Spring 4-13-2022

Simple but not Easy: An Examination of the Test-Retest Reliability of the Perceived Stress Scale (PSS-10) and Exploration of a Meditation App for Stress Management with High School Students

Eileen Kaskons
ekaskons@lesley.edu

Simple but not Easy:
An Examination of the Test-Retest Reliability of the Perceived Stress Scale (PSS-10) and Exploration of a Meditation App for Stress Management with High School Students

A Dissertation Submitted by

Eileen Kaskons

In partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate School of Arts and Sciences

Lesley University

Cambridge, Massachusetts

May 21, 2022

Dissertation Final Approval Form
Division of Counseling and Psychology
Lesley University

This dissertation, Titled:

Simple but not Easy: An Examination of the Test-Retest Reliability of the Perceived Stress Scale (PSS-10) and Exploration of a Meditation App for Stress Management with High School Students.

as submitted for final approval by Eileen Kaskons under the direction of the chair of the dissertation committee listed below. It was submitted to the Counseling and Psychology Division and approved in partial fulfillment of the requirements for the degree of Doctor of Philosophy Degree at Lesley University.

Approved:

Adam Meiselman, Ph.D., Committee Chair, Assistant Professor Counseling and Psychology, Lesley University

Susan Gere, Ph.D., Committee Member, Professor Counseling and Psychology, Lesley University

Elizabeth Often, Ed.D., Committee Member, University of Massachusetts Lowell

Director of Doctoral Program in C & P. Lesley University: Peiwei Li, Ph.D.

Dean of GSASS: Sandra B. Walker, MBA

4-13-2022
Date of Final Approval

ACKNOWLEDGMENTS

First, I am indebted to the students who graciously participated and the health teacher and principal who allowed me to conduct this study. Second, I am truly grateful to my chair, Dr. Adam Meiselman, for all his time, guidance, and mentorship and to my committee members, Dr. Susan Gere and Dr. Elizabeth Often, for their wise counsel and ongoing support. Third, I sincerely appreciate the many kindnesses of my friends, Lesley professors, and fellow Lesley students as I worked to complete this project. Most of all, I cannot begin to thank my husband Peter, sons Max and Gus, and future daughter-in-law Lindsey enough for their love, encouragement, and faith in me. I love you and could not have finished without you!

ABSTRACT

Stress can have a negative effect on overall well-being, trigger physical disease, and contribute to mental health disorders. The 10-item Perceived Stress Scale (PSS-10) is frequently used in research. However, there are limited studies on the test-retest reliability of the PSS-10 with adults, and no studies could be identified that tested this issue with adolescents. Meditation apps are often recommended for stress reduction, but they have not been widely studied for use in high school. This study investigated the test-retest reliability of the PSS-10 and examined the feasibility of using a meditation app with high school students. First, a comprehensive review examined prior studies of PSS-10 test-retest reliability. Second, the test-retest reliability of the PSS-10 was investigated with a sample of high school students over 24-hour, 3-week, 6-week, and 9-week time intervals. Third, the test-retest reliability results from the high school study were compared to the results from comparable intervals in prior studies. Finally, the frequency of independent use of the meditation app was measured, and thematic analysis was used to explore the students' experience. Most significantly, this study appeared to be the first one to investigate the test-retest reliability of the PSS-10 exclusively with an adolescent population. Quantitative results showed that the PSS-10 demonstrated excellent 24-hour test-retest reliability but may not be stable for longer intervals. Frequency analysis confirmed that most students did not use the meditation app independently. However, evidence from qualitative findings indicated that the meditation app showed promise for use in a high school curriculum. This study provides several recommendations for the use, analysis, and interpretation of the PSS-10 with both adults and adolescents. It also outlines a plan for high school stakeholders who would like to advance the use of a meditation app.

Keywords: Perceived Stress Scale, test-retest reliability, mindfulness meditation

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Hans Selye (1978), widely regarded as the "father of stress research" (Tan & Yip, 2018, p. 170), stated, "Stress can be avoided only by dying" (Selye, 1978, p. 63). While some stress is unavoidable, and occasionally helpful (Selye, 1978), many people suffer from chronic stress, which greatly affects their well-being and prompts them to seek methods to alleviate it. The last few decades have seen an explosion of research and media coverage on the increased levels of stress and its significant negative effect on work, school, relationships, and almost every facet of our lives (O'Connor, Thayer, & Vedhara, 2021). Persistent stress can contribute to unhealthy physical conditions such as obesity and cardiovascular disease (O'Connor et al., 2021) and lead to an increased risk of developing mental health conditions such as chronic anxiety and depression (Sapolsky, 2004).

The World Health Organization (as cited in Konaszewski, Niesiobędzka, & Surzykiewicz, 2021) reported that over 50% of mental health conditions arise during the teenage years. In 2014, the APA survey *Stress in America: Are Teens Adopting Adults' Stress Habits?* measured and compared the stress of 1,950 adults over 18 and 1,018 teens ages 13–17 and noted the stress levels of all age groups are on the rise but are at unprecedented levels in teenagers (APA, 2014). Stress can significantly impact a student's ability to learn due to a decrease in attention, memory, and focus (Metz et al., 2013). However, stress in teenagers has routinely been neglected since adolescent health measurement is lacking (Guthold et al., 2021) and adult stress instruments may not be appropriate for use with adolescents (Byrne, Davenport, & Mazanov, 2007). While researchers have underscored the importance of using appropriate instruments to measure adolescent stress, they also have advocated for school-based intervention programs since teens may be reluctant to seek outside care for stress management (van Loon et al., 2020).

Although programs for anxiety and depression are plentiful, few school programs target stress, and those that do vary in effectiveness (Feiss et al., 2019; van Loon et al., 2020).

This study focused on the measurement of stress in high school students with a treatment and control group and examined the feasibility of a stress reduction intervention with the treatment group. Five components were investigated in this study. First, the psychometric properties of an adult-based widely used stress measurement instrument were explored. Second, the instrument was tested with an adolescent population, and the results were compared to those of adults. Third, a mindfulness meditation smartphone application (app) intervention was introduced, and the frequency of student compliance was measured. Fourth, data related to compliance with the intervention as a stress reduction method were examined for feasibility and used to determine further analysis. Finally, because compliance was not achieved, a research question regarding the effect of the intervention on stress levels was not analyzed. Instead, qualitative data exploring student barriers and motivations to use the intervention were examined.

Key Terms

Emerging adult: a transitional developmental phase that spans from the late teenage years into the mid-20s (Arnett, 2000).

Meditation: a practice in which a person may focus on a word, phrase, their breath, or the present moment in an alert, aware, and non-judgmental state (Sedlmeier et al., 2012).

Mindfulness: “the awareness that arises by paying attention on purpose, in the present moment, and nonjudgmentally” (Kabat-Zinn, 2006, p. 145).

Perceived stress: “the degree to which situations in one’s life are appraised as stressful” (Cohen, Kamarck, & Mermelstein, 1983, p. 385).

Psychometrics: "the theory and application of principles of psychological measurement and testing" (Holden, 2000, p. 417)

Psychometric scale: "used to capture a behavior, a feeling, or an action that cannot be captured in a single variable or item" (Boateng, Neilands, Frongillo, Melgar-Quiñonez, & Young, 2018, p. 1).

State-trait: "In psychological measurement, the distinction between trait (enduring or stable) components and state residual (variable or fluctuating) components" (Geiser, Götz, Preckel, & Freund, 2017, p. 219).

Test-retest reliability: "the systematic examination of consistency, reproducibility, and agreement among two or more measurements of the same individual, using the same tool, under the same conditions" (Aldridge, Dovey, & Wade, 2017, p. 208).

Significance of the Problem

Stress is a significant adolescent health issue that needs to be addressed to ensure a healthy transition to adulthood (Alderman & Breuner, 2019). Stress measurement is complicated, but it is important to seek valid measures to proactively identify elevated stress levels that could lead to the development of negative physical or psychological effects (Cohen, Gianaros, & Manuck, 2016) particularly with adolescents (Azzopardi, Kennedy, & Patton, 2017). Stress management education in high school is recommended (Massachusetts Department of Elementary and Secondary Education, 2018), but administrators have found it challenging to implement stress management into curricula and practice during the school day (Colbert, 2013).

Byrne, Davenport, and Mazanov (2007) noted adolescent stress was routinely assessed with instruments designed for adults and developed the 58-item Adolescent Stress Questionnaire to be used with teens. However, researchers argued that the Adolescent Stress Questionnaire was

too long to be easily and quickly administered (McKay, Andretta, & Perry, 2019). The 10-item Perceived Stress Scale (PSS-10) (Cohen et al., 1983; Cohen & Williamson, 1988) is a simple and brief instrument (Ali et al., 2021) and arguably the most widely used and established measure of psychological stress (Arza et al., 2019; Makhubela, 2020). The PSS-10 possesses adequate psychometric properties with adults (Cohen & Janicki-Deverts, 2012; Lee, 2012; Leung, Lam, & Chan, 2010; Roberti, Harrington, & Storch, 2006) and has been used in research with teens (Bluth, Roberson, & Gaylord, 2015; M. Braun, Levy, Collins, & Mogilner, 2014; Foret et al., 2012; Kohn & Milrose, 1993; Lemon & Watson, 2011; Siqueira, Diab, Bodian, & Rolnitzky, 2000; Wu et al., 2021).

Three versions of the Perceived Stress Scale—the PSS-14, PSS-10, and PSS-4 (Cohen et al., 1983; Cohen & Williamson, 1988)—exist, but the PSS-10 is the arguably the most popular and has been used in countless studies to measure perceived stress and gauge the effect of interventions (Galante et al., 2021). On his lab website, Sheldon Cohen (Carnegie Mellon University, 2015) has generously granted permission to use all PSS versions for nonprofit academic research or nonprofit educational use. A review of the literature, which focused on the PSS-10 because it was the version used in this high school study, showed limited research and conflicting results on the temporal stability of the instrument. During development, Cohen et al. (1983) predicted that the instrument would remain stable until approximately 4 weeks and provided data for 2-day and 6-week intervals. A statement on Cohen's lab website revealed that, except for the 2-day and 6-week time intervals from the original study of the PSS-14 (Cohen et al., 1983), no data have been collected for other time intervals (Carnegie Mellon University, 2015). Additionally, the website stated that scores on the PSS-10 are expected to become less accurate over time but asserted that the instrument should be stable over daily intervals.

However, no studies were provided or referenced to support this claim. Considering the popularity of the PSS-10, and its wide use in research over the past 30+ years, the dearth of information on temporal stability is surprising. In fact, several researchers have emphasized that the PSS-10 needs to be examined over a variety of time intervals to determine how long the scores will remain stable (Lee, 2012; Y. R. Miller, Medvedev, Hwang, & Singh, 2020; Roberti et al., 2006).

Although literature on psychometric data for the PSS-10 exists in studies with adults, there is a dearth of data with adolescents. Moreover, there is limited research on the test-retest reliability of the PSS-10 with adults, and almost all of the studies have been conducted with participants over 18 years of age (Lee, 2012). To date, no researcher has examined test-retest reliability with a predominantly adolescent population. Therefore, there was a need to analyze the test-retest reliability of the PSS-10 with adolescents, examine the effects of duration and magnitude, and compare the results to existing adult data to confirm reliability.

In a comprehensive review, Lee (2012) identified that PSS-10 test-retest correlation results met an acceptable level of > .70 in four studies that examined intervals between 1 and 4 weeks and noted that shorter intervals generally yielded more acceptable results than longer intervals. Lee (2012) also noted that the original test-retest reliability data of the PSS-14 (Cohen et al., 1983) showed an acceptable test-retest reliability of .85 for a 2-day interval but an unsatisfactory .55 for a 6-week interval. Similar to Cohen et al. (1983), who estimated the PSS-14 may become less predictive over time, Lee (2012) questioned the stability of the PSS for intervals longer than 4 weeks. Due to these concerns, and the limited availability of test-retest reliability reported in the literature, Lee recommended conducting studies that focus on revaluating PSS scores at various time intervals with an emphasis on checking scores at 6 weeks.

In addition, researchers recommended using the intraclass correlation coefficient (ICC) since it is a better way to evaluate test-retest reliability for continuous score instruments such as the PSS (Lee, 2012; J. Liu et al., 2016).

Furthermore, while test-retest reliability coefficients for the English version of the PSS-14 are known (Cohen et al., 1983), test-retest reliability results for the English version of PSS-10 are unknown. Miller, Medvedev, Hwang, and Singh (2020) noted that test-retest correlations for versions of the PSS-10 translated into other languages were at acceptable levels of > .70. After a thorough search of the literature, it appeared that test-retest correlation coefficients of the PSS-10 have only been reported in studies from non-English speaking countries with adults over 18 using translated versions of the scale. In addition, a test-retest reliability analysis of the PSS-10 with high school-aged adolescents has not been conducted in any country. The only study with an approximately similar age group (Chinese undergraduate students, $M = 18.3$) was conducted by Lu et al. (2017) who examined a Simplified Chinese translation of the PSS-10 with a 2-week interval and reported a test-retest reliability coefficient of .70. This noteworthy lack of information on the test-retest reliability of the English version of the PSS-10, particularly with high school students who may be vulnerable to physical and psychological damage from elevated stress levels, pointed to a significant gap in the literature that warranted attention.

Test-retest reliability is a valuable metric that is often used to determine if an instrument is reliable while also determining if it is a state or trait measure (Medvedev, Krägeloh, Narayanan, & Siegert, 2017). In psychological measurement, *state* refers to components that can vary or change; conversely, *trait* components are considered enduring or stable (Geiser et al., 2017). Test-retest reliability data on the PSS-10 may help establish if the instrument measures a fleeting state or an enduring trait. In addition, a confirmation of state or trait measurement may

help determine if an intervention, such as stress reduction, will be effective (Medvedev et al., 2017).

As previously mentioned, effective stress interventions for adolescents are needed. Wisner, Jones, and Gwin (2010) reported that high school administrators are searching for new methods to meet students' social-emotional needs and highlighted the dearth of stress research and methods for teaching stress reduction skills to adolescents. Studies have indicated that mindfulness meditation programs showed promise as a feasible intervention for high school students (Elder et al., 2011; Erbe & Lohrmann, 2015; Metz et al., 2013; Wisner et al., 2010). However, researchers noted that measurement of compliance was critical because consistent adherence to a meditation program may be essential for success and affect the effectiveness of the intervention (Antonson, Thorsen, Sundquist, & Sundquist, 2018; Flett, Hayne, Riordan, Thompson, & Conner, 2019; Foret et al., 2012; Goldberg, Knoeppel, Davidson, & Flook, 2020; Quach, Gibler, & Jastrowski Mano, 2017).

Smartphone apps are a promising intervention that may meet the need for an effective, scalable method to deliver mindfulness meditation training for stress management (Flett, Hayne, et al., 2019). Apps can provide an objective measure of adherence via electronic data (Flett, Fletcher, et al., 2019). Miller et al. (2015) revealed that college students wanted an app to help manage their stress, and Eva and Thayer (2017) reported that high school staff believed a smartphone app may motivate students to engage in mindfulness meditation practice at home. Specifically, researchers recommended investigating the feasibility of delivering mindfulness meditation instruction via a smartphone app with emerging adults and examining the effects on stress (Eva & Thayer, 2017; T. Miller et al., 2015). Moreover, the mindfulness meditation app Stop, Breathe & Think (SBT) (Stop, Breathe & Think, 2019), used in this study, had the capacity

to provide digital information to determine frequency of use. This information was essential since the ability to collect data to measure compliance has been critical for research (Flett, Fletcher, et al., 2019; Vettese, Toneatto, Stea, Nguyen, & Wang, 2009).

Theoretical Basis for the Study

Stress is a construct that is difficult to define and can be difficult to measure. Stress defies a consistent definition, and measurement instruments vary according to scientific discipline. For example, life event scales are used in epidemiology, global scales in psychology, and heart rate and blood pressure readings in biology (Cohen et al., 2016; Epel et al., 2018). The traditional definition of psychological stress arose from Lazarus and Folkman's (1984) stress and coping theory, which asserted that an event was deemed stressful if a person *perceived* they were unable to meet the demands of the situation (Epel et al., 2018). This study focused on the psychological aspect of perceived stress, which is commonly measured as a global appraisal on the Perceived Stress Scale (Epel et al., 2018).

