of young women. *Journal of Abnormal Psychology, 122*, 445.

Stice, E., Telch, C. F., & Rizvi, S. L. (2000). Development and validation of the Eating Disorder Diagnostic Scale: A brief self-report measure of anorexia, bulimia, and binge-eating disorder. *Psychological Assessment, 12*, 123.

Striegel-Moore, R. H., Leslie, D., Petrill, S. A., Garvin, V., & Rosenheck, R. A. (2000). One-year use and cost of inpatient and outpatient services among female and male patients with an eating disorder: Evidence from a national database of health insurance claims. *International Journal of Eating Disorders, 27*, 381-389. doi: https://doi.org/10.1002/(SICI)1098-108X(200005)27:4<381::AID-EAT2>3.0.CO;2-U

Striegel-Moore, R. H., Wilson, G. T., DeBar, L., Perrin, N., Lynch, F., Rosselli, F., & Kraemer, H. C. (2010). Cognitive behavioral guided self-help for the treatment of recurrent binge eating. *Journal of Consulting and Clinical Psychology, 78*, 312-321. doi:10.1037/a0018915

Swoboda, C. M., Kratochwill, T. R., & Levin, J. R. (2010). *Conservative dual-criterion method for single-case research: A guide for visual analysis of AB, ABAB, and multiple- baseline designs (WCER Working Paper No. 2010-13)*. Retrieved from University of Wisconsin–Madison, Wisconsin Center for Education Research website: http://www.wcer.wisc.edu/publications/workingPapers/papers.php

Tarlow, K. R. (2017). An improved rank correlation effect size statistic for single-case designs: Baseline corrected tau. *Behavior Modification, 41*, 427-467. doi:http://dx.doi.org.www2.lib.ku.edu/10.1177/0145445516676750

Taylor, C. B., Bryson, S., Luce, K. H., Cunning, D., Doyle, A. C., Abascal, L. B., ... & Wilfley, D. E. (2006). Prevention of eating disorders in at-risk college-age women. *Archives of General Psychiatry, 63*, 881-888. doi: 10.1001/archpsyc.63.8.881

Taylor, C. B., Kass, A. E., Trockel, M., Cunning, D., Weisman, H., Bailey, J., ... & Wilfley, D. E. (2016). Reducing eating disorder onset in a very high risk sample with significant comorbid depression: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 84*, 402. doi:http://dx.doi.org/10.1037/ccp0000077

Traviss, G. D., Heywood-Everett, S., & Hill, A. J. (2011). Guided self-help for disordered eating: A randomised control trial. *Behaviour Research and Therapy, 49*, 25-31.

University of Kansas – International Support Services. (2020). *Total Enrollment: Spring 2020 (Lawrence and Edwards Campuses).* https://iss.ku.edu/sites/iss.ku.edu/files/docs/Statistics/2020/s20_total_pie.pdf

Vannest, K. J., & Ninci, J. (2015). Evaluating intervention effects in single-case research designs. *Journal of Counseling & Development, 93*, 403-411. doi: 10.1002/jcad.12038

Vannest, K.J., Parker, R.I., Gonen, O., & Adiguzel, T. (2016). *Single Case Research: Web based calculators for SCR analysis. (Version 2.0)*. College Station, TX: Texas A&M University. Retrieved Tuesday 7th July 2020. Available from singlecaseresearch.org

Wagner, G., Penelo, E., Nobis, G., Mayrhofer, A., Wanner, C., Schau, J., Spitzer, M., Gwinner, P., Trofaier, M. L., Imgart, H., Fernandez-Aranda, F., &Karwautz, A. (2015). Predictors of good therapeutic outcome and drop-out in technology assisted guided self-help in the treatment of bulimia nervosa and bulimia like phenotype. *European Eating Disorders Review, 23*, 163-169. doi: 10.1002/erv.2336

Waller, G., Tatham, M., Turner, H., Mountford, V. A., Bennetts, A., Bramwell, K., ... & Ingram, L. (2018). A 10-session cognitive-behavioral therapy (CBT-T) for eating disorders: Outcomes from a case series of nonunderweight adult patients. *International Journal of Eating Disorders, 51*, 262-269.

Watkins, D. C., Hunt, J. B., & Eisenberg, D. (2012). Increased demand for mental health services on college campuses: Perspectives from administrators. *Qualitative Social Work, 11*, 319–337. https://doi.org/10.1177/1473325011401468

Watson, D., & Clark, L. A. (1999). The PANAS-X: Manual for the positive and negative affect schedule-expanded form. New York: Springer Publishing.