Stress measurement notwithstanding, compliance is also difficult to measure and is a key component to consider in evaluating an intervention for feasibility. Self-report data can be subject to biases such as recall bias and response bias (Flett, Fletcher, et al., 2019). In this study, an app-based intervention was employed to avoid misreporting of compliance and aid in data collection. The mindfulness meditation app Stop, Breathe & Think (SBT) (Stop, Breathe & Think, 2019), which was rebranded to MyLife on May 11, 2020, was linked to either the Apple Health (iOS) or Google Fit (Android) app, which provided access to a history of meditation data for each student in the study. As previously mentioned, if students had been compliant with the intervention, the plan was to determine the effect on the stress level of high school students by comparing the treatment and control group scores on the PSS-10.

Relevant Literature

The adolescent students in this study were in a transitional stage marked by profound physical and cognitive growth (Dahl, 2004). This developmental phase spans the late teenage years to the mid-20s and has been identified as "emerging adulthood" (Arnett, 2000, p. 469). Many emerging adults face new social, academic, career, and financial responsibilities and may experience problems related to independence, family dynamics, career paths, social connections, and formation of identity (Kadison & DiGeronimo, 2005; Ramasubramanian, 2017).

Overall, adolescence and emerging adulthood can be overwhelming and stressful (APA, 2014). Juniors and seniors in high school face stress from interpersonal relationships, conflict with parents, identity formation, body image, and post-secondary pursuits (Seiffge-Krenke, Aunola, & Nurmi, 2009). Research has shown that individuals with increased stress levels are at a greater risk of developing depression and anxiety (Kass, 2017; Sapolsky, 2004). In fact, teenagers and emerging adults have consistently reported concerns related to mental health such as anxiety and depression, as well as ideations of suicide and/or self-harm, eating disorders, and substance abuse (APA, 2014; Bamber & Kraenzle Schneider, 2016; Bland, Melton, Welle, & Bigham, 2012; Byrd & McKinney, 2012; Crowley & Munk, 2017; Dvořáková et al., 2017; Greeson, Juberg, Maytan, James, & Rogers, 2014; T. Miller et al., 2015; Oman, Shapiro, Thoresen, Plante, & Flinders, 2008). A national survey conducted from 2005–2017 showed that there has been a significant increase in mood disorders and psychological distress among adolescents aged 12–17 (Twenge, Cooper, Joiner, Duffy, & Binau, 2019). Due to these possible negative effects on mental health, researchers have noted the critical need to measure the stress levels of adolescents (McKay et al., 2019).

As previously mentioned, researchers have frequently evaluated the psychometric properties of the Perceived Stress Scale with adults (Cohen & Janicki-Deverts, 2012; Lee, 2012; Leung et al., 2010; Roberti et al., 2006), but research is lacking with adolescents. Furthermore, researchers have emphasized the importance of test-retest reliability data and recommended using it to assess psychometric instruments for stability over time (Creswell & Creswell, 2018; McCrae, Kurtz, Yamagata, & Terracciano, 2011). Studies that analyzed test-retest reliability have reported results with three different coefficients, the Pearson product-moment correlation, Spearman rank correlation, and the intraclass correlation coefficient (ICC). Lee (2012) stated that the Pearson or Spearman correlation coefficients are often reported as measurements of association but argued that the ICC is a more refined approach for assessing test-retest reliability for an instrument such as the PSS. Lee's view is supported by other researchers who asserted the ICC is appropriate to calculate test-retest reliability when participants complete more than one self-report survey under the same conditions (Koo & Li, 2016; J. Liu et al., 2016; Perinetti, 2018; Vetter & Schober, 2018). In addition, test-retest reliability coefficients have been used to determine if an instrument measures a changing state or stable trait (Medvedev et al., 2017) and can be helpful to gauge the effect of an intervention.

Limited studies of mindfulness meditation interventions with adolescents have shown a promise in stress reduction and improvement in well-being, but there is a dearth of research with adolescents (Elder et al., 2011; Metz et al., 2013; Wisner et al., 2010). While there is a gap in the extant literature with adolescents, evidence presented from research with college students indicated mindfulness meditation may be useful. In studies with college students, researchers reported a decrease in stress and increase in well-being, as well as improved sleep and grades (Crowley & Munk, 2017; Dvořáková et al., 2017; Greeson et al., 2014; Oman et al., 2008;

Ramasubramanian, 2017; Warnecke, Quinn, Ogden, Towle, & Nelson, 2011). In particular, studies of mindfulness meditation with first-year college students, who are developmentally close to high school students, showed an increase in physical and psychological well-being (Bamber & Kraenzle Schneider, 2016; Dvořáková et al., 2017; Loi, Spencer, & Willam, 2008; Oman et al., 2008).

Specifically, delivering a mindfulness meditation intervention with a smartphone app may prove feasible and provide the ability to deliver instruction to a large population without the expense, time, and space associated with ongoing classroom instruction (Adams et al., 2018; Bostock, Crosswell, Prather, & Steptoe, 2019; T. Miller et al., 2015). Apps have shown promise with adults (Adams et al., 2018; Bostock et al., 2019; Champion, Economides, & Chandler, 2018), and researchers have reported that college students indicated they would use an app as a stress intervention (T. Miller et al., 2015). Moreover, studies have shown that 11% of teenagers and young adults have already tried a mindfulness app (Rideout, Fox, & Trust, 2018).

However, researchers have also noted that compliance with digital interventions can be low, and app-based interventions may not be ideal for everyone (Weber, Lorenz, & Hemmings, 2019). Concerns related to compliance include limited engagement (Gál, Ștefan, & Cristea, 2020), lack of completion rate (Mrazek et al., 2019), and discontinued use of the app (Psihogios, Stiles-Shields, & Neary, 2020). Furthermore, measurement of compliance is critical since researchers have cautioned that app-based interventions may result in increased attrition rates due to lack of in-person participation, distraction, lack of interest, and technical difficulties (Howells, Ivtzan, & Eiroa-Orosa, 2016).

In summary, the primary goals of this study were to examine the PSS-10 as a viable adolescent stress measure and determine the feasibility of using the SBT app as a stress reduction

intervention by assessing student compliance with use. Investigating the PSS-10 as a valid measure of stress with high school students, as well as introducing them to a free app that provided basic instruction in mindfulness meditation to reduce stress, may have had a positive impact on their well-being. An app was chosen to motivate students because it was portable and easy to use so students could engage in mindfulness meditation sessions at a time and place most convenient for them. In addition, data from the app assisted in determining meditation frequency to measure compliance. This simple stress reduction intervention had the potential to benefit students given the limited resources in time and funding that students, teachers, and school districts often have. Finally, based on a thorough literature review, this study was the first to investigate the test-retest reliability of the PSS-10 over different time intervals with an adolescent population and provided valuable information on the feasibility of implementing an app-based stress reduction strategy with high school students.

Problem Statement

To answer the research questions in this study, it was important to test an appropriate stress measurement instrument. Moreover, there was a need to calculate the ICCs of the English version of the PSS-10 with a high school population at different time intervals, examine the test-retest results for stability, and compare the results to existing adult data to confirm reliability. In addition, it was critical to investigate if the app was a feasible (i.e., accepted and used by students) stress reduction intervention by measuring student compliance. The results of these two research questions determined the direction of a subsequent research question that looked at the effect of the intervention or examined motivations and barriers to compliance. Specifically, this study examined stress measurement on the PSS-10 with an intervention and control group of high school students and introduced and measured the feasibility of use of the SBT mindfulness

meditation app as a stress reduction method with the intervention group. Compliance was a pivotal aspect to the feasibility component of this study. If compliance was achieved, the plan was to measure and compare the stress levels of the students in the intervention group to those of the control group to determine if the intervention had a statistically significant effect. If compliance was not achieved, qualitative information about students' motivations and barriers to compliance would be gleaned from an anonymous exit survey.

Research Questions

Three research questions were investigated in this study with a plan to pursue a fourth based on the results of the third question. First, what is known about the test-retest reliability of the PSS-10? Second, is the PSS-10 a reliable psychometric instrument to measure stress in adolescents? Third, is the free version of the SBT mindfulness meditation app a feasible stress reduction method (i.e., could be effectively introduced in the classroom and the students would use it independently at least 4 days a week)? If enough participants complied with using the SBT app four times a week, a fourth question would have sought to determine if meditation with the app had an effect on the stress levels of high school students. However, because participants did not comply, a fifth question sought to analyze qualitative data from an anonymous exit survey to determine students' motivations and barriers to using the app.

- **Research question 1:** What is the range of test-retest reliability of the PSS-10 with consideration to language, time interval, and age?
- **Research question 2:** Does the PSS-10 demonstrate good reliability when measuring stress in adolescents?
    - **Research question 2A:** What is the 24-hour test-retest reliability of the PSS-10 in the high school study's sample of adolescent students?

  - **Research Question 2B:** How does the test-retest reliability in the high school study change as a function of time, specifically focusing on 24-hour, 3-week, 6-week, and 9-week intervals?
  - **Research question 2C**: What is the test-retest reliability of the PSS-10 in the high school study compared to test-retest reliability of the PSS-10 in prior studies?
- **Research (exploratory) question 3:** Did the participants adhere to the baseline of compliance: Meditating using the smartphone meditation app for a total of 4 days per week during the 8-week intervention?
- **Research question 4:** If Q3 demonstrated better compliance, Q4 would have asked about the effect of the app-based mindfulness meditation program on the stress level of high school students as measured on the Perceived Stress Scale (PSS-10), compared to a control group.
- **Research question 5:** What were the barriers and motivations for student compliance to meditation with the app?

Summary

This study sought to examine the Perceived Stress Scale (PSS-10) as a stress measurement instrument and a meditation app as a stress reduction method with high school students. It consisted of five research questions and used quantitative and qualitative methods. The quantitative results and qualitative findings are discussed in Chapter 5. The research questions sought to investigate five key areas. The first question focused on a comparative review of prior research of the test-retest reliability of the Perceived Stress Scale (PSS-14 and PSS-10). The second question investigated the test-retest reliability of the PSS-10 over different time intervals with the current study's high school participants and compared the results to prior

research. The third question determined compliance by measuring the frequency of student adherence to the study protocol of meditating independently 4 times a week. As noted, the fourth question could not be answered but allowed for the development of a qualitative anonymous exit survey. The fifth question focused on feasibility and explored qualitative data focused on discovering the barriers and motivations students may have experienced with meditation, independent use of the app, and the study. The components of this study were designed to increase our knowledge of stress measurement, with a focus on adolescent stress, and investigate the feasibility of using a meditation app with high school students by exploring the details of their experience.

CHAPTER 2: LITERATURE REVIEW

This chapter addresses five themes. First, it defines stress and current concerns, with attention given to marginalized populations and the developmental needs of adolescents as emerging adults and traces the history of stress research from the perspective of physical and physiological response to one of psychological and cognitive appraisal. Second, it examines the Transactional Model of Stress and Coping (Lazarus & Folkman 1984) as a viable theoretical framework used in current research related to mindfulness meditation as a coping strategy to fight stress and increase well-being. Third, it explores the psychometric properties of the Perceived Stress Scale (PSS-10) (Cohen et al., 1983; Cohen & Williamson, 1988) as an instrument used to assess stress level. Fourth, it reviews the history of meditation brought from Eastern religious traditions and incorporated into Western methods and provides a current definition of mindfulness meditation. Finally, it investigates the emerging use of technology in the form of smartphone applications (apps) as a widescale delivery method for stress management with an emphasis on introduction in a high school setting.

Search Method/Keywords

A search of Web of Science for full text studies published in English during the years 2008–2018 using keywords “stress,” “mindfulness meditation,” “emerging adult*,” “smartphone,” and “app” produced zero results. However, removal of the words “smartphone” and “app” yielded 23 results. A search from 2008–2018 of peer-reviewed, full text, published in English, PsycINFO information using keywords “stress,” “mindfulness meditation,” and “emerging adult*” produced one article. Therefore, additional keywords of “high school,” “high-school,” “college and university,” and “student*” were explored to review material related to the targeted students and developmental level. In addition, PsycINFO and Web of Science databases

were searched from 2010–2021 for peer-reviewed, full text, published in English information using keywords “Perceived Stress Scale,” “PSS,” “PSS-10,” “psychometrics,” and “test-retest reliability.” Google Scholar and other sources were also examined.

Definition and Types of Stress

Although stress is subjective and can be difficult to describe, it is widely studied in many branches of biology and social science and not always regarded as harmful or negative. Indeed, life itself is not possible without some level of stress (Selye, 1978). Even when asleep, the body is required to meet energy demands to sustain life. The heart beats, muscles help the lungs respire, the digestive system processes food, and even the brain is active while dreaming. In fact, the only way humans can avoid stress is by dying (Selye, 1978).

Dr. Hans Selye introduced the term “stress” into popular use to describe the exposure of unpleasant conditions to animals, but the word is often used to describe psychologically harmful events that occur in all aspects of daily human life (Kabat-Zinn, 2013). The term also denotes and describes negative experiences within relationships, with friends or family, at school or work, or in situations involving physical or mental health. Although it has become part of everyday vocabulary, and often everyday life, the term “stress” as it is used today did not exist 100 years ago.

In a review of the history of stress research, Robinson (2018) explained the word stress is derived from the Latin verb *strictus*, which means “to draw tight,” and was used to describe a physical compression force on an object. The word evolved into the Middle English term “distress,” which described “hardship or force exerted on a person,” but the psychological term and aspect of stress was not used or studied until the late 1800s (Robinson, 2018, pp. 335–336).

Stress can be further classified as helpful, *eustress*, or harmful, *distress*, and categorized as either occurring in short *acute* doses or as a long-term *chronic* condition (Kabat-Zinn, 2013).

Eustress

The term eustress, literally "good stress," is defined as a low, short-term level of stress (Kabat-Zinn, 2013). Several researchers have noted that eustress may benefit a person by providing the necessary physical and psychological motivation to perform a task well, motivate learning, and may even result in euphoria when a difficult task is accomplished or a thrilling event is experienced (Bamber & Kraenzle Schneider, 2016; Crowley & Munk, 2017; Kang, Choi, & Ryu, 2009; Oman et al., 2008; Selye, 1978). However, not all stress is physically or psychologically helpful.

Distress

The term distress, or "bad" stress, describes stress that is damaging or difficult to bear (Selye, 1978). Robinson (2018) explained that early research on stress focused on physical causes such as dangerous materials, surgery, and extreme environmental conditions. The author pointed out that psychological stress was not recognized until after World War II when doctors realized that physical symptoms could be linked to the emotional damage seen in soldiers returning from battle. Kabat-Zinn (2013) stated that psychological distress can be categorized as *acute*, a generally brief but taxing routine occurrence (e.g., engaging in an argument or running late for an appointment), but certain experiences (e.g., a death in the family, job loss, or a serious accident) may result in long periods of acute stress. The author emphasized that if prolonged acute stress is not addressed to allow the individual to heal from the incident, it may become chronic. Additionally, Kabat-Zinn (2013) clarified that acute stress and chronic stress can have different immediate and long-term effects.

Acute Stress

From an evolutionary point of view, Sapolsky (2004) explained that acute stress has protective aspects and allows an organism to summon resources in case of an emergency. For example, when an animal is hunted by a predator, bodily systems are deployed as the animal prepares to fend off or evade the attack. Stress hormones adrenaline and noradrenaline immediately flood the bloodstream, and the hormone cortisol, which enters the blood more slowly and over a longer period of time, is released. Heart rate, blood pressure, and respiration increase, muscles are activated, senses are heightened, memory is sharpened, and perception of pain is diminished as the animal prepares to engage in battle or make a quick escape. In addition, activity that is not directly related to self-preservation is put on hold; food is not digested, and systems involved in growth, repair, development, reproduction, and immunity are suppressed. This is an example of fight, flight, or freeze, a primitive physical stress response that occurs whenever an organism is faced with a potentially life-threatening emergency (Siegel, 2012). Although we have evolved beyond the primitive state, the fight, flight, or freeze response is still encoded in our genes (Jensen & Nutt, 2015) and will be elicited whether the threats are real, perceived, or imagined (Siegel, 2012). No matter the response, Sapolsky (2004) emphasized that once the emergency is over (e.g., predator evaded, fire extinguished, or drowning child rescued), all systems should return to normal. However, as Goleman and Davidson (2017) asserted, life or death physical stress episodes are no longer common, and while most modern-day stress is psychological, it still triggers the same primitive stress response. In addition, the authors explained that some individuals experience or perceive stress as ongoing and, therefore, relentless.

Chronic Stress

Chronic stress can occur when an individual is constantly experiencing acutely stressful events and the stress response is repeatedly summoned. Sapolsky (2004) explained that repeated incidences of acute stress episodes can lead to high levels of stress over long periods of time, resulting in physical and psychological damage. In other words, “if you experience every day as an emergency, you will pay the price” (Sapolsky, 2004, p. 13). Research has shown that unmanaged chronic stress can impair an individual physically and contribute to high blood pressure, diabetes, heart disease, decreased immune response, and physical exhaustion (Adams et al., 2018; American Psychological Association [APA], 2014; Bostock et al., 2019; Oman et al., 2008). In addition to the physical toll, researchers have reported that high levels of stress can negatively impact an individual both psychologically and emotionally as chronic stress can adversely affect all levels of cognitive performance and academic achievement and contribute to anxiety and depression (Bamber & Kraenzle Schneider, 2016; Bland et al., 2012; Bostock et al., 2019; Crowley & Munk, 2017; Kang et al., 2009; Shapiro, Brown, & Astin, 2011). Researchers have reported chronic stress can damage areas of the brain associated with essential learning skills (e.g., imagination, memory, attention, and problem-solving) and impact the ability to self-regulate (Metz et al., 2013; Tang, Hölzel, & Posner, 2015). Although intended to be protective, stress is particularly damaging if an individual is unable, or perceives they are unable, to cope and the stress response itself becomes harmful and may have lifelong impact on physical and mental health (Lazarus & Folkman, 1984; Robinson, 2018; Sapolsky, 2004; Selye, 1978). Importantly, while stress is perceived on a personal level, it is often dependent on the larger environment a person may have experienced.

An individual’s larger environment may include social justice issues that are behind possible sources and origins of stress. As previously mentioned, Lazarus (2006) explained there

is a sociocultural component to stress, and individuals living with highly stressful situations (e.g., war, racism, immigration, unemployment, or poverty) may experience increased levels of stress. The author emphasized that it is important to look at the intersection of social systems and the individual difference between the concept of stress, regarding the larger social system, and the psychological concept of stress that is experienced by an individual or by the social group(s) to which they may belong. Some marginalized groups may experience increased levels of stress (e.g., women, members of the LGBTQ community, racial and ethnic minorities, those living in poverty, and individuals with disabilities). As previously mentioned, some emerging adults face increased stress due to developmental issues and may also belong to one or more marginalized populations.

Stress in Populations

Gender

The American Psychological Association (APA; 2014) reported that females are at an increased risk of elevated stress. The APA affirmed teenage girls reported a higher level of stress than teenage boys, and the rate of depression was 37% for girls and 23% for boys. Stress-related anxiety is higher for college-age women than men and can lead to a lifetime of stress-related diet issues (APA, 2014; Bamber & Kraenzle Schneider, 2016).

Mayor (2015) explained the socially constructed term *gender* differs from the physical and biological term *sex* as gender is related to events that happen to individuals identified as male or female. The author claimed incidents of stress occur more frequently in women than men and women perceive stress as more harmful. The author noted women who work in roles that are considered equal to men did not experience increased stress, but women who occupy positions considered less powerful exhibited increased levels of stress. Mayor also stated that in many

families, women are more likely than men to be in a caregiver role, which may create increased stress levels.

LGBTQ Community

Research has shown that members of the LGBTQ community are more likely to encounter stressful events in their life (Mozumder, 2017). Byrd and McKinney (2012) reported students considered different due to sexual orientation have an increased level of mental health issues compared to students who identify as heterosexual. In a study of first-year college students, Riley, Kirsch, Shapiro, and Conley (2016) stated those in the sexual and gender minority experience increased stress when they attempt to conceal their sexual identity due to shame, fear of bias, and intolerance or to prevent harm. In addition, the authors revealed that while LGBTQ and heterosexual college students encounter similar stressors, the LGBTQ students' stress may be compounded by the additional challenge of their sexual identity and therefore perceived as higher. As a result of increased stress levels, researchers have reported that LGBTQ students are more likely to become isolated, depressed, or anxious (Riley et al., 2016; Singh & McKleroy, 2011).

Race and Ethnicity

Research has indicated that students of color have increased levels of stress, and those attending predominantly white institutions are at the greatest risk (Byrd & McKinney, 2012; Colbert, 2013). In a study of high school students, Elder et al. (2011) reported Hispanic students experienced difficulty with language and other acculturation factors, African American students were exposed to increased violence, and American Indian students did not feel supported by their local school and battled cultural and identity issues. The researchers noted these students were at a greater risk of displaying negative academic and health behaviors due to these experiences.

At the college level, Kadison and DiGeronimo (2005) reported that international students may experience increased levels of stress due to discrimination, completing schoolwork in a second language, and acculturation factors. In addition, Lazarus (2006) indicated the stress of immigration is often rooted in erroneous speculations regarding race, ethnicity, religion, or the economy, which resulted in some of the native population members resenting newcomers. The author specified that the process of entering a new country, learning a new language, and other aspects of acculturation are rife with stress. Furthermore, Lazarus clarified that additional tension may arise from a mistaken, but frequently accepted, belief that immigrants are receiving money and other economic resources that are unavailable to the local population, particularly if the locals are poor.

Socioeconomic Status

Sapolsky (2004) maintained that belonging to a lower socioeconomic class is related to high levels of chronic physical and emotional stress and emphasized that individuals born into poverty do not have the financial resources to access proper care and are often operating in a constant vigilant state of crisis. The author noted that stress is not just the result of a lack of money but is also the result of being subjected to living in a society that allows poverty to exist. Importantly, Sapolsky (2004) stated people of lower socioeconomic status often lack the resources to access stress reduction activities in order to manage their stress. Similarly, Kabat-Zinn (2013) explained poverty can be related to other socially exploitive conditions such as living in toxic environments and being subjected to poor working conditions, which can result in increased stress and other health concerns. In addition, growing up in poverty is a significant source of stress for adolescents and may increase problems with behavior and decrease academic performance in school (M. Braun et al., 2014; Eva & Thayer, 2017; Mendelson et al., 2010).

Ability

Rhode, Froehlich-Grobe, Hockemeyer, Carlson, and Lee (2012) believed it is vital to address stress with disabled people as they experience a significant increase in stress-related health problems such as obesity, high blood pressure, diabetes, and sleep problems compared to those who are not disabled. The authors noted these conditions may affect the disabled at a younger age and have an impact during earlier developmental phases. Rhode et al. (2012) stated people with disabilities are at a higher risk of experiencing a broad range of stressors such as unemployment, poverty, and other health-related issues that may further compound stress. Kabat-Zinn (2013) explained disability stress can also impact those who are not disabled, since caregivers for disabled family members are also at a higher risk of experiencing chronic stress. In a study of adolescent students with learning disabilities, Beauchemin, Hutchins, and Patterson (2008) noted that a decrease in self-confidence or fear of failure may have caused students to experience elevated levels of stress.

Stress related to gender, sexual identity, race, ethnicity, socioeconomic status, or ability is not limited by age group. Before developmental stress is examined, it should be noted that there may be an intersection of stressors. Some individuals may belong to more than one marginalized group and therefore have increased stress depending on several different factors.

Emerging Adults

Emerging adulthood has been identified by Arnett (2000) as a distinct developmental phase that spans the late teenage years into the 20s. In support of the rationale for stress intervention with high school junior and senior students, it is logical to first examine the impact of stress through results of research conducted with populations that have graduated from high school. Although research on stress with high school students is examined, a focus on research

conducted with students at colleges and universities may provide evidence for a need to proactively intervene and mitigate stress before students enter college.

Recently, Arnett, Žukauskienė, and Sugimura (2014) augmented Arnett's (2000) original theory and stated emerging adults (EAs) can now be categorized in a stage of life development from the late teenage years until the age of 29. This range includes many high school students and most college students (Rogers, 2013). Whether they attend college or not, emerging adults are primarily concerned with exploring their identity and are subjected to numerous unpredictable and lifechanging events related to personal growth (Greeson et al., 2014; Rogers, 2013). In addition, Bland et al. (2012) stated those who choose to attend college will face new situations without their previous support systems of family, friends, or teachers. These changes may contribute to an increase in the stress levels of college students.

Indeed, chronic stress during the college years can result in low self-esteem, frustration, substance abuse, relationship difficulties, lack of engagement in school, poor academic achievement, decreased graduation rates, depression, and anxiety (Bamber & Kraenzle Schneider, 2016; Crowley & Munk, 2017; Dvořáková et al., 2017; Greeson et al., 2014; Kang et al., 2009; Oman et al., 2008; Ramasubramanian, 2017). In addition, continued high levels of stress hormones may impair functions related to learning such as paying attention, the ability to solve problems, and memory-related tasks (Metz et al., 2013; Sapolsky, 2004). These functions are necessary in most post-secondary endeavors but are particularly important for students who choose to attend college.

Once reserved for a select few, the college experience has become mainstream but may create additional stress. Bland et al. (2012) reported that 64% of women and 60% of men attend college after high school, and 85% enroll as fulltime students. In a survey of 117 colleges and

universities, Byrd and McKinney (2012) found that one third of students listed the stress of educational requirements, such as studying and taking tests, as the most prevalent issue affecting both health and academics. This increase in academic work, with less support in a new setting, can impact personal growth, transformation, and the exploration of individuality (Bland et al., 2012). In addition, Crowley and Munk (2017) reported students may experience stress due to learning difficult material in a short time while balancing other responsibilities such as jobs. The challenges of transition to college, adaptation to new surroundings, change in social networks, continued exploration of self, independence, and decisions regarding life goals can be stressful, create anxiety, negatively impact academics, and decrease well-being in all aspects of student life (Bamber & Kraenzle Schneider, 2016; Bland et al., 2012; Crowley & Munk, 2017).

Continued high stress levels in college students can contribute to substance abuse, create eating disorders, cause a decline in mental health, increase depression, negatively impact self-esteem, increase rumination, and decrease resilience (Bland et al., 2012; Oman et al., 2008). Greeson et al. (2014) stated approximately 50% of college students reported high levels of anxiety and depression, and 16.5% revealed suicidal or self-harm actions. In addition to anxiety and depression, a lack of stress management can lead to headaches, sleep problems, injuries, and colds (Bamber & Kraenzle Schneider, 2016; Oman et al., 2008). There is also indication that a high level of stress is correlated with an increase in autoimmune diseases, worsening of HIV symptoms, and cellular changes that contribute to physical aging (Oman et al., 2008). Byrd and McKinney (2012) confirmed that 95% of college counseling centers reported heightened levels of mental health issues in students. Given this increase in mental health issues reported by college counselors, research related to the experience of first-year college students is examined,

as this information may provide knowledge that could be used to develop programs designed to prepare students before they enter college as freshmen.

For example, Kass (2017) reported that while the period of change freshman students experience at the beginning of college life can be thrilling, it is also daunting, and this population is at risk to the danger of chronic stress. The author reviewed several surveys of psychological stress in college students and reported stress was evident at all levels, but there was an increased chance of heightened stress during the first year. Compared to older students, Dvořáková et al. (2017) noted freshman students consistently reported higher levels of stress than their upper-class peers, which may have resulted in poor coping strategies, negatively impacted relationships, and contributed to a decline in academic progress. Indeed, research showed that in addition to the developmental hurdles of emerging adulthood, first-year college students are also dealing with factors related to transitioning to a new environment such as forming new social connections; handling finances; attempting to juggle social, personal, and academic priorities; and determining career paths (Dvořáková et al., 2017; Loi et al., 2008; Ramasubramanian, 2017). While studies have highlighted the vulnerability of freshman college students to stress (Bamber & Kraenzle Schneider, 2016; Dvořáková et al., 2017; Loi et al., 2008; Oman et al., 2008), there is less information on the stress levels of the high school students who are poised to enter the freshman college class. Therefore, it is prudent to examine stress in high school students who have one foot in adolescence but are on the cusp of emerging adulthood.

Adolescents

In a policy statement from the American Academy of Pediatrics (AAP), Alderman and Breuner (2019) emphasized the importance of focusing on adolescent health issues to promote a healthy transition to adulthood and cited “toxic stress” as a factor to be addressed (p. 3).

Researchers have noted stress levels are on the rise in all age groups but are at unprecedented levels in teenagers (APA, 2014). The APA survey *Stress in America: Are Teens Adopting Adults' Stress Habits?* measured and compared the stress of 1,950 adults over 18 and 1,018 teens ages 13–17 and depicted a country suffering from increased levels of stress and lacking in effective coping strategies (APA, 2014). Overall, the APA survey detailed that high levels of stress and unsuccessful coping strategies seem to have become a way of life, promoting and sustaining harmful and lifelong unhealthy practices and maladaptive behaviors that may impact future generations.

Specifically, teenagers reported higher stress levels during the school year, 5.8 for teens compared to 5.1 for adults as measured on a 10-point scale (APA, 2014). The study noted that the top three sources of stress identified by teenagers were school (83%), decisions regarding college or post-secondary pursuits (69%), and family financial issues (65%). While these sources of stress may be unavoidable and require long-term management, the APA also examined coping skills and discovered many teens were unaware of strategies to manage stress, and those who were aware used stress management techniques infrequently. Statistics from the APA study showed 31% of high school-aged teens reported increased stress in the past year, and 34% estimated their stress levels would rise in the next year.

Although data on stress management in high school students are limited, the research that does exist is revealing. Stress may become more frequent during this developmental period as adolescents attempt to meet the cultural expectations of adulthood, and chronic stress can lead to poor physical health, increase in risk-taking, antisocial behavior, and emotional distress (Krapić, Hudek-Knežević, & Kardum, 2015). Emotional distress may contribute to reduced academic achievement, increased dropout rate, and decreased college acceptance (Colbert, 2013; Lemon &

Watson, 2011), particularly in students who belong to racial and ethnic minorities (Elder et al., 2011) or are exposed to poverty (Mendelson et al., 2010). Students who are unable to manage emotions in times of stress are at an increased risk of developing anxiety and depression, engaging in self-harm, abusing substances, and developing unhealthy sleep and eating behaviors (M. Braun et al., 2014; Feld & Shusterman, 2015; Metz et al., 2013). In addition, Mendelson et al. (2010) asserted that chronic stress during adolescence can affect brain development and emotional regulation.

Adolescence is a time of brain plasticity (Dunning et al., 2019). Giedd (2008) explained the increased "plasticity of the teen brain make adolescence a time of great risk and great opportunity" (p. 341). Executive functions (i.e., abilities like emotional regulation, judgement, attention, organization, goal setting, and planning) develop in the frontal lobe and can affect adolescent thinking and behavior but will continue to develop well beyond the teenage years until approximately age 25 (Arain et al., 2013; Giedd, 2008). These changes in the prefrontal cortex (PFC) allow for learning and adapting but may make it difficult for teens to make rational decisions (Blakemore & Choudhury, 2006). Additionally, because adolescent brains are malleable (i.e., still under construction), they can be influenced by external factors such as stress (Arain et al., 2013).