Watson, D., O'Hara, M. W., Naragon-Gainey, K., Koffel, E., Chmielewski, M., Kotov, R., Stasik, S. M., & Ruggero, C. J. (2011). Development and validation of new anxiety and bipolar symptom scales for an expanded version of the IDAS (the IDAS-II). *Assessment, 19*, 399-420. doi:10.1177/1073191112449857

Wilson, S. M., Darling, K. E., Fahrenkamp, A. J., D'Auria, A. L., & Sato, A. F. (2015). Predictors of emotional eating during adolescents' transition to college: does body mass index moderate the association between stress and emotional eating? *Journal of American College Health, 63*, 163-170. doi: 10.1080/07448481.2014.1003374

Wolfe, K., Seaman, M. A., Drasgow, E., & Sherlock, P. (2018). An evaluation of the agreement between the conservative dual-criterion method and expert visual analysis. *Journal of Applied Behavior Analysis, 51*, 345-351. doi:10.1002/jaba.453

Xiao, H., Carney, D. M., Youn, S. J., Janis, R. A., Castonguay, L. G., Hayes, J. A., & Locke, B. D. (2017). Are we in crisis? National mental health and treatment trends in college counseling centers. *Psychological Services, 14*, 407-415. doi:http://dx.doi.org.www2.lib.ku.edu/10.1037/ser0000130

Yager, J., Devlin, M. J., Halmi, K. A., Herzog, D. B., Mitchell, J. E., Powers, P. S., & Zerbe, K. J. (2005). Guideline watch: Practice guideline for the treatment of patients with eating disorders. *Focus, 3*, 546-551.

Zerwas, S. C., Watson, H. J., Hofmeier, S. M., Levine, M. D., Hamer, R. M., Crosby, R. D., . . . Zimmer, B. (2017). CBT4BN: A randomized controlled trial of online chat and face-to-face group therapy for bulimia nervosa. *Psychotherapy and Psychosomatics, 86*, 47-53.

Table 1

Study Aims, Hypotheses, Outcomes and Measures

Study Aim	Hypothesis	Outcome	Outcome measure
Aim 1: Test the efficacy of a novel, mobile CBT-gsh intervention for reducing overall eating-disorder psychopathology within-person over time using a multiple-baseline design.	• Hypothesis 1: Participants will demonstrate reductions in eating-disorder psychopathology, as documented by significant within-person reductions of Total Composite scores on a self-report eating-disorder measure.	• Primary	• EPCOT
Aim 2: Test the efficacy of a novel, mobile CBT-gsh intervention for reducing eating-disorder behaviors and cognitions within-person over time using a multiple-baseline design.	• Hypothesis 2: Participants will demonstrate reductions in eating-disorder behaviors, as documented by significant within-person reductions of Binge Eating, Restricting, Purging, Excessive Exercise, and Body Dissatisfaction scores on a self-report eating-disorder measure.	• Secondary	• EPCOT
Aim 3: Test the efficacy of a mobile CBT-gsh intervention for reducing psychiatric impairment.	• Hypothesis 3: Participants will have significantly lower within-person scores on a self-report measure of psychiatric impairment secondary to an eating disorder at end-of-treatment compared to pre-treatment.	• Secondary	• CIA

Note: CBT-gsh=Cognitive-Behavior Therapy guided self-help; EPCOT=Eating Pathology Symptoms Inventory; CIA=Clinical Impairment Assessment; IDAS-II=Inventory of Depression and Anxiety Symptoms; PANAS-X=Positive and Negative Affect Schedule-Expanded Form.

Table 1 Continued

Study Aim	Hypothesis	Outcome	Outcome Measure
Aim 4: Test the efficacy of a mobile CBT-gsh intervention improving overall mental health.	• Hypothesis 4: Participants will have significantly lower within-person scores on measures of comorbid psychopathology (i.e., self-reported general depression, panic, social anxiety symptoms) at end-of-treatment compared to pre-treatment.	• Secondary	• IDAS-II • PANAS-X
Aim 5: Collect qualitative data about user experiences.	• No *a priori* hypotheses were made for this study aim.	• Secondary	• Participant satisfaction survey

Note: CBT-gsh=Cognitive-Behavior Therapy guided self-help; EPCOT=Eating Pathology Symptoms Inventory; CIA=Clinical Impairment Assessment; IDAS-II=Inventory of Depression and Anxiety Symptoms; PANAS-X=Positive and Negative Affect Schedule-Expanded Form.

Table 2

Assessment Schedule

Questionnaire	Screening	Psychological Evaluation	Weekly Questions	Weeks 4, 8 Questions	Post-Study Survey
CIA	X			X	X
EDDS	X				
Demographics	X	X			
Intake Interview		X			
EPCOT		X	X		
IDAS-II		X		X	
PANAS-X		X		X	
Objective Weight and Height Measurements		X			
Medical Evaluation		X			
Post-study survey					X

Note: CIA=Clinical Impairment Assessment; EDDS=Eating Disorder Diagnostic Scale; EPCOT=Eating Pathology Clinical Outcomes Tracking; IDAS-II=Inventory of Depression and Anxiety Symptoms; PANAS-X=Positive and Negative Affect Schedule-Expanded Form.