Siegel (2014) explained that stress during the teenage years can negatively impact the *pruning* process that shapes neural connections, resulting in damage to the brain's organization and balance. Research has shown that stress affects areas of the brain responsible for emotional control, memory, and problem-solving. For example, Eagleman (2015) explained that under stress, the amygdala, an area of the brain that controls anger, aggression, and fear, becomes activated and overrides other brain structures to deal with the stressor. In addition, the

hippocampus, a structure vital to memory creation, is damaged by the high cortisol levels that stress produces (Kass & Trantham, 2013). Furthermore, chronic stress can disrupt the function of the PFC and interfere with problem-solving and creativity (Liston, McEwen, & Casey, 2009). The PFC, housed in the frontal lobe of the brain, normally regulates and controls the amygdala by calmly filtering information to allow for a rational response, but stress may cause the PFC to temporarily go "offline," which can result in impulsive and harmful behavior (Siegel, 2014). Moreover, adolescents are particularly vulnerable to stress because the PFC does not fully mature until the mid-20s (Eagleman, 2015; Wu et al., 2021). Consequently, young adults, whose amygdalae are under less control due to their underdeveloped frontal lobes, are likely to respond rapidly to stressful situations with more acute emotions than adults, who can depend on the mitigating influence of their prefrontal cortex to regulate their anger and fear (Erbe & Lohrmann, 2015; Jensen & Nutt, 2015; G. C. Patton et al., 2016; Siegel, 2014). Additionally, young adults may be vulnerable to perceived stress (Wu et al., 2021) and may make impulsive decisions without the ability to rationally calculate the consequences (Eagleman, 2015).

Although increased stress levels are a concern, 100 years ago the current definition of stress did not exist, and the concept and understanding of the stress response, stress management, and coping strategies had not yet been discovered (Robinson, 2018). The definition of stress, types of stress, and evidence of increased stress in certain populations are important to note. However, in order to design an intervention to address stress, it is helpful to understand the evolution of knowledge regarding the stress response from physical to psychological. It is also helpful to historically and chronologically examine the research that has led to our current understanding of stress before reviewing possible stress reduction methods. Therefore, the

following review of the history of stress research highlights the development of one the most widely used theories of stress and coping.

History of Stress Research

Physical Stress

Selye (1978), one of the early stress researchers, explained and outlined the beginnings of the study of stress in his seminal book *The Stress of Life*, which was first published in 1956. The author recounted that during the late 1800s, "French physiologist Claude Bernard, at the Collège de France in Paris, taught that one of the most characteristic features of all living beings is their ability to maintain the constancy of their *internal milieu*, despite changes in the surroundings" (p. 12). As Selye explained, Bernard discovered the concept of homeostasis through observation that, even when subject to extremes of heat or cold, humans still manage to regulate their normal body temperature by the dilation and constriction of blood vessels. Selye explained that Bernard believed disease occurred when the body was unable to physically self-regulate.

In a comprehensive review of the history of stress research, Robinson (2018) explained that Selye is widely credited for borrowing the word stress, commonly used in physics and engineering to describe a force or pressure, to represent conditions in living organisms. The author recounted that Selye "defined *stress* as mutual actions of forces that take place across any section of the body, physical or psychological," renamed it General Adaptation Syndrome when he observed that rats subjected to a variety of miserable physical experiences developed symptoms (e.g., stomach ulcers, larger adrenal glands, and diminished immune tissue), and eventually coined the term "stress response" (Robinson, 2018, pp. 337–338).

While Selye is credited with identifying the psychological aspect of stress, Robinson (2018) confirmed Bernard's early contribution and chronologically highlighted key

developments in the study of stress. The author documented that, concurrent with Bernard's work, Canadian physician William Osler noted that internal physiological conditions may have a lasting impact on health and recognized that driven and determined individuals were more likely to experience heart disease. Robinson (2018) explained that while Bernard and Osler focused on the physical body, in the late 1800s psychologist William James began to study emotion—the feelings that can arise from physical experience—and the impact of adrenaline and introduced the psychological aspect of stress. The psychological effects of stress are explored next.

Psychological Stress

Robinson (2018) chronicled that the James-Lange Theory of Emotion was developed when James hypothesized that emotion was the result of the mind perceiving the physical effect of a stimulus and Danish physician Carl Lange posited that emotion was a subsidiary reaction to a stimulus. In the early 1900s, building on the work of Bernard and Osler, and disagreeing with but taking into consideration the emotional aspect of the James-Lange Theory, Harvard physiologist Walter Cannon connected stress to the experience of extreme duress and emotion in his work with soldiers during World War I. Through this work, Cannon was the first to postulate the role of adrenaline as a result of individual response to stressful situations and coined the term "fight or flight" (Robinson, 2018, p. 337).

While Cannon is responsible for identifying the concept of fight or flight, O'Connor, Thayer, and Vedhara (2021) noted that Selye's early work was a catalyst for an explosion of research into understanding the effect of stress and its impact on well-being. The authors stated that Selye believed adapting to stress was essential to life. Furthermore, Sapolsky (2004) noted that Selye drew two crucial conclusions from his early observations: First, no matter what the

stressful event is, the physical reaction is similar, and second, if the stress is sustained, it can result in illness.

Stress Theory

Robinson (2018) stated that psychologist Richard Lazarus (1966), in his seminal book *Psychological Stress and the Coping Process,* challenged the reductive stimulus-response explanation of Selye's General Adaptation Syndrome theory and focused on psychological stress and individual experience (p. 339). Robinson (2018) noted that Lazarus claimed psychological stress was unique, and there was a considerable difference in how individuals viewed or appraised stressful conditions because psychological stress involved personal meaning, which Lazarus called appraisal, and emotions. Furthermore, Lazarus believed this difference was due to independent appraisal, the variety of individual thoughts and motivations that came between the stressor and a personal reaction, which explained why one person may perceive a situation as harmless while another person may interpret an identical situation as stressful. Robinson emphasized that this insight laid the foundation for Lazarus to develop a theoretical model of a cognitive process of a person's mind interacting with their environment and that thoughts and feelings were an important aspect of the stress response.

Transactional Model of Stress and Coping

In the mid 1980s, Lazarus and his graduate student Susan Folkman released their influential book *Stress, Appraisal, and Coping* (Lazarus & Folkman, 1984) and acknowledged that stress in life was unavoidable but individual response varied. They asserted that psychological stress was determined first by cognitive appraisal, which is an assessment of a "transaction or series of transactions" (Lazarus & Folkman, 1984, p. 19) between a person and their environment, and second by their ability to cope. In other words, stress depended on how a

person interpreted their surroundings in relation to their well-being and if they believed they were capable of handling the situation.

In their model, Lazarus and Folkman (1984) identified two types of cognitive appraisal as *primary* or *secondary*. The authors stated that during *primary* appraisal, an individual decided if an event was irrelevant, benign-positive, or stressful. They clarified that a stressful situation was further evaluated as harm/loss, threat, or challenge. Harm/loss is the result of damage that has already occurred to the person, threat is the fear of harm or loss, and challenge is an opportunity to gain knowledge or skill. Lazarus and Folkman stated *secondary* appraisal involves deciding what to do, applying coping strategies, and evaluating the results. They explained that appraisal is at the heart of the transactional process as a bridge between stimulus and response and allows an individual the ability to be aware of, cognitively assess, and decide how to cope with stress. Lazarus and Folkman maintained that cognitive appraisal would result in two possible types of coping mechanisms: They believed that a person would either attempt to seek a solution, "problem-focused coping," or try to regulate their own inner response, "emotion-focused coping" (p. 44).

Lazarus and Folkman (1984) introduced and defined the concept of coping as "constantly changing cognitive and behavioral efforts to manage specific external and internal demands that are appraised as taxing or exceeding the resources of the person" (p. 141). In most cases, problem-focused coping is perceived as being within some control of the individual and is centered on altering or ameliorating the cause of stress, whereas emotion-focused coping is used when control of the stressor is not possible and involves "regulating emotional response to the problem" (p. 150). Lazarus and Folkman stated emotion-focused coping may result in *reappraisal*, whereby a person reconsiders the situation to make it less threatening (e.g.,

determining that things could be worse, or the situation was not really important) (p. 150). The authors emphasized that both problem-focused and emotion-focused coping must be understood in context, and any strategies employed are specific to the individual situation. This is consistent with Epel et al. (2018), who asserted that perceived stress is not the same as trait level depression or anxiety but is the result of specific conditions, and Siegel (2012), who explained the process of emotional appraisal is non-linear but occurs and reoccurs as information is processed and then recalibrated and can be reinforced.

Specific to coping strategies, Lazarus and Folkman (1984) explained that problem-focused interventions often failed to address the negative thoughts and feelings that may have arisen from previous personal traumas and believed that new interventions should be developed. They explained that cognitive behavior therapists often incorporate strategies for both mind and body and stated that "meditation…is commonly taught in programs whose goal is to lower tension and hence control stress" (p. 369).

When viewed with regard to Lazarus and Folkman's (1984) research, studies have shown that students who are mindful found it easier to adapt to the stress of personal, social, and academic demands. Ramasubramanian (2017) noted that students who practiced mindfulness methods, including meditation, used coping strategies that were both "problem-focused and emotion-focused" (pp. 309–310). Students who appraised stressful situations and incorporated coping techniques, learned through mindfulness meditation, into everyday activities saw a reduction in stress (Oman et al., 2008). In a study of work-related stress, Bostock et al. (2019) stated mindfulness could encourage "positive reappraisal of stressful circumstances as benign or meaningful" (p. 2) and therefore improve an individual's chance to heal from distressing circumstances. Researchers have reported that mindfulness was associated with an increase in

positive reappraisal that resulted in decreased stress (Garland, Gaylord, & Fredrickson, 2011; Hölzel et al., 2011). In a study of college students, Loi et al. (2008) reported that students who appraised stress in a negative manner and were not able to employ mitigating coping strategies were at a greater risk of depression and other health-related issues. Kang et al. (2009) further highlighted the importance of the relationship between an individual and the stressor and the importance of choosing their own stress management methods, specific to each individual and their own appraisal, even if their methods are different from others.

Additionally, researchers who study stress have used terminology from Lazarus and Folkman's (1984) theory. Sapolsky (2017) explained that reappraisal occurs when a reaction to a disturbing emotion is controlled by evaluating it in a new way. Kabat-Zinn (2013) stated it is "the *meaning* we bring to the transaction" that will establish whether an experience is determined as stressful or benign (p. 292). In addition, in a narrative review of 57 studies, Bamber and Kraenzle Schneider (2016) noted that while only 12 studies used a theoretical framework, Lazarus and Folkman's Transactional Model of Stress and Coping was the most commonly used. Therefore, the Transactional Model of Stress and Coping (Lazarus & Folkman, 1984) is used as a theoretical framework in this study. In addition, the instrument used in this study, the Perceived Stress Scale (Cohen et al., 1983), is based on the original work of Lazarus, which contributed to the development of the Transactional Model of Stress and Coping (Lazarus & Folkman, 1984).

The Perceived Stress Scale

Valid psychometric instruments are critical for effective research (Crosswell & Lockwood, 2020). Watson (1988) identified the construct of stress as a series of complex components that included a type of stimulus, a stressor, a response, psychological factors that influence the response, individual coping abilities, somatic responses, and observable changes in

behavior. Cohen, Gianaros, and Manuck (2016) confirmed that stress has been studied in many disciplines to address specific issues. For example, studies of stress in epidemiology have assessed individual life events, research in biology has explored the impact of stress on homeostasis and metabolism, and investigations in psychology have focused on individual perceptions of and reactions to stress. In addition, each of these disciplines has used different methods to measure stress. For instance, an epidemiologist may have employed life event scales; a biologist could have measured heart rate, blood pressure, or cortisol level; and a psychologist may have used subjective measures of specific social roles (e.g., work, marriage, or parenthood) or employed a global scale that was independent of specific experiences. Differences aside, Crosswell and Lockwood (2020) noted the primary goal of stress measurement was to advance, develop, and refine evidence-based instruments focused on the association between stress, health, and well-being in order to help individuals flourish in a stress-filled environment. This section focuses on the Perceived Stress Scale, devised by Cohen, Kamarck, and Mermelstein (1983) as a global stress measure. It outlines the development of the scale and details studies that examined general psychometric properties, investigates if the instrument is a state or trait measure, and compares studies of unidimensional and two-factor structure. Most importantly, it centers on studies of test-retest reliability (also known as retest reliability, temporal consistency, or temporal stability) as researchers have recommended that element of the scale requires further attention.

Development of PSS-14, PSS-10, and PSS-4

The original Perceived Stress Scale (PSS) (Cohen et al., 1983), now known as the PSS-14, was a 14-item instrument designed to measure the “degree to which situations in one’s life are appraised as stressful” (p. 385). Cohen et al. (1983) explained they developed the PSS-14

based on the early theoretical work done by Lazarus, who postulated that an individual's *perception* of stress plays a part in the stress response. The authors proposed that a psychometrically sound measure of perceived stress could contribute to research that studied the relationship between stress and disease. Currently, researchers studying stress have three versions of the Perceived Stress Scale available—the PSS-14, PSS-10, and PSS-4—but the PSS-10 is arguably the most widely used and has been translated into over 25 languages (Makhubela, 2020). Additionally, the PSS has been one of the most common instruments used by researchers to measure psychological distress and determine the effect of interventions (Galante et al., 2021)

During development of the PSS-14, Cohen et al. (1983) evaluated the psychometric properties in three studies. The first two of the three studies were conducted with 332 and 114 college students respectively. The third study included 64 participants in a smoking-cessation program. The authors found coefficient alpha reliability for the PSS-14 was .84, .85, and .86 in each of the three samples. They also examined test-retest reliability, identified the PSS-14 as a state measure, and explained that test-retest correlation coefficients were expected to be higher for short retest intervals than for longer ones. In support of this expectation, they confirmed a test-retest correlation of .85 for 82 college students who retook the PSS-14 within a 2-day interval. The authors reported the students were specifically instructed to aim for accuracy on the retest rather than for consistency across time. In contrast to .85 for the 2-day retest, Cohen et al. (1983) acknowledged the test-retest correlation was reduced to .55 for the 64 smoking-cessation participants who retook the PSS-14 after an interval of 6 weeks. In addition, the authors reported the PSS-4, consisting of four items with the highest correlation to the 14-item scale, was conducted with the smoking cessation group in follow-up telephone calls 1 month and 3 months

post-treatment. The test-retest reliability of the PSS-4 was .55 within the 2-month interval between calls.

As previously mentioned, Cohen et al. (1983) believed the instrument would become less predictive of stress-related health issues over time since stress levels are influenced by day-to-day difficulties, important events, and fluctuations in coping resources. Due to this, the authors surmised the predictive validity of the PSS-14 would diminish sharply after 4 to 8 weeks. In contrast to Cohen et al.'s (1983) beliefs, Epel et al. (2018) questioned the ability of global measures of perceived stress to predict health. Epel et al. asserted such measures are useful to gauge recent perceptions but do not accurately represent an accumulation of experiences and therefore are not a reliable estimate of long-term health outcomes. In fact, Cohen et al. (1983) emphasized that the PSS-14 was a state measure and was not considered to be predictive of any specific psychological pathology even though a stressful state may contribute to, or be associated with, certain psychological disorders.

In further investigation of the PSS-14 with a large sample of 2,387 respondents, Cohen and Williamson (1988) continued to analyze the psychometric properties but did not evaluate test-retest reliability. In this noteworthy analysis, the authors found that 10 of the 14 items had factor loadings of .48 or above compared to four items with lower factor loadings ranging from .39 to .11. This discovery led to the development of the PSS-10, which retained the 10 high factor items. In addition, Cohen and Williamson determined the PSS-10 possessed a superior internal reliability and reported a Cronbach's alpha coefficient of .78 compared to .75 for the PSS-14.