Table 3

Example Contingency Management Pay Schedule

	--	1	2	3	4	5	6	7	8	9	10	11	12	13	Total
Weekly Survey	--	$2	$2	$2	$4	$4.50	$5	$5.50	$6	$9	$10	$11	$12	$13	$86
Weeks 4, 8 Survey	--	--	--	--	--	--	--	$4	--	--	--	$4	--	--	$8
Psychological/Medical Evaluations	$50	--	--	--	--	--	--	--	--	--	--	--	--	--	$50
Sub-Total	$50	$2	$2	$2	$4	$4.50	$5	$9.50	$6	$9	$10	$15	$12	$13	$144
Bonus	--	--	--	**$10**	--	--	--	--	--	--	--	--	--	**$10**	$20
Total	$50	$2	$2	$12	$4	$4.50	$5	$9.50	$6	$9	$10	$15	$12	$23	**$164**

Note: Participants must have completed all components to earn payment for that week. Blue shading denotes the baseline phase. Yellow shading denotes the first five modules of the intervention phase. Green shading denotes the second five modules of the intervention phase. Reimbursement increased between completed weeks by $0.50 for the first five modules and by $1.00 for the second five modules. The numbers in the top row represent each week of study enrollment. This example demonstrates the pay schedule for a participant who was enrolled for 13 weeks and who completed 100% of the study components (i.e., minimum study duration).

Table 4

Participant Demographic Data

	Participant 1	**Participant 2**	**Participant 3**
Age	18	23	26
Race	White	White	Asian Indian
Ethnicity	Non-Hispanic	Non-Hispanic	Non-Hispanic
Gender	Female	Female	Male
Sexual Orientation	Bisexual	Gay	Gay
College-Level Study	<1	5	6+
BMI	29.8	28.1	27.1
Diagnosis	BN	BN	OSFED

Note: College-Level Study=Years of college-level study completed at the time the screening survey was completed; BMI=Body Mass Index at the psychological evaluation; BN=bulimia nervosa; OSFED=other specified feeding or eating disorder. Participant 3's OSFED was characterized by subjective binge-eating episodes and restricting, which can be referred to as 'compensatory eating disorder' (i.e., a specific type of OSFED).