During the aforementioned analysis, Cohen and Williamson (1988) reported that the PSS-14 had been used as an outcome variable that included stressful events, coping, and personality

factors. The link between personality factors and stress, and conversely stress and personality factors, has been addressed in studies. For example, Cohen and Williamson (1988) stated that it can be difficult to tell the difference between perceived stress and psychological distress as there may be an overlap of the two. In a later article, Cohen et al. (1995) explained this overlap by using Negative Affect (NA) as an example. The authors defined NA as individual distress that can include a wide range of states such as anxiety, hostility, and depression and stated NA can be measured as either a brief change in mood or as an established difference in personal affective level. Similar to Cohen et al.'s (1983) identification of the PSS-14 as a state measure, Cohen et al. (1995) clarified that temporary fluctuations in negative mood are considered to be *state*, but an enduring negative affect that could continue for months, years, or an entire lifetime is considered to be *trait.*

State Versus Trait Measurement

The concept of state versus trait in psychological measurement has been widely discussed in the literature. In an editorial statement, Geiser, Götz, Preckel, and Freund (2017) explained that the difference between state components, those that can vary or change, and trait components, those that are enduring or stable, has received increased attention since 1980. Scherer (2005) emphasized that when researchers decide on vocabulary to characterize state or trait, it is vital to specify if the term is used to describe an episodic state or a personality quality. The author explained that terms such as irritable or anxious can be used to describe dispositions as well as fleeting moods or emotions. Specific to measures of stress, Barron and Gore (2020) asserted that the construct of stress has been defined and studied in various contexts and examined as both state and trait, but the definitions are often conflated. This issue can be particularly challenging since researchers need to rely on native language, as opposed to

scientific descriptors, and it is easy to confuse some terms because they can be used to describe a concept as either a state, trait, or both.

This dilemma is not new as the concept was exhaustively studied and categorized over 80 years ago by Allport and Odbert (1936), who distinguished between enduring personality traits and temporary states of mood and placed them into distinct and separate categories. In an illustration of their process, the authors used anxiety as an example. They explained that most people have experienced an anxious state at some point in their life, but some people suffer from anxiety neurosis, which is a true trait of personality. Counter to Allport and Odbert's (1936) categorization, Allen and Potkay (1981) argued that the distinctions between trait and state were not completely rigid and declared them to be arbitrary. In regard to instrument development, which is of prime concern to this section, Allen and Potkay (1981) further contended that if states and traits are considered to be nonarbitrary, three issues should be addressed. First, instruments that assess personality would need to be identified as state or trait but could not be both. Second, researchers would be required to clarify which of the two was being measured, and the research community would need to agree on the decision. Finally, if state measures were being used to obtain an indicator of trait, the researchers would need to determine where the state ended and the trait began. To be clear, Allen and Potkay referenced personality tests, and not instruments used to measure stress, but the concern and confusion is noteworthy and directly relevant to this project's focus on mindfulness meditation and stress as measured by the PSS-10. For example, Medvedev, Krägeloh, Narayanan, and Siegert (2017) noted an increase in studies using mindfulness interventions to alleviate symptoms and increase the ability to cope with stress. Similar to Allport and Odbert (1936), Medvedev et al. (2017) clarified trait as a relatively fixed pattern of behavior and state as a momentary experience or situation. Although the authors

measured mindfulness, not perceived stress, they emphasized that determining whether state or trait was being measured has become an important concern in research.

Specific to the Perceived Stress Scale, Miller, Medvedev, Hwang, and Singh (2020) emphasized that stress can be categorized as an enduring trait where an individual tends to experience stress across different circumstances or as a dynamic state dependent on specific situations or circumstances that elicit an instant stress response. More specifically, the authors contended that in conducting research, it was critical to determine if a measurement assessed state or trait. Miller et al. stated this was particularly important if the research included an intervention because a change in state would be temporary and the effects of the intervention would not last. Since this project employed an intervention, it was important to determine if the PSS-10 measured state or trait and consult the literature for research that addressed the concern.

The question of state versus trait was examined in the original Perceived Stress Scale development article and addressed in subsequent research. Cohen, Tyrrell, and Smith (1991) revisited the work of Cohen et al. (1983) and their original statement that the PSS-14 was a state measure. In a study that examined the relationship between stress levels and the risk of acquiring the common cold, Cohen et al. (1991) employed the PSS-10 to measure stress but also included personality measures of self-esteem, personal control, and introversion/extraversion. The authors hypothesized that stress levels may be partially due to differences in personality characteristics instead of purely limited to stressful elements in the surrounding environment. Interestingly, the findings indicated that none of the three personality traits accounted for the effect between stress and susceptibility to infection.

However, other studies of personality traits and stress have asserted that personality has influenced the stress response and affected individual perception, coping methods, and recovery

(Childs, White, & de Wit, 2014; Luo, Derringer, Briley, Roberts, & Mõttus, 2017). Specifically, researchers have asserted neurotic individuals tended to be exposed to interpersonal stress events more often, regarded such events as highly threatening, and appraised their ability to cope as low (Carver & Connor-Smith, 2010). In addition, several researchers reported a positive association between neurotic individuals and perceived stress and theorized that they possessed less effective coping strategies (Barron & Gore, 2020; Ebstrup, Eplov, Pisinger, & Jørgensen, 2011; Piekarska, 2020; You, Laborde, Dosseville, Salinas, & Allen, 2020).

Researchers have also asserted that perceived stress is a subjective state and is expected to increase and decrease over time, but personality traits are considered stable over time and can influence the way people perceive and cope with stress (Christensen et al., 2019; Luo et al., 2017). Further, studies have shown that states are context-dependent and reflect minute-to-minute responses to changes in environment as evidenced by fluctuations in thoughts, feelings, and behaviors (Roberts, 2018). In addition, Epel et al. (2018) asserted self-report responses on global stress scales are a snapshot of short-term experiences (e.g., 1 month) and do not accurately measure long-term experience but may indicate a trait-like pattern of an individual's response to stress. This research supports Cohen and Williamson (1988), who suggested that any correlation between a scale assessing perceived stress and a scale assessing psychological distress may be partially or completely due to the fact that a portion of the items in each scale measures a comparable or identical concept. Notwithstanding, it is critical for researchers to determine if an instrument used in their study measured a state or trait. As mentioned, Cohen et al. (1983) emphasized that the PSS-14 was a state measure. Although the scale has been widely used, Miller et al. (2020) reported research has not adequately addressed or identified the Perceived Stress Scale as a measure of state or trait. According to Medvedev et al. (2017), calculation of

test-retest reliability coefficients has historically been used during psychometric analysis to determine state or trait components in a scale (e.g., < 0.60 for state, > 0.70 for trait).

Test-Retest Reliability

Test-retest reliability of a psychometric instrument investigates the agreement or consistency of more than one measurement of a construct. Aldridge, Dovey, and Wade (2017) stated that test-retest reliability ensures that if a person repeats a test, when they are not expected to experience a change in the construct being measured, under the same testing conditions on two or more occasions they will produce similar results. The authors further explained that in psychological testing, it is important to use instruments with adequate test-retest reliability so that any change in the results will indicate that a true change has occurred in the individual and that a change in results cannot be attributed to an unstable instrument. Matheson (2019) reported that test-retest reliability results have been interpreted with great leniency. Matheson further explained that while many researchers have agreed that > 0.75 is considered excellent, ranges such as 0.60–0.74 may be interpreted as good, and 0.40–0.59 may be considered as either fair or good. However, Koo and Li (2016) argued that only values > .90 should be considered excellent, between .75–.90 good, between .50–.75 moderate, and .50 poor. Despite the lack of consensus, Schober, Boer, and Schwarte (2018) asserted that, while < 0.10 would be considered a weak correlation and > 0.90 would be considered strong, standard thresholds for test-retest reliability have not been firmly established and any interpretation should be approached with caution.

Psychometric Analysis of the Perceived Stress Scale

As previously mentioned, researchers have frequently evaluated psychometric properties of the Perceived Stress Scale such as internal consistency, reliability, and, less frequently, test-retest reliability. Researchers have emphasized it is important to examine test-retest reliability as

a component of an instrument's internal consistency and to ensure that items measure the same construct and recommended assessing test-retest reliability to check for stability over time (Creswell & Creswell, 2018; McCrae et al., 2011). In regard to reliability, Helms, Henze, Sass, and Mifsud (2006) stated Cronbach's alpha (alpha) is the most frequently used indicator. The authors explained alpha measures the degree to which responses to items are consistent with coefficients that can range from 0 to 1.00. In fact, the studies chronicled in this section have consistently reported alpha reliability. In contrast, and often causing confusion for the reader, studies that analyzed test-retest reliability have reported results with three different coefficients: Pearson product-moment correlation, Spearman rank correlation, and the intraclass correlation coefficient (ICC).

Reliability Coefficients

Schober, Boer, and Schwarte (2018) noted that both the Pearson product-moment correlation (r) and Spearman rank correlation (r_s) examined the relationship between two variables. They explained the strength and direction of the relationship ranged from -1 to +1, where 0 represents no relationship, but the relationship becomes stronger as it approaches an absolute value of 1. The authors clarified that the Spearman coefficient is basically a Pearson correlation that uses ranks instead of actual values, is helpful for use with nonnormally distributed continuous data, can be employed with ordinal data, and is less sensitive to outliers. Additionally, J. Liu et al. (2016) indicated that the Pearson correlation is appropriate when two variables follow a linear relationship, but Spearman can be used with non-linear relationships.

In their description of the similarities and differences between the Pearson, Spearman, and ICC, J. Liu et al. (2016) explained that the three correlations are used to gauge the strength of association between variables but vary slightly in concept, with Pearson and Spearman viewed

as measures of correlation and the ICC as a measure of agreement. J. Liu et al. asserted that a correlation measured with either Pearson or Spearman is primarily focused on change and can often measure constructs that differ greatly. In contrast, they emphasized the ICC measured agreement and focused on the degree of consistency, either between individual raters or between two or more assessments, and can include test results repeated in the same place or under the same conditions.

In a methodical review of the three versions of the PSS, Lee (2012) stated that the Pearson or Spearman correlation coefficients are often reported as measurements of association but argued that the ICC is a better way to evaluate test-retest reliability for an instrument such as the PSS, which has a continuous score. Lee's opinion is supported by J. Liu et al. (2016), who asserted the ICC is widely used as a measure of agreement for continuous outcomes and allows for differences in the means of the measures being examined.

While both the Pearson product-moment correlation and the ICC have often been used to determine test-retest reliability (Vaz, Falkmer, Passmore, Parsons, & Andreou, 2013), Yen and Lo (2002) pointed out that the Pearson correlation is designed to measure the relationship between different variables and therefore should not be used when two measures are taken from the same variable (e.g., test-retest reliability) and that the ICC is a more suitable correlation. In addition, Koo and Li (2016) declared a preference for the ICC since it indicates both correlation and agreement between measures. Accordingly, use of the ICC is appropriate to calculate test-retest reliability when raters are not involved, such as when participants complete more than one self-report survey under the same conditions (Koo & Li, 2016; Perinetti, 2018; Vetter & Schober, 2018).

Intraclass Correlation Coefficient

To date, there are 10 forms of ICC. Shrout and Fleiss (1979) defined the original six models, and McGraw and Wong (1996) added four more—two forms separated into interpretations for single and average scores (see Koo & Li, 2016 for a full description of all 10 forms). In calculating ICCs, IBM SPSS Statistics predictive analytics software—beginning with version 8.0—uses the McGraw and Wong (1996) system to identify the different forms (Hansen, Lehn, Evensmoen, & Håberg, 2016; Weir, 2005). Sainani (2017) recommended that authors should identify which ICC method was used in analysis, and Koo and Li (2016) cautioned readers to determine if the researchers used the correct ICC form before relying on the results of any study. Specifically, in reporting test-retest reliability, researchers have employed two ICC forms. For example, Matheson (2019) employed the two-way *mixed* effects, absolute agreement, single rater/measurement, while Hansen et al. (2016) used the two-way *random* effects, absolute agreement, single assessment. Hansen et al. explained that while the results of mixed or random effects will be the same, if the researcher decides to use random effects, the calculation can be generalized to a larger population while mixed effects results must remain within the confines of the study. In agreement with Matheson (2019), Koo and Li (2016) and Qin et al. (2019) asserted that the two-way mixed effects is the appropriate selection for test-retest reliability because repeated tests using the same instrument at the same time (i.e., repeated measures) are not random.

Test-Retest Reliability of the Perceived Stress Scale

In a comprehensive review, Lee (2012) examined 19 articles that investigated the psychometric properties of three versions of the Perceived Stress Scale: PSS-14, PSS-10, and PSS-4. The author reported Cronbach's alpha met the minimum measure of internal consistency of > .70 for both the PSS-14 and PSS-10 but did not for the PSS-4 since it was < .70. During the

review, Lee (2012) examined test-retest reliability and discovered that it was reported in only six of the studies (Almadi, Cathers, Hamdan Mansour, & Chow, 2012; Chaaya, Osman, Naassan, & Mahfoud, 2010; Cohen et al., 1983; Remor, 2006; Siqueira Reis, Ferreira Hino, & Romélio Rodriguez Añez, 2010; Wongpakaran & Wongpakaran, 2010). Three of the six studies that investigated test-retest reliability assessed the PSS-14, four of the six assessed the PSS-10 (one study assessed both), and none of the six assessed the PSS-4. Furthermore, three studies employed either the Pearson or Spearman correlation coefficient, and three studies used the ICC. As previously discussed, Lee (2012) claimed the ICC was a more sophisticated approach (see Appendix A).

In the review, Lee (2012) astutely noted that test-retest correlation results varied, as the time interval between test and retest changed, and shorter intervals generally yielded more acceptable results than longer intervals. In the four studies that examined test-retest reliability of the PSS-10, with intervals between 1 and 4 weeks, an acceptable level of > .70 was met in all. Lee reported that test-retest reliability of the PSS-14 was examined in three studies. Two of the three studies had intervals of 2 weeks and reported an acceptable coefficient value of > .70. In the third study, which consisted of the original test-retest reliability data from Cohen et al. (1983), Lee (2012) reported an acceptable test-retest reliability of .85 for a 2-day interval and an unsatisfactory .55 for a 6-week interval. Similar to the concern expressed by Cohen et al. (1983), who estimated the PSS-14 may become less predictive after 4 to 8 weeks, Lee (2012) also indicated that “the stability of PSS might be less than six weeks” (p. 126). Due to concerns regarding stability and the limited availability of test-retest reliability reported in the literature, Lee recommended researchers continue to explore this and conduct “a systematic, longitudinal study of changes in PSS scores” (p. 126).

Although Lee (2012) included a study by Wang et al. (2011) in the review, the test-retest results of that study were not reported. Examination of the Wang et al. study showed the PSS-10 was administered to 240 Chinese policewomen and re-administered 2 weeks later to a random selection of 36 participants. In addition to examining the composite scores of the PSS-10, Wang et al. also reported on the subscales Factor 1 and Factor 2 of the PSS-10 and stated the Spearman correlation of the PSS-10 was 0.68 for the composite scale with 0.72 for the Factor 1 subscale and 0.63 for the Factor 2 subscale (see Appendix A for a list of all test-retest studies reviewed).

Factor Structure

Since the development of the Perceived Stress Scale, the issue of two factors has emerged as a discussion point. The original PSS-14 instrument was considered unidimensional by Cohen et al. (1983). However, Cohen and Williamson (1988) explained there were two factors and stated Factor 1 was comprised of negatively worded items associated with feeling upset or not in control, and Factor 2 consisted of positively worded items associated with successful coping and confidence. Even with an explanation of a two-factor structure, the authors indicated the distinction between factors was irrelevant for measuring perceived stress.

Many researchers have since disagreed and suggested the two factors are distinct and separate as indicated by the wording of the items (Makhubela, 2020; Sun, Gao, Kan, & Shi, 2019). For the 10-item PSS-10, Wang et al. (2011) clarified that Factor 1 consisted of negatively worded scale items 1, 2, 3, 6, 9, and 10, and Factor 2 consisted of positively worded scale items 4, 5, 7, and 8. Although researchers have agreed on the specific scale items that are negatively worded or positively worded, they have used different terms for Factor 1 and Factor 2 such as stress and counter-stress (Chiu et al., 2016) and “Perceived Helplessness and Perceived Self Efficacy” (X. Liu et al., 2020 p. 2).

Michaelides, Christodoulou, Kkeli, Karekla, and Panayiotou (2016) found very small distinctions between the two factors. Makhubela (2020) reported that many researchers have supported the two-factor model over the unidimensional one but conceded that studies have also found the differences to be minor, and in some instances conflicting, which may support Cohen and Williamson (1988), who originally suggested a composite score for the instrument should be used. Although factor structure of the PSS-10 was recently examined by X. Liu et al. (2020) with Chinese adolescents, the study relied on the eight-year-old psychometric data reported by Lee (2012), who indicated that research on test-retest reliability of the PSS-10 was insufficient and recommended that researchers continue to investigate this topic.

Research released following Lee (2012) that reported one, two, and/or composite factor structures was examined for results of test-retest reliability of the Perceived Stress Scale and to investigate if the issue of state and trait was addressed. Several subsequent studies have assessed the test-retest reliability of the PSS-14 and PSS-10 in many languages, with various age groups, in many countries, and with different test-retest time intervals. Similar to the review by Lee (2012), these studies reported test-retest reliability with three different coefficients. Of the three, the ICC was most frequently used, followed by the Pearson and Spearman. The following summary of studies includes those conducted after 2012 and details research of test-retest reliability of the Perceived Stress Scale. Studies that reported the ICC are examined first, followed by the Pearson and the Spearman.

As previously stated, the ICC was the most frequently used coefficient of test-retest reliability. Al-Dubai, Alshagga, Rampal, and Sulaiman (2012) explained the values of the ICC vary from 1 to 0, with 1 being perfectly reliable and 0 totally unreliable. In a study with 70 university students in Malaysia, Al-Dubai et al. (2012) reported a test-retest of the PSS-10, with

a 3-week interval, resulted in an acceptable (e.g., > 0.75) ICC of 0.82. Ben Loubir, Serhier, Battas, Agoub, and Bennani Othmani (2014) conducted research with a large Moroccan population using both electronic and paper versions of an Arabic translation of the PSS-10. The authors reported that test-retest results from 45 of the original 799 participants within a 1-week interval, using the paper version, showed an ICC of 0.91. A study of middle-aged nurses in Malaysia with the Malay PSS-10 conducted by Sandhu, Ismail, and Rampal (2015) retested 25 of the original 229 participants after 7 days and reported the ICC was 0.81.

Three recent studies used the two-factor model and determined the ICC for the entire scale and each of the two subscales. The first study examined the PSS-14, and the other two reported on the PSS-10. In a population of adults over the age of 70 living in the United States who did not suffer from dementia and were enrolled in an aging study to assess the possible predictive validity of stress on cognitive impairment, Jiang et al. (2017) examined test-retest of the English PSS-14 over a 1-year interval. The authors reported the ICC for the negatively worded (PSS-NW) scale (Factor 1) was 0.55, the positively worded (PSS-PW) scale (Factor 2) was 0.49, and the composite scale was 0.62. It is interesting to note that due to the longitudinal study design, Jiang et al. (2017) studied a test-retest interval of 1 year. As previously noted, Cohen et al. (1983) have stated shorter test-retests generally resulted in more significant correlations. In a second examination including the two-factor subscales, Khalili, Sirati nir, Ebadi, Tavallai, and Habibi (2017) studied the Persian translation of the PSS-10 with 100 adult patients attending a clinic for headache pain in Iran. Although the researchers reported intraclass correlations from 30 participants, no interval time between tests was given. However, Dr. Khalili confirmed a 2-week interval between test and retest (R. Khalili, personal communication, November 19, 2020). Khalili et al. (2017) reported the ICC of Factor 1 (Distress) was 0.95,

Factor 2 (Coping) was 0.90, and the entire scale was 0.93. A third study that included subscales of the PSS-10 was conducted by Sun et al. (2019) with 205 predominantly female Chinese patients, aged 18 to 66, who had systemic lupus erythematosus. The authors stated PSS-10 test-retest reliability, over a 7-day interval, was satisfactory and confirmed the ICC for the entire instrument was 0.954 with an ICC for Factor 1 of 0.820 and Factor 2 of 0.993.

Three research studies of the PSS-10 reported Pearson coefficients. The first study focused on subscales only, and the second reported the composite scale. The first, conducted by Chiu et al. (2016), examined 37 college student-athletes in Taiwan and test-retest of the Chinese PSS-10 over an 8–9-day interval. Chiu et al. explored the two-factor subscales of the PSS-10 and stated the Pearson coefficient for perceived stress was $r = .66$, $p < .00$ and counter stress was $r = .50$, $p < .00$, and they deemed the two subscales significantly reliable. In the second study, Lu et al. (2017) employed a version of the PSS-10 translated into Simplified Chinese, the primary language of mainland China, with 1,096 university students who were approximately 18 years old. Lu et al. reported a 2-week test-retest reliability of 129 subjects, randomly selected from the original sample, resulted in a Pearson correlation of 0.70 for the total scale, but they did not report subscales. Figalová and Charvát (2021) created and administered Czech versions of all three forms of the PSS (i.e., PSS-14, PSS-10, and PSS-4) to 1,725 participants aged 18–91 (M = 44.32). After 14 days, the researchers retested 159 participants with the PSS-14, PSS-10, and PSS-4 and reported retest scores of $r = .85$, .88, and .83, respectively.

Only one study was found that reported results with the Spearman correlation coefficient. Dao-Tran, Anderson, and Seib (2017) employed the Vietnamese PSS-10 with a sample of 28 participants aged 60–84 years. The authors reported most of the participants had limited formal education, were unemployed, had a low-middle income, and lived with a partner. Dao-Tran et al.

stated the test-retest Spearman was 0.43 with a 1-month interval. The aforementioned researchers reported test-retest reliability coefficients of the PSS-10. However, with the exception of Cohen et al. (1983), none of them stated if the PSS-10 measured state or trait.

It is important to highlight that Medvedev et al. (2017) attested that test-retest reliability coefficients have historically been used to determine state or trait components in a scale. However, a study conducted by Miller et al. (2020) used a data set from a study of 122 educators in Australia, randomized to treatment or control groups, on the effect of an 8-week mindfulness program on perceived stress. Miller et al. applied generalizability theory (G theory) on the temporal stability results to determine if the PSS-10 measured a dynamic state or stable trait. The PSS-10 was administered at three points: preintervention, postintervention, and 6 weeks postintervention. The authors specifically stated an accurate measure of stress, as a dynamic state or enduring trait, was essential to determine if the mindfulness intervention was effective. They noted that although Cohen et al. (1983) maintained the PSS-14 was a subjective measure, a determination of whether the PSS-14 measured state or trait had not yet been thoroughly investigated. Miller et al. (2020) agreed with Medvedev et al. (2017), who asserted that test-retest reliability should be used to distinguish between state or trait and concurred with them that coefficients are expected to be < .60 for a state scale and > .70 for a trait scale. Specifically, Miller et al. (2020) asserted the coefficient number would determine the overall effect of the intervention as a temporary state (i.e., ICC < .60) and would result in relapse, but an enduring trait (i.e., ICC > .70) would predict long-lasting change. In psychometric analysis of PSS-10 scores at preintervention, postintervention (8-week interval), and 6 weeks postintervention, Miller et al. reported strong internal consistency (e.g., alpha .85, .87, .87). The authors reported temporal stability across three occasions, measured with ICC, was .65 for the intervention group,

.69 for the control group, and .66 for the full sample but did not delineate test-retest reliability for each time interval. The authors acknowledged that although ICC measures were > .60, they were lower than anticipated for a trait measure (i.e., > .70) but stated the .69 ICC for the control group was narrowly acceptable.

In order to further determine if the PSS-10 measured state or trait, Miller et al. (2020) applied G theory, which addressed true person variance due to measurement error, and explained that a generalizability (G) score ≥ .80 indicates a trait measure. After removing the intervention group from analysis, Miller et al. reported relative and absolute G scores of .86 and .85, which indicated the PSS-10 was a trait measure. In addition, the authors experimented with the measurement design of the PSS-10, conducted a second analysis of the intervention group, reported relative and absolute G scores of .86 for that group, and concluded the PSS-10 measured trait. Significantly, prior to their study, Miller et al. discovered a gap in the extant research and stated, "to the best of our knowledge, state and trait measure of perceived stress is not available to date" (p. 2).

In examining the aforementioned studies, it is also important to note that only one study (Lu et al., 2017) reported test-retest results for a predominantly adolescent population (18-year-old university students). Although the Perceived Stress Scale was not specifically designed for adolescents, Cohen et al. (1983) explained it was intended for use in populations with a minimum of a junior high school education. The items and response options are simple to understand and designed to be "free of content specific to any subpopulation group" (Cohen et al., 1983, pp. 386–387). Cole (1999) investigated differential item functioning of the instrument with regard to education, race, and gender and reported the PSS-10 was appropriate for general use. In addition, several researchers have used the PSS-10 in studies with adolescents (Bluth et

al., 2015; M. Braun et al., 2014; Foret et al., 2012; Kohn & Milrose, 1993; Lemon & Watson, 2011; Siqueira et al., 2000). As indicated at the beginning of this section, the PSS-10 is a well-known and highly regarded stress measurement instrument that is currently widely used. Recently, researchers have examined its psychometric properties and confirmed that the PSS-10 has exemplary internal consistency across many studies (Creswell & Creswell, 2018), is considered generally valid (Taylor, 2015), and possessed an acceptable reliability of > .70 (Makhubela, 2020). However, test-retest reliability data are still not routinely measured or reported.

Mindfulness Meditation

The Transactional Model of Stress and Coping (Lazarus & Folkman, 1984) and the application to mindfulness and perceived stress is further supported by neuroscientific findings. For example, Sapolsky (2004) posed the question, “are there ways to change the world around us and to alter our perceptions of it so that psychological stress becomes at least a bit less stressful?” (p. 395). Sapolsky acknowledged research that indicated daily meditation was helpful in reducing the physical stress response during meditation but was quick to question if this reduction in stress was due to meditation itself or was a result of the personality of those who chose to meditate. Researchers Goleman and Davidson (2017) have pursued a similar line of inquiry in their work. They examined whether the enjoyable experience that many participants claim meditation produces can create a long-term change in individuals even when they are not meditating. The authors explained their research in meditation was rooted in the search for evidence of a long-term change in behavior. They defined this possible change as “An altered trait—a new characteristic that arises from a meditation practice” and explained that it “endures apart from meditation itself” (p. 6). Goleman and Davidson also noted that the process and

research behind their study of state and trait aspects of meditation was built on the work of other neuroscientists.

In their groundbreaking book *Altered Traits: Science Reveals How Meditation Changes Your Mind, Brain, and Body*, Goleman and Davidson (2017) credited neuroscientist Bruce McEwen, who presented evidence that stressful events can diminish neural areas in the brain and produce lasting, damaging, neurological changes, and neuroscientist Marion Diamond, whose research showed that positive experiences can increase neural growth, particularly in the areas of the brain that involve focus and the ability to regulate emotion. Goleman and Davidson stated that these early findings led to the concept of neuroplasticity and the hypothesis that negative and positive experiences can alter brain structure and result in a functional change in the brain. Kabat-Zinn (2013) reported evidence that supported those early findings when he cited two studies conducted at Massachusetts General Hospital and Harvard University that used fMRI, a brain scanning technology, to study structural changes in the brains of individuals who had completed 8 weeks of a Mindfulness Based Stress Reduction (MBSR) practice. He stated that after completing the MBSR course, the results indicated there was an increase in areas of the brain correlated with acquisition of knowledge, feelings, and self-awareness. In addition, Kabat-Zinn reported the studies demonstrated a decrease in the amygdala, a structure in the brain used to appraise stressful, threatening, and emotional events, and the decrease was consistent with improved scores on perceived stress.

Goleman and Davidson (2017) expanded this area of research and conducted studies based on previous neuroscientific findings. They explained that during practice, mindfulness meditation can produce a calm and beneficial state. However, they also claimed prolonged practice can change the brain and produce “altered traits” (p. 17). Accordingly, Bamber and

Kraenzle Schneider (2016) reported, “state mindfulness increases trait mindfulness, and higher levels of trait mindfulness reduces stress and anxiety” (pp. 3–4).

This concept of change in brain structure is important to consider in light of research on the brains of emerging adults. Neuroscientific findings indicated that until about age 25, the brain is not fully developed and undergoes a process of reconstruction that is particularly evident in the frontal cortex region (Eagleman, 2015; Sapolsky, 2017; Siegel, 2012). Eagleman (2015) explained that such changes in the brains of adolescents and individuals in their early 20s are vital for access to higher levels of thinking, judgement, and self-regulation and noted that impulse control is one of the last skills to develop. He astutely pointed out that this empirical information has long been understood by companies who sell car insurance, which is more expensive for young adult drivers, and the court system, which does not treat adolescent offenders the same as adults. In agreement, Kass (2017) noted the ability to self-regulate and to be self-reflective begins to arise during emerging adulthood. Roeser and Pinela (2014) described late adolescence as a “window of opportunity” for learning mindfulness meditation since there is an increase in neuroplasticity during this developmental period when a teenager’s identity is capable of being molded and habits and choices can be influenced (p. 10).

As previously noted, studies have shown that mindfulness meditation showed promise in helping individuals reassess their ability to face stressful situations (Bamber & Kraenzle Schneider, 2016; Bostock et al., 2019; Kang et al., 2009; Loi et al., 2008; Oman et al., 2008; Ramasubramanian, 2017). Goleman and Davidson (2017) reported that meditation practice over the long term appears to alter the brain by creating increased connections between areas of the brain that regulate emotional control and inhibiting or suppressing connections in areas of the brain associated with “wanting or attachment” (p. 252).

In summary, research has shown that the brains of emerging adults are still developing (Eagleman, 2015; Sapolsky, 2017; Siegel, 2012), and researchers asserted that mindfulness meditation may have a positive impact on brain function (Goleman & Davidson, 2017; Kabat-Zinn, 2013). Due to this information, research on mindfulness meditation is reviewed. In addition, studies that explored the use of mindfulness meditation as a possible method to provide individual stress reduction for students in the target population of high school students are examined. Although the subject of this review is mindfulness meditation, there are several other methods individuals may use to cope with stressful events. Researchers have shown that in addition to meditation, stress management can include journaling, engaging in exercise, listening to music, obtaining sufficient sleep, and enjoying social connections (Bamber & Kraenzle Schneider, 2016; Kabat-Zinn, 2013; Sapolsky, 2004). However, given the reported benefits and the focus of this study, a review of the history of meditation practice is presented followed by an examination of studies on mindfulness meditation.

History of Mindfulness Meditation

Goleman and Davidson (2017) explained that meditation, a type of contemplative practice, has been part of Eastern religious practices in India and other Asian countries and is an essential element found in all major spiritual, religious, and philosophical traditions. The authors explained that meditation was not originally designed to address stress but focused on an intense examination of the mind that was designed to move toward a deep and insightful change of self. They emphasized that religious forms are still practiced and should not be confused with secular forms.

Sedlmeier et al. (2012) described the two main distinctions in meditation as concentrative, which focuses on a word, phrase, or the breath, or mindfulness, which relies on

focusing on the present moment (and may include the breath) in an alert, aware, and non-judgmental state. The authors explained that both forms originated as spiritual practices; concentrative forms tend to be derived from Hindu, and sometimes Buddhist, traditions, and mindfulness forms are deeply rooted in Buddhism. Wisner et al. (2010) further clarified that concentrative mantra-based methods include the relaxation response, developed by Herbert Benson, and Transcendental Meditation (TM), developed by Maharishi Mahesh Yogi, while mindfulness meditation, such as that practiced in MBSR, relies on being aware of and accepting the present moment. Further details of the relaxation response, TM, and MBSR are described later in this literature review.

Goleman and Davidson (2017) offered a comprehensive explanation of five levels of meditation ranging from narrow but deep to wide and inclusive. They described the first level as an ancient intense form of meditation (e.g., practiced as a complete lifestyle in the tradition of Theravada Buddhist monks from Southeast Asia or yogis in Tibet) and the second level as still spiritual but not an entire lifestyle. They further clarified the third level as a wide secular approach, available to a much larger population and often learned by taking a course such as the 8-week MBSR program or through an experienced private teacher such as in TM. Goleman and Davidson (2017) explained that the fourth level included methods contained on recordings or apps, available to a larger audience because they do not require enrollment in a course or tutelage from an authorized teacher, but the authors believed such forms are inevitably diluted in order to make them easily accessible to most people. The authors also stated that they are developing a form of meditation, based on scientific research, that they envision will become a fifth level. Although Goleman and Davidson revealed that they practice meditation at the second level (i.e., within a spiritual framework), they affirmed that their own Asian teachers believed that

meditation should be available to all who seek relief from suffering and not just those who practice through religious methods.