Table 5

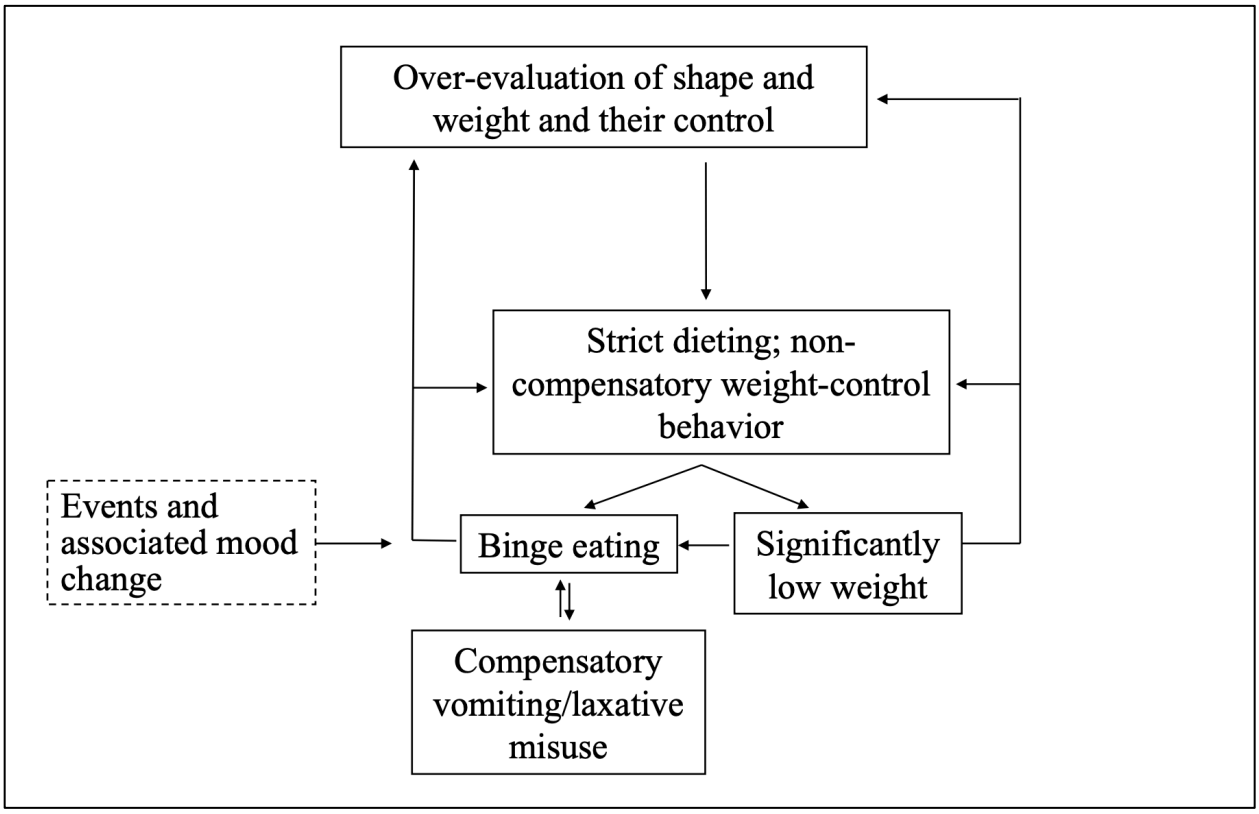
Raw Scores and Normed Means

	BE			R			EE			P			BD		
Participant	**bL**	**Tx**	**Norm M (*SD*)**	**bL**	**Tx**	**Norm M (*SD*)**	**bL**	**Tx**	**Norm M (*SD*)**	**bL**	**Tx**	**Norm M (*SD*)**	**bL**	**Tx**	**Norm M (*SD*)**
1	17.33	5.00	11.21 (*5.45*)	0.33	1.00	5.44 (*4.53*)	4.00	7.33	6.90 (*5.07*)	0.00	0.00	0.86 (*2.13*)	20.00	19.00	13.68 (*6.89*)
2	22.00	8.00	11.21 (*5.45*)	4.33	2.67	5.44 (*4.53*)	4.33	0.33	6.90 (*5.07*)	0.00	0.00	0.86 (*2.13*)	18.33	12.00	13.68 (*6.89*)
3	5.67	2.33	11.52 (*7.12*)	0.33	1.00	5.83 (*5.18*)	0.00	3.00	4.36 (*4.85*)	0.00	0.00	1.85 (*3.56*)	8.67	12.33	15.45 (*7.69*)

Note: All scores reported are raw scores; BE=EPCOT Binge Eating; R=EPCOT Restricting; EE=EPCOT Excessive Exercise; P=EPCOT Purging; BD=EPCOT Body Dissatisfaction; bL=Mean of last three baseline-phase measurements; Tx=Mean of last three intervention-phase measurements; Norm=gender norms for college students and general psychiatric outpatients as reported in Forbush et al. (2013); M=mean; *SD*=standard deviation.

Figure 1

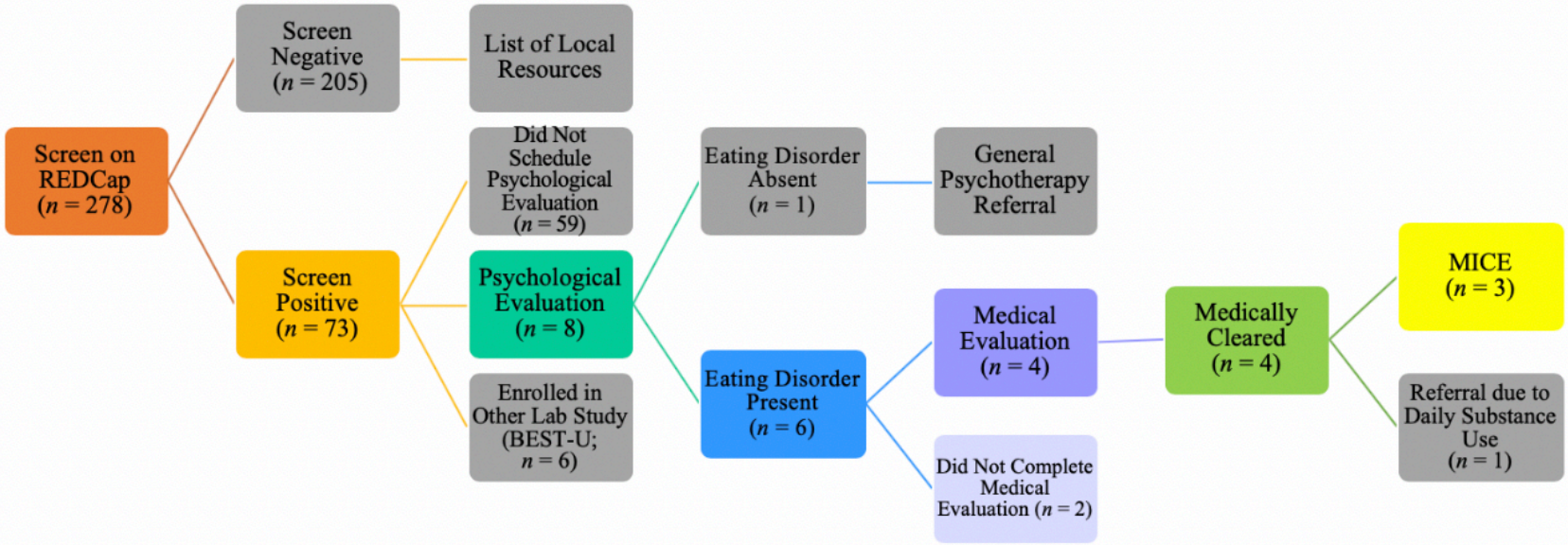
Transdiagnostic Cognitive-Behavioral Model of Eating Disorders



Note: This figure depicts the Enhanced Cognitive Behavioral Therapy for Eating Disorders (CBT-E) model of transdiagnostic eating-disorder psychopathology (CREDO, 2020); the "significantly low weight" part of the model did not apply to the current sample as persons with low body weight were excluded from the current study.