In accordance with the belief that meditation should be accessible to all, Sedlmeier et al. (2012) explained that Eastern meditative practices have been adapted to our Western culture and are now taught with secular methods that are not associated with any culture or religious beliefs. Goya et al. (2014) clarified that Eastern meditative traditions were designed to be lifelong, personal, and introspective pursuits to gain spiritual insight, while Western meditation methods are less spiritual and more concerned with general well-being. Goya et al. are skeptical of the ability Western forms of mindfulness meditation may possess to reach an optimal level of effective skill within the limited time frame of 8 weeks investigated in most studies. In addition, the concern for recognizing and honoring the spiritual core, particularly of the Buddhist tradition, is shared by others. Li and Ramirez (2017) asserted that Western forms of mindfulness meditation have reduced traditional Zen practice to a method devoid of its original spiritual roots. They questioned the commercial use of the practice and the ethics of how it should best be integrated, to provide access to those who may benefit, without totally losing the principles and spirit that it was initially based on.

Originating as an Eastern form, TM was first introduced to the West in the early 1960s by the Maharishi University of Management, previously known as the Maharishi International University, and is based on classic Sanskrit mantras (Goleman & Davidson, 2017). In TM, the student is instructed to sit comfortably and silently repeat a mantra, a word or phrase bestowed by a teacher who claims that it was specifically selected for the student and must never be revealed, twice a day for 20 minutes (Benson & Klipper, 2000). Colbert (2013) stated that TM can be learned by anyone regardless of religious belief and is a basic technique designed to allow

the practitioner to experience thoughts and "transcend" into awareness but must be learned from instructors trained in the TM method (p. 497). Interestingly, Selye (1978) advocated relaxation into altered states and noted that Harvard University's Dr. Benson modified TM and called it "the relaxation response" (p. 421).

Originally published in 1975, in the updated version of *The Relaxation Response* (Benson & Klipper, 2000), Dr. Benson explained that his interest in stress arose from his work as a cardiologist at Boston City Hospital and expressed gratitude to Maharishi Mahesh Yogi for the development of TM as a simple, consistent technique that lent itself to scientific scrutiny and for allowing researchers to study the method. Benson and Klipper reported that the relaxation response was developed as the result of research that showed a pattern of decrease in blood pressure, heart rate, and respiration, the opposite of that which occurred during periods of high stress. They clarified that the relaxation response is a non-religious method, similar to meditation, and provided step-by-step instruction on the method that includes a quiet environment, focus on an object (a symbol, the breath, a word, or phrase), a passive attitude, and a comfortable posture that can be held for 20 minutes twice a day. Dr. Benson believed regular engagement in such a practice would be beneficial, have a protective effect on health, and mitigate damage from stress (Benson & Klipper, 2000). In addition, the relaxation response is taught at The Benson-Henry Institute for Mind Body Medicine at Massachusetts General Hospital ("Benson-Henry Institute," n.d.).

Mindfulness meditation was first introduced to Western culture by the Vietnamese Zen Buddhist monk Thich Nhat Hanh and has roots in a meditative practice called Vipassana (Bamber & Kraenzle Schneider, 2016) but was popularized in the West by Jon Kabat-Zinn, who earned a Ph.D. in molecular biology and has spent decades studying the mind/body connection

and mindfulness meditation with chronic pain and disorders related to stress (Kabat-Zinn, 2013). Kabat-Zinn (2013) explained his own role in mindfulness meditation and described that he adapted the original Buddhist technique in 1979 when he developed the 8-week MBSR course—taught at the Stress Reduction Clinic at the University of Massachusetts Memorial Medical Center in Worcester, Massachusetts—designed to educate patients in self-care techniques and help them cope with chronic illness. Kabat-Zinn stated he developed the MBSR curriculum in an effort to use non-religious techniques, while still maintaining the mind-body philosophy of Buddhist teachings, to help a wide range of patients relieve their suffering from pain and illness. In addition, if the clinic was successful, he intended to use it as a model for other medical facilities (Kabat-Zinn, 2006). While based on Zen philosophy, some researchers such as Li and Ramirez (2017) have noted that Kabat-Zinn totally removed any affiliation with religion or spirituality from MBSR and have asserted that it amounts to "a secularized and Westernized form of Buddhist practice" (p. 184).

The original MBSR program still exists at the University of Massachusetts Memorial Medical Center through 8-week workshops. Kabat-Zinn (2013) stated thousands of people have enrolled in over 700 programs based on MBSR curriculum in medical settings in the United States and other countries. In addition to sitting meditation, the author stated the clinic teaches and advocates other stress reduction techniques such as walking meditation, a body scan, yoga, and mindful eating. Kabat-Zinn is well-known in the Western world for his widely used definition of mindfulness as "the awareness that emerges through paying attention on purpose, in the present moment, and nonjudgmentally to the unfolding of experience moment by moment" (Kabat-Zinn, 2006, p. 145).

The concurrent development of both the relaxation response and mindfulness meditation for use in stress reduction by doctors in medical settings may be confusing. However, Kass (2017) carefully noted that both Benson and Kabat-Zinn are credited with the early introduction of meditation to deal with stress. In addition, the author explained that while Benson and Kabat-Zinn conducted similar research, they took different paths in their teaching methods. Kass (2017) clarified the two are not in opposition to each other, but Benson's methods required mental concentration whereas Kabat-Zinn's methods focused on mindful awareness.

In this literature review, Transcendental Meditation (TM), Mindfulness Based Stress Reduction (MBSR) meditation, the relaxation response, and other forms of secular meditation are grouped into a generic category of Mindfulness Meditation. Therefore, studies including all forms of mindfulness meditation used to mitigate stress are considered. In support of grouping similar forms of meditation into a single category of mindfulness meditation, a study conducted by Oman et al. (2008) employed two different meditation-based interventions. They used two forms of meditation that included attention and reflection in the practice, an adapted version of MBSR, and an intervention based on Easwaran's Eight-Point Program (EPP). As Oman et al. (2008) originally hypothesized, and subsequently discovered, MBSR and EPP had such similar reductions of perceived stress at posttest and follow-up that they combined the results. However, Bamber and Kraenzle Schneider (2016) reported it is important to note that the curriculum of MBSR is highly consistent, and this must be considered during any comparison of methods as mindfulness meditation interventions can differ throughout various studies. In addition, with regard to both TM and MBSR, researchers cautioned, "uncertainty remains about what these distinctions mean and the extent to which these distinctions actually influence psychosocial stress outcomes" (Goyal et al., 2014, p. 358).

Although mindfulness meditation forms may employ different methods, whose background and practice slightly vary, they all share a similar intention to educate participants and provide them with an approach to the practice of mindfulness. Central constructs of mindfulness include an attention to the present moment, awareness, non-reactivity, and non-judgement (Bamber & Kraenzle Schneider, 2016; Bostock et al., 2019; Shapiro et al., 2011) and are intended to produce an overall increase in positive psychological aspects such as improved capacity to cope, increase in well-being, compassion, wisdom, insight, emotional regulation, and a decrease in reactivity (Bostock et al., 2019; Crowley & Munk, 2017; Metz et al., 2013). In contrast, renowned neuroscientist Robert Sapolsky, who has written on the subject of stress (Sapolsky, 2004), pointed out that most studies did not randomize participants and relied on those who decided to meditate, which could affect results. He acknowledged that meditation appeared to improve health but also noted the similarity of meditation types and cautioned against "anyone who says that their special brand has been proven scientifically to be better for your health than the other" (p. 402).

In fact, researchers have reported a positive impact in the use of mindfulness meditation to lower stress, increase awareness and attention, and decrease reactivity and judgement (Bamber & Kraenzle Schneider, 2016; Bostock et al., 2019; Oman et al., 2008; Shapiro et al., 2011). Studies have shown the practice of mindfulness meditation may increase the ability to cope and regulate emotions (Bostock et al., 2019; Crowley & Munk, 2017; Metz et al., 2013). An increase in awareness, attention, and coping coupled with a decrease in emotional reactivity and judgement may be helpful to high school students. However, before research at the high school level is examined, studies conducted at colleges and universities are reviewed to derive information regarding stress experience and management with older students.

Mindfulness Meditation in School Settings

As previously mentioned, high levels of chronic stress can affect students both physically and psychologically (Bamber & Kraenzle Schneider, 2016). Due to the similar developmental level, and the limited research on high school students, college students are included in the following review. In addition, since older high school students and freshman college students are close in age, information from research with first-year college students is highlighted. Many studies that focused on meditation and stress employed quantitative methods, but mixed methods and qualitative studies are included to develop a more nuanced picture of any relationship between stress and meditation. Research with college and university students is reviewed first, followed by a review of research of students at the high school level.

College and University

Bamber and Kraenzle Schneider (2016) conducted a comprehensive narrative review that examined 57 studies on the effect of mindfulness meditation on stress and anxiety in college students. The authors combed several databases and other valid sources to uncover 176 studies, screened them for inclusion criteria of the use of psychometric instruments to measure stress or anxiety, and excluded those that did not include meditation as an intervention. As noted, only 12 studies cited a theoretical framework, and "the most common of these was the Transactional Model of Stress and Coping" (p. 6). The authors found that while most of these studies indicated a reduction in stress in college students, there were significant common flaws such as small samples, homogeneous populations, variation of dose (duration, frequency, and weeks of meditation), meditation method, and lack of information regarding the connection between state and trait mindfulness and/or stress and anxiety, as well as a lack of physiological measures of stress. They stated that, of the 34 studies that included mindfulness meditation to address self-

reported stress, 25 studies reported a significant decrease in stress. As a result of this comprehensive analysis, the authors concluded mindfulness meditation interventions seem promising but indicated more rigorous research is needed before educational institutions design large-scale interventions for stress reduction.

Many researchers employed quantitative methods to study the effect of various mindfulness meditation programs on stress in college populations with a Random Control Trial (RCT) design and used an intervention and wait list model. Dvořáková et al. (2017) conducted a pilot study in a large public university with a self-selected, ethnically diverse, and gender-balanced population of 109 freshman students that examined the effect of an 8-week mindfulness program (Learning to BREATHE), delivered by trained facilitators, with first-year college students. The authors reported that students found the program aided in stress reduction and stress management and concluded mindfulness interventions may promote an increase in well-being, decrease mental health issues, and be useful to include in prevention programs in higher education. Dvořáková et al.'s (2017) study was significant because it highlighted the success of early intervention in college students and sought interventions to reduce student stress through institutional programs and practice. In addition, the authors reported that most students indicated they felt equipped to use the stress reduction skills in future situations. However, the study used a small sample and relied on both self-motivation and self-report, and the authors believed future studies should include physical measures of stress.

As previously mentioned, Oman et al. (2008) conducted an RCT study with 47 college students with two similar (i.e., MBSR and EPP) mindfulness meditation programs. The authors measured the effect of the interventions on stress at pretest, 8-week posttest, and 8-week follow-up. Due to the predicted lack of statistical difference between MBSR and EPP, the authors

combined them into Meditation Management of Stress to enhance statistical power. They reported results that indicated the treatment group showed a marginally significant decrease in stress scores on the PSS-10 at posttest but admitted it did not reach statistical significance until follow-up data revealed a small but statistically significant maintenance of perceived stress reduction at the 2-month follow-up. Although the study was limited to a small, self-selected sample size, which consisted of a large percentage of white, first-year, Roman Catholic female undergraduates, the authors reported that students may benefit from future interventions using meditation to address stress and demonstrated that similar non-religious meditation interventions may be equally effective. In addition, echoing the concern of other researchers, the authors indicated a lack of physiological stress measurement data was a limitation.

Greeson et al. (2014) examined the effects of a 4-week mindfulness meditation intervention called *Koru* in an RCT study of 90 students with an average age of 25. The authors reported data analysis of the PSS-10 showed medium effect sizes on perceived stress. Age, gender, ethnicity, race, religion, level of education, and sleep habit data were provided. However, the study consisted of a homogeneous sample consisting of 60% female, 62% white, and 85% non-Hispanic, so the findings may be limited. In addition, the authors reported compliance was an issue as only 33% of the Koru group attended all four classes. Despite limitations, the study analyzed an evidence-based mindfulness program, employed the widely used PSS-10, experienced a retention rate of 82%, and targeted the specific developmental needs of this population by using a brief intervention designed to maintain motivation for practice. The authors recommend a similar study be conducted with a more diverse population.

Warnecke et al. (2011) conducted an RCT and studied the effect of a 30-minute voice-recorded guided mindfulness meditation, delivered via compact disc, on the stress levels of

medical students who were instructed to listen to the meditation every day for 8 weeks. At posttest, the authors reported a decrease in scores on the PSS-10 and a slight but significant decrease in scores on the section of the Depression, Anxiety and Stress Scale (DASS) that measured stress. Although the medical student population represented an older portion of the developmental level targeted by this literature review, the use of a compact disc recording as a standardized form of delivery for independent practice was a new design method and pertains to this study.

A mixed methods study conducted by Ramasubramanian (2017) explored a 14-week, one-college-credit freshman mindfulness course with a convenience sample of 18 intervention and 35 control students who completed all aspects of the study, on physical, mental, and emotional well-being co-taught by instructors with experience in mindfulness. The researcher used the PSS-10 and the Coping Inventory for Stressful Situations (CISS) as measurements of stress and coping; analyzed, coded, and sub-coded qualitative journals by themes; and provided specific narrative examples. The researcher reported that qualitative thematic analysis included "Fatigued and restless, Anxious and stressed, and Overwhelmed" prior to intervention while post-intervention themes included "Relaxed, Focused, Well-rested, Aware and alive, Joyful, Optimistic, and Grateful" (pp. 315–316). Ramasubramanian (2017) noted rare reports of positive feelings pre-intervention and negative feelings post-intervention and stated that analysis of the CISS showed statistically significant findings of everyday stress positively correlated with emotional coping, task-oriented coping positively correlated with both avoidance coping and emotional coping, and emotional coping positively correlated with avoidance coping. The author also reported that the PSS-10 showed a marginal reduction in stress for the treatment group and a slight increase in the control group. The strengths of this study included that both genders were

equally represented (race and ethnicity were not provided), and rich descriptive narrative qualitative data were gleaned, but the study may have been biased since participants were recruited from a credit course and were given extra credit for participation. Ramasubramanian (2017) was consistent with the recommendation of Bamber and Kraenzle Schneider (2016) and used a theoretical framework, the Transactional Model of Stress and Coping (Lazarus & Folkman, 1984) to examine the construct of stress. Ramasubramanian (2017) proposed that students who are mindful may find it easier to adapt to stress and particularly noted that while the quantitative data were minimally significant, the qualitative was promising, showed a need for stress management in emerging adults, and indicated more research on this population is needed. In addition, Ramasubramanian emphasized that the qualitative data provided a more holistic approach to the study by providing reflective and nuanced narratives that the pretest and posttest data were unable to convey. In agreement with this view, qualitative research is examined next.

Crowley and Munk (2017) conducted a qualitative study of 28 college students engaged in a variety of meditation practices twice a week for 15 weeks and hypothesized that meditation practice would decrease stress levels and promote student well-being. The authors used grounded theory to capture rich, descriptive narrative data, which produced 12 consensual themes that were consolidated to three major themes of mindfulness, psychological well-being, and compassion. The authors' study was limited by a population of predominantly white, female students, the possibility of researcher bias exists because the study was conducted by the instructor of the elective course, and no data were provided regarding specific meditation methods. Despite these concerns, the authors concluded meditation appears to promote emotional well-being in college students and submitted that future research should use both

qualitative and mixed methods to explore the benefits of meditation in groups of different age, gender, and social class.

This review of studies of mindfulness meditation interventions conducted with college students indicated benefits that included perceived stress reduction, increased emotional well-being, and improved relationships and showed promise in positively affecting grades (Crowley & Munk, 2017; Dvořáková et al., 2017; Greeson et al., 2014; Oman et al., 2008; Ramasubramanian, 2017; Warnecke et al., 2011). In addition to aiding in self-regulation and stress management, qualitative data gleaned from students who practiced mindfulness meditation indicated students slept better, felt happier, and felt more capable of addressing stressful situations (Ramasubramanian, 2017), while Crowley and Munk (2017) stated students reported heightened attention to events that increased "joy, peace, gratitude and self-acceptance" (pp. 95–96).