Figure 2

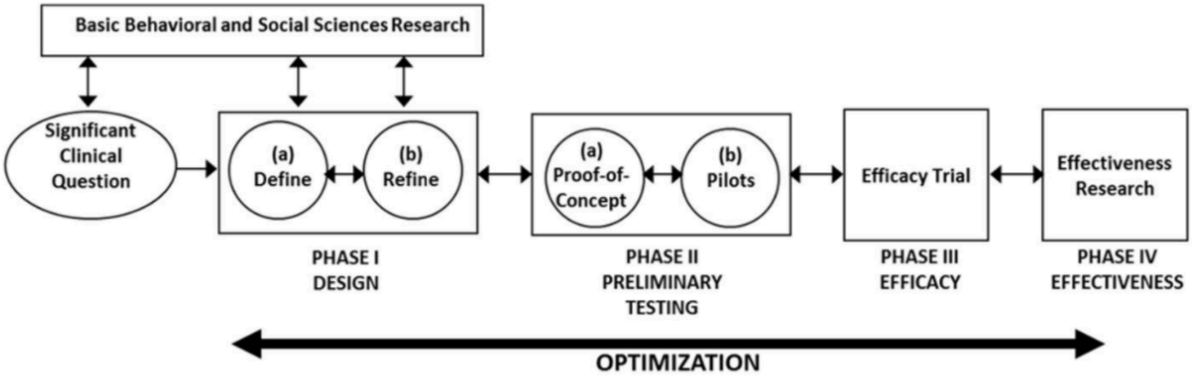
Study Recruitment Flow



Note: This figure shows the recruitment flow for the present study.

Figure 3

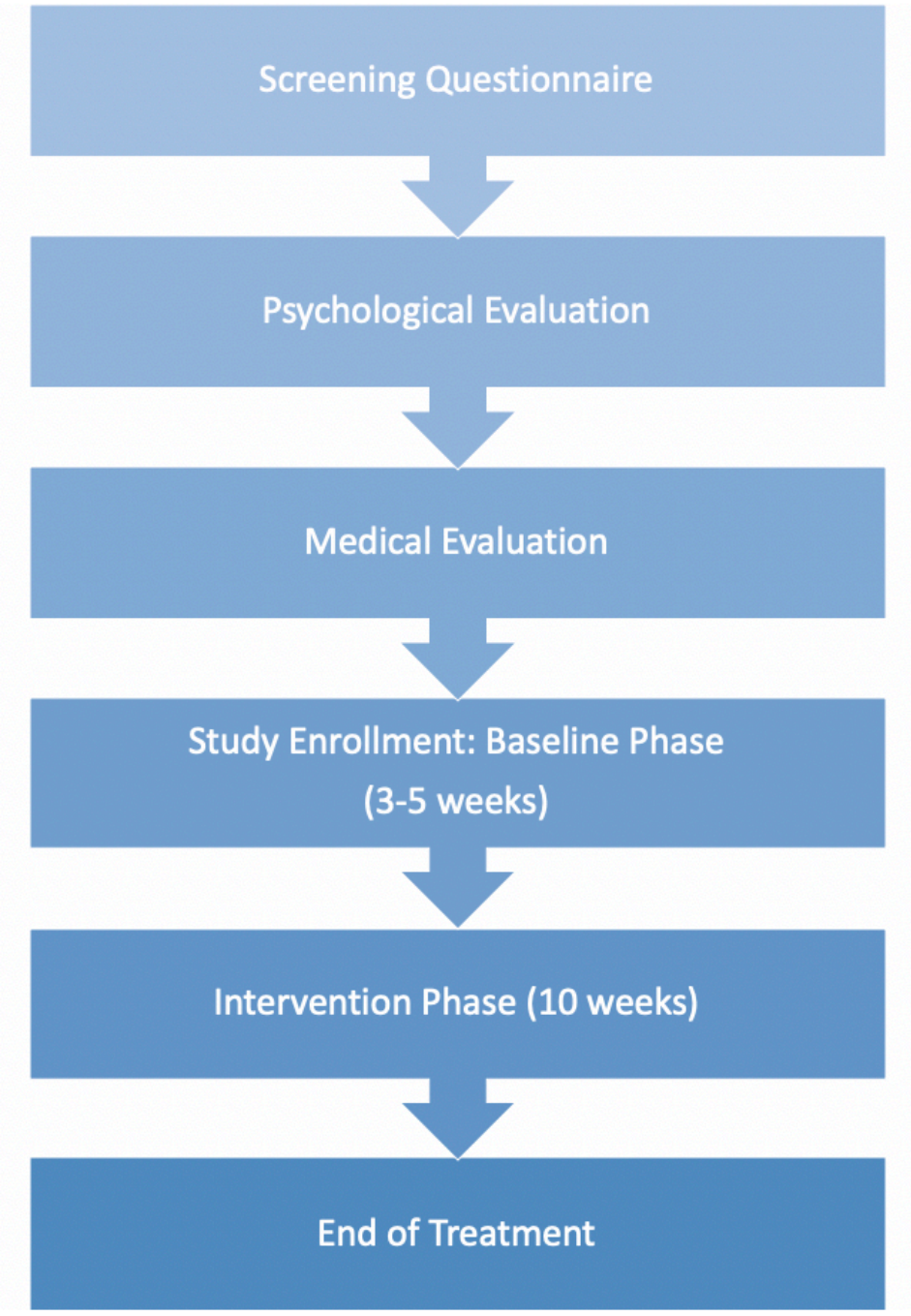
The Obesity-Related Behavioral-Intervention Trials (ORBIT) Model



Note: This figure, adapted from Czajkowski et al. (2015), shows the conceptualization and flow for intervention research studies.

Figure 4

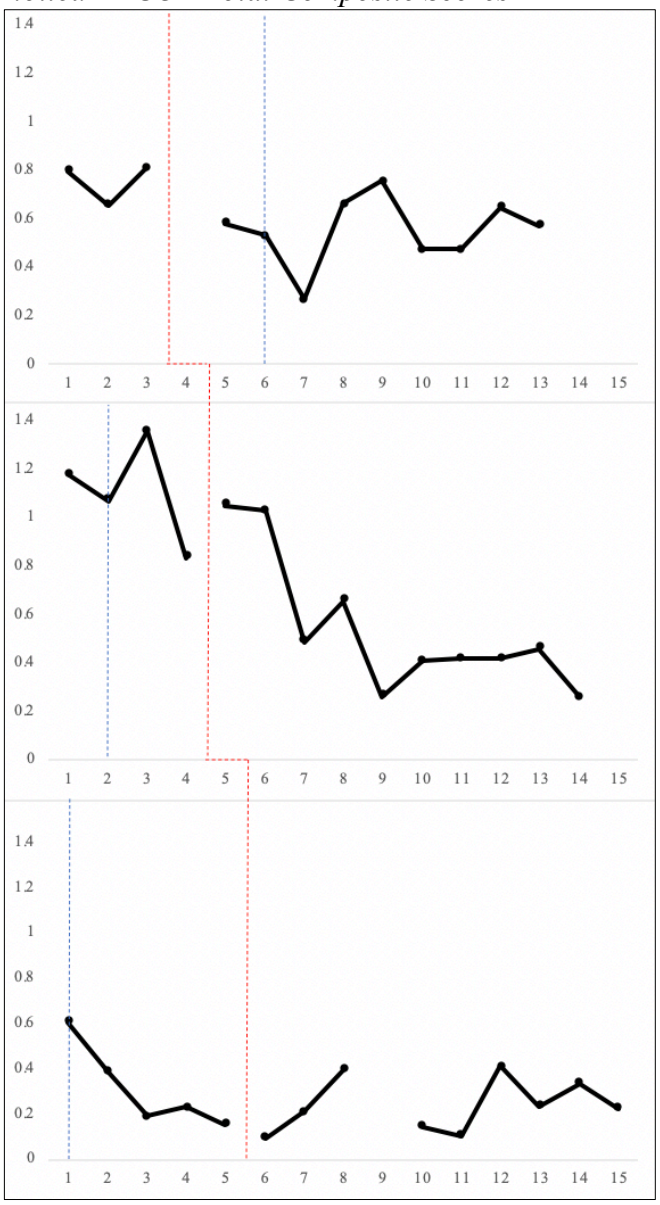
Study Flow



Note: This figure depicts the sequential steps each participant completed from the screening questionnaire to the end of treatment.

Figure 5

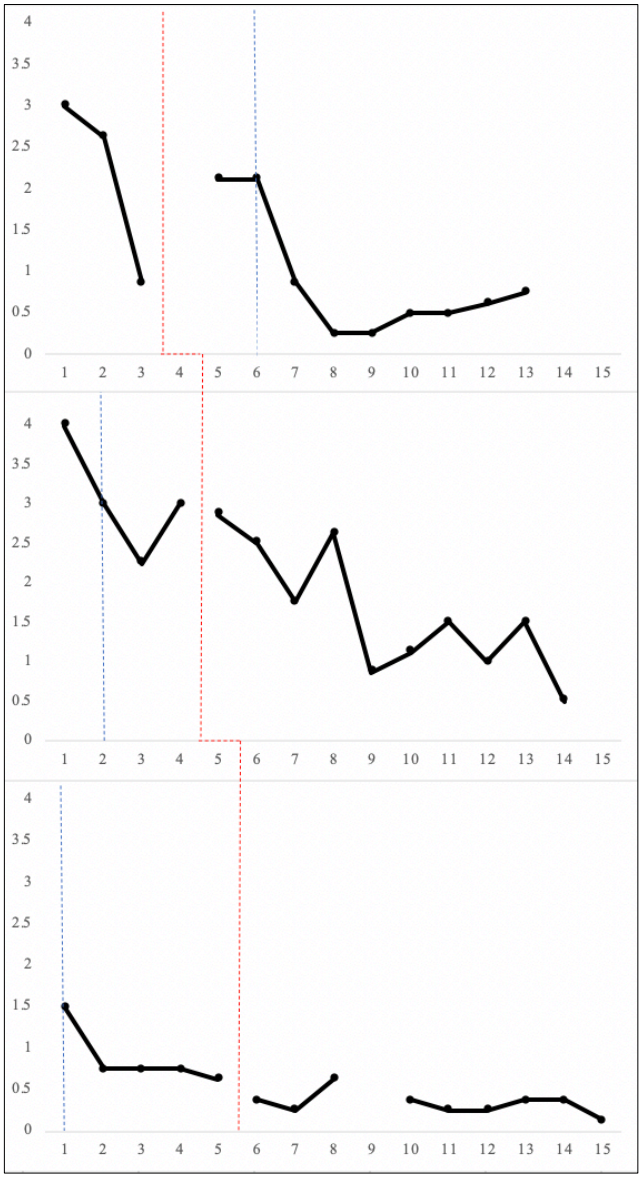
Plotted EPCOT Total Composite Scores



Note: X-axis values represent time in weeks; Y-axis values represent EPCOT Total Composite Scores. The red dashed line indicates implementation of the intervention for each participant. The blue dashed line represents the week within each participant's timeline when KU ceased on-campus operations for students. Participants 1-3 are graphed in descending order. Participant 1 had missing data for Week 1 of the intervention phase; Participant 3 had missing data for Week 4 of the intervention phase.

Figure 6

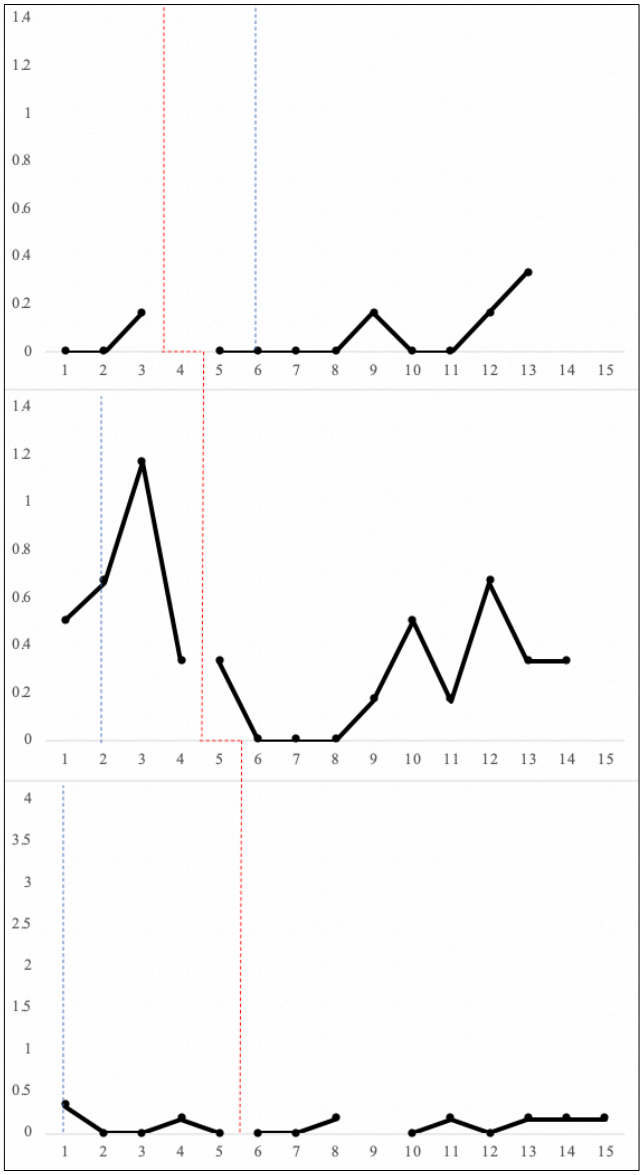
Plotted EPCOT Binge Eating Scores



Note: X-axis values represent time in weeks; Y-axis values represent EPCOT Binge Eating scores. The red dashed line indicates implementation of the intervention for each participant. The blue dashed line represents the week within each participant's timeline when KU ceased on-campus operations for students. Participants 1-3 are graphed in descending order. Participant 1 had missing data for Week 1 of the intervention phase; Participant 3 had missing data for Week 4 of the intervention phase.

Figure 7

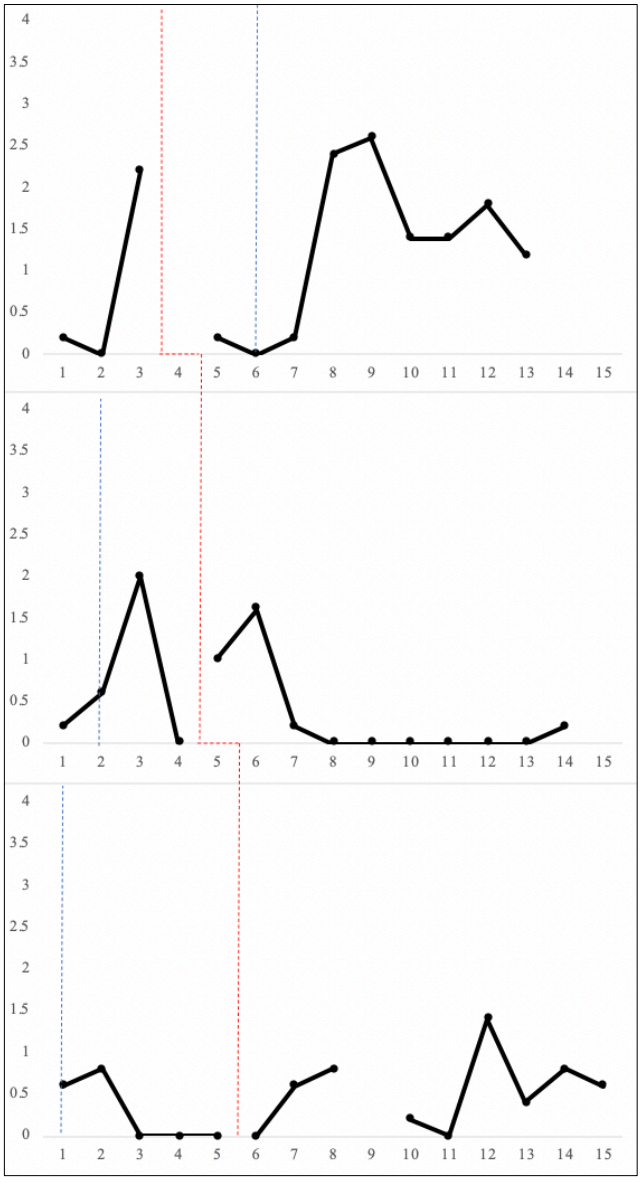
Plotted EPCOT Restricting Scores



Note: X-axis values represent time in weeks; Y-axis values represent EPCOT Restricting scores. The red dashed line indicates implementation of the intervention for each participant. The blue dashed line represents the week within each participant's timeline when KU ceased on-campus operations for students. Participants 1-3 are graphed in descending order. Participant 1 had missing data for Week 1 of the intervention phase; Participant 3 had missing data for Week 4 of the intervention phase.

Figure 8

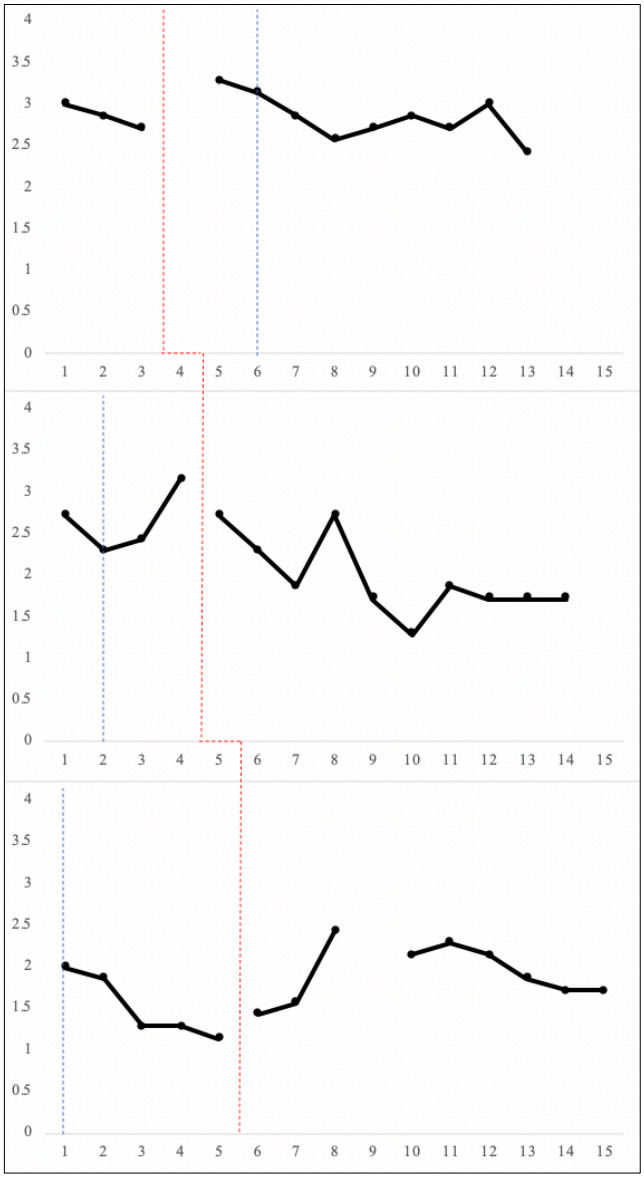
Plotted EPCOT Excessive Exercise Scores



Note: X-axis values represent time in weeks; Y-axis values represent EPCOT Excessive Exercise scores. The red dashed line indicates implementation of the intervention for each participant. The blue dashed line represents the week within each participant's timeline when KU ceased on-campus operations for students. Participants 1-3 are graphed in descending order. Participant 1 had missing data for Week 1 of the intervention phase; Participant 3 had missing data for Week 4 of the intervention phase.

Figure 9

Plotted EPCOT Body Dissatisfaction Scores



Note: X-axis values represent time in weeks; Y-axis values represent EPCOT Body Dissatisfaction scores. The red dashed line indicates implementation of the intervention for each participant. The blue dashed line represents the week within each participant's timeline when KU ceased on-campus operations for students. Participants 1-3 are graphed in descending order. Participant 1 had missing data for Week 1 of the intervention phase; Participant 3 had missing data for Week 4 of the intervention phase.