Specifically, many researchers (Bamber & Kraenzle Schneider, 2016; Dvořáková et al., 2017; Loi et al., 2008; Oman et al., 2008) indicated mindfulness meditation interventions should be targeted for the vulnerable population of first-year college students. In contrast to subsequent years, researchers reported that first-year students experienced increased levels of sustained and persistent stress (Loi et al., 2008; Ramasubramanian, 2017). In a wellness survey of 1,007 first-year U.S. college students, Loi et al. (2008) reported that students scored low on their ability to manage stress. Bamber and Kraenzle Schneider (2016) submitted that mindfulness meditation interventions should be employed with first-year students to help encourage and promote positive adjustments to the college or university setting. Researchers using mindfulness meditation practice as an intervention have indicated that mindfulness meditation assisted first-year students in their transition and adjustment to the college setting, aided in academic

struggles, facilitated and helped manage positive coping strategies, and mitigated the negative aspects of physical and psychological stress (Bamber & Kraenzle Schneider, 2016; Dvořáková et al., 2017; Oman et al., 2008; Ramasubramanian, 2017). Colbert (2013) reported on the difficulties with transition that college students experienced and believed that high school students are greatly in need of stress management intervention so that "college will be a viable post secondary option" (p. 495). In addition, Eva and Thayer (2017) noted that teenagers from disenfranchised districts are not always given the supports needed to help them navigate difficulties as they move toward independence, post-secondary education, or careers. Although there is minimal research with high school students, the following studies are examined.

High School

In a review of studies of meditation with adolescents in a school setting, Wisner et al. (2010) noted that research with adolescents is limited, but relevant studies indicated that mindfulness meditation increased student well-being, self-control, focus, and academic and psychosocial strengths and enhanced self-regulation abilities and coping strategies. The authors stated that studies showed decreased incidence of student behavior issues and the physical benefits of reduced heart rate and blood pressure. Based on these findings, Wisner et al. (2010) concluded that more studies are needed to investigate interventions and examine the benefits of mindfulness meditation programs with teenagers in schools. In addition, in a meta-analysis of school-based intervention programs, van Loon et al. (2020) noted that while there are programs designed to reduce anxiety and depression, there is a lack of programs that directly address stress. In addition, the authors asserted that stress management programs could benefit both students and schools by warding off future mental health conditions that may impact academic success.

Elder et al. (2011) conducted a study of a mindfulness meditation program in four U.S. high schools with 106 (68 volunteer meditators and 38 control: mean age 16.8) racially and ethnically diverse (Hispanic, African American, and American Indian) students. The researchers explained volunteers were taught meditation and practiced, alongside non-meditating students, during a supervised class designated for quiet activities (e.g., reading or homework) for 10–15 minutes twice a day (morning and afternoon) for 4 months. The researchers measured stress levels pre- and post-intervention on the Strengths and Difficulties Questionnaire (SDQ) and the Spielberger State-Trait Anxiety Inventory for Children and found the treatment group had a reduction in psychological stress, evidenced by a significant reduction in emotional symptoms on the SDQ, and reduction in trait anxiety with no differences between ethnic groups. In addition, the authors discovered a reduction in anxiety in the control group and believed it may have been due to the calm environment generated in the classroom during the study. Elder et al. (2011) concluded that meditation was effective in reducing stress with adolescents, had a positive effect on academic and emotional levels, and recommended subsequent studies use a longer time frame with a larger population in a single school since they believed meditation has the potential to increase the physical and mental health of high school students.

Similarly, Metz et al. (2013) emphasized that research on meditation with adolescents is limited and conducted a pilot mindfulness meditation study with 216 (129 treatment, 87 control) students in two high schools, using the same Learning to BREATHE mindfulness program studied at the college level by Dvořáková et al. (2017) but delivered in the classroom by a teacher trained in mindfulness. Metz et al. (2013) found that participants in the mindfulness meditation treatment group had statistically significant improvements in emotional regulation and reported a decrease in perceived stress, measured as a single rating from 1–10 where 1

indicated no stress and 10 indicated significant stress. Erbe and Lohrmann (2015) conducted a literature review on mindfulness meditation interventions and stress with adolescents in clinics and schools. The authors noted that the studies of high school students appeared encouraging but used small sample sizes and did not employ RCT designs and therefore emphasized that more research with the high school population is needed. In addition, Yeager, Dahl, and Dweck (2018) suggested that school-based interventions for teens are more effective if they demonstrate respect for the participant and do not threaten social status.

Foret et al. (2012) conducted a study of a 4-week relaxation response curriculum with 86 students, 44 11th grade students in the intervention group and 42 10th grade students in the control group, who volunteered to participate in lieu of attending physical education classes. In addition to eight classroom training sessions, students in the intervention group were asked to practice 5–10 minutes of guided meditation daily. Audio tracks of guided meditations were provided by a website monitored by the researchers. The researchers reported 41% of students accessed the website once and 41% did not log on at all. In the post-study survey, students stated that accessing meditations on a computer was not convenient or relaxing. The researchers recommended that future studies use a portable method for students to access meditations.

As with any review of research, these findings must be interpreted with caution. A rigorous review and meta-analysis conducted by Goyal et al. (2014) examined meditation interventions that consisted of both mindfulness- and mantra-based meditation programs in adult clinical populations. While the authors reported that too few mantra meditation studies met the strict criteria to report results on stress, the mindfulness-based interventions analyzed showed small improvements in stress reduction. However, they noted that most of the 47 studies consisted of 8-week interventions and believed that this brief intervention time may have been

too short for participants to achieve sufficient skill in meditation to experience a measurable change. The authors also identified that "dose" (i.e., length and duration of meditation) was not adequately described in most of the studies but stated that most studies implied that a longer time spent meditating may produce better outcomes.

To summarize, studies of mindfulness meditation interventions of those attending college showed benefits such as stress reduction, increase in well-being, and better relationships, as well as improvement in sleep levels and academic grades (Crowley & Munk, 2017; Dvořáková et al., 2017; Greeson et al., 2014; Oman et al., 2008; Ramasubramanian, 2017; Warnecke et al., 2011). In addition, studies of first-year college students reported similar results of increased physical and psychological well-being (Bamber & Kraenzle Schneider, 2016; Dvořáková et al., 2017; Loi et al., 2008; Oman et al., 2008). Research of mindfulness meditation with adolescents has shown promise in stress reduction and improvement in well-being, but studies have been limited in size, and researchers have indicated there is a dearth of research with adolescents (Elder et al., 2011; Metz et al., 2013; Wisner et al., 2010). While there is a gap in the extant literature, the evidence presented in studies with college students indicated mindfulness meditation may be useful as a proactive stress reduction intervention for high school students. However, as Colbert (2013) noted, it is often difficult to implement interventions and instruction into programs and practice that can easily reach each student. In addition, Wisner et al. (2010) reported that schools are searching for new methods to deliver instruction to meet social-emotional needs, and Foret et al. (2012) recommended that employing a portable method of meditation instruction may be beneficial for students. Considering these concerns regarding access and method, this literature review next examines research using smartphone apps as a possible innovative delivery method for meditation instruction.

Mindfulness Meditation and Smartphone App Technology

Smartphone use is on the rise, and the technology associated with it is having an impact on all wellness fields "with preventative health and clinical interventions leading the way" (Howells et al., 2016, p. 164). There is also a documented increase in the popularity of mobile health apps that promote emotional wellness and adaptive coping (Lau et al., 2020). With the advent of such technology, smartphone and mobile health apps may provide the ability to deliver widescale mindfulness meditation instruction to a large population without the cost of training or hiring trained teachers and the additional expense, time, and space associated with ongoing classroom instruction (Adams et al., 2018; Bostock et al., 2019; T. Miller et al., 2015).

Research on mindfulness meditation programs delivered via smartphone apps is limited (Bostock et al., 2019). However, two studies of the effect of mindfulness meditation on stress conducted with smartphone apps in adult populations reported a reduction in stress-related high blood pressure, similar to reductions shown in results from other meditation programs (Adams et al., 2018), and a decrease in work-related stress (Bostock et al., 2019). Additionally, Champion, Economides, and Chandler (2018) conducted a 30-day study of a mindfulness meditation app that encouraged 10–20 minutes of daily meditation with 74 adult participants (41 female, 33 male, mean age 39.4) who were randomized to treatment (n = 38) and waitlist (n = 36). The authors reported a reduction in PSS-10 scores, compared to waitlist, after 10 days with an additional decrease at day 30 and therefore concluded that 10 days of mindfulness meditation with an app can reduce stress enough to reach statistical significance.

Miller et al. (2015) assessed 219 college students on the feasibility of using smartphone apps to target health-related topics. The researchers revealed that students listed stress in their top three categories of need and indicated they specifically wanted an app to address stress

management. Although a mindfulness meditation intervention was not offered, the same study compared the effectiveness of a weight loss intervention via web, app, or written log and found the app group had the highest retention rate at 93%. In addition, Adams et al. (2018) reported participants indicated a strong preference for an app versus a face-to-face intervention. Eva and Thayer (2017) studied a mindfulness meditation program at a high school and reported that the staff recommended students continue to use the method on their own time and suggested making a smartphone app available so the students could engage in mindfulness meditation practice at home. The results of these studies indicate a gap in the research and a rationale for designing a study to investigate the effect of mindfulness meditation instruction delivered by a smartphone app on perceived stress in the high school population (Eva & Thayer, 2017; T. Miller et al., 2015). Researchers stated a mindfulness meditation app meets a need for a beneficial, low-cost, accessible, easy-to-learn, and portable stress intervention (Adams et al., 2018; Bostock et al., 2019; Champion et al., 2018; Howells et al., 2016).

Goleman and Davidson (2017) claimed they are supportive of the digital delivery of contemplative practice and agreed it can reach more people but are concerned about the lack of stringent scientific research and are cautious regarding the results of studies that mindfulness meditation apps report. The authors stated, “the apps typically cite studies done elsewhere on some kind of meditation (and not necessarily the best such studies), while failing to be transparent about their own effectiveness” (p. 283). The authors emphasized that research in meditation is inconsistent, and this fact may be overlooked particularly when research results are being used to further the progress of a certain type of meditation method or merchandise. However, they noted the effectiveness of a web-based intervention that consisted of twenty sessions lasting 10 minutes each. Goleman and Davidson suggested there may be evidence that

well-designed digital apps, which were found to be effective after rigorous scientific study, may be useful. In support of this view, Davidson is part of a team that is developing a digital product designed to teach meditation called “Healthy Minds” (p. 275). In fact, initial evidence from an 8-week RCT comparing meditation interventions delivered by a smartphone app as part of the Healthy Minds Program (HMP) to waitlist control (Goldberg, Imhoff-Smith, et al., 2020) indicated a small reduction in stress.

Voicing a similar concern for replacing face-to-face mindfulness meditation instruction with an app, Australian researchers Mani, Kavanagh, Hides, and Stoyanov (2015) reviewed 700 meditation-related smartphone apps to determine if apps could compare to in-person mindfulness instruction. They narrowed the field to 560 mindfulness meditation apps and found 23 met the inclusion criteria of a non-religious mindfulness meditation practice that provided instruction through guided mindfulness meditation training and could be purchased for less than $10. The authors of each study that Mani et al. reviewed had completed mindfulness meditation training, and two of the researchers used mindfulness meditation in their practice as clinical psychologists. Mani et al. employed the Mobile Application Rating Scale (MARS), which they helped to develop, to assess each app. Mani et al. reported the MARS has shown “excellent internal consistency ($\alpha = 0.92$) and interrater reliability” (p. 2). It was noted that mindfulness meditation training is a skill that requires commitment and regular practice to be effective. The authors submitted that pleasant, engaging apps that are available for use at all times may encourage consistent use. In addition, the researchers stated apps that can track personal data and connect users to a wider supportive network may also promote regular use. Results of the review showed four apps—Headspace, Mindfulise, Buddhify 2, and Smiling Mind—scored above acceptable levels on the engagement portion of the MARS instrument. In addition, the authors cautioned

there is a lack of evidence, and more research is needed to assess the effectiveness of mindfulness meditation apps. They recommended that future studies should use randomized controlled trials.

Although mindfulness meditation apps are just emerging as a delivery method in health-related prevention and intervention, studies have shown they can be useful and are low-cost, portable, easy to use with a large group of individuals, and available at any time (Adams et al., 2018; Bostock et al., 2019; Howells et al., 2016; T. Miller et al., 2015). Researchers have stated that mindfulness meditation apps have shown promise in reducing stress in adult populations (Adams et al., 2018; Bostock et al., 2019; Champion et al., 2018). A study of college students indicated they preferred an app delivery model for weight loss and would like to use an app for stress management (T. Miller et al., 2015). A survey of digital health practices of 1,137 youth aged 14–22 (Rideout et al., 2018) reported approximately 11% have used an app for mindfulness or meditation practice.

Participant Compliance With Apps

Weber, Lorenz, and Hemmings (2019) noted that non-adherence to digital interventions can be high, and app-based interventions may not be ideal for everyone. Several researchers have reported concerns regarding compliance with digital interventions including limited engagement (Gál et al., 2020) and lack of completion rate (Mrazek et al., 2019). In addition, Psihogios, Stiles-Shields, and Neary (2020) reported the average app is uninstalled in less than 9 days and cautioned that evaluating user engagement is critical if the app is planned for long-term use.

In a randomized control trial, Huberty et al. (2019) studied the effect of the commercial app Calm with 88 college students (mean age intervention 20.41, control 21.85). After familiarizing themselves with mindfulness meditation by listening to a 7-day introductory

program, the intervention group was asked to meditate for 10 minutes a day during the 8-week intervention. The authors reported the intervention group showed a significant reduction in scores on the PSS-10 after 8 weeks that persisted at the 12-week follow-up and acknowledged that participants received $5, $10, and $15 gift cards as incentives for completing the pre-, post-, and follow-up assessments respectively. Huberty et al. (2019) noted that adherence declined during the study but reported the intervention group averaged 38 minutes of meditation per week. Moreover, they stated that college students have many obstacles (e.g., a busy schedule, academic requirements, and relationships) that could interfere with daily meditation practice, so a lack of compliance was not surprising.

In a review of 70 studies of smartphone app interventions that targeted mental health problems, Linardon and Fuller-Tyszkiewicz (2020) reported that adherence rates were below par. The authors believed that researchers could increase adherence by offering monetary incentives, prompting participants with frequent reminders, and emphasizing the commitment required to complete the intervention. In a study of the meditation app Headspace On-The-Go, which consisted of 10 minutes of meditation a day for 10 days, Howells et al. (2016) reported higher than expected attrition rates during both recruitment and retention. The authors asserted that, compared to other study methods, app-based studies may result in increased attrition rates due to lack of in-person participation, distraction, lack of interest, and technical difficulties.

In a 10-day RCT with a convenience sample of 208 university students in New Zealand (M = 20.8), Flett, Hayne, Riordan, Thompson, and Conner (2019) randomized students to one of two treatment groups, using the mindfulness apps Headspace or Smiling Mind, or a control group with a notetaking app called Evernote. Students in the treatment groups were instructed to complete 10 minutes of their respective app's introductory program for 10 days, which contained

similar exercises such as mindful breathing and body scans. Students in the control group were asked to take 10 minutes each day to list all the activities they had completed during the day. The study relied on self-report of app use. After the 10-day study period, each group was given free access to the app for an additional 30 days. Flett, Hayne, et al. (2019) reported that while daily use during the 10-day period was high (Headspace 8.24, Smiling Mind 8.00, and Evernote 8.74), over half of the participants discontinued app use after 10 days. The researchers believed that self-report may have led to erroneous estimations of app use and asserted that such a large decrease in use is common. Moreover, further research (Flett, Fletcher, et al., 2019) reported that adherence to self-guided digital interventions is often low and recommended that studies should make use of the electronic record of sessions maintained by most apps instead of relying solely on self-report.

This section has shown that there is ample research with college students. Currently, however, there is a gap in research on the use of mindfulness meditation apps with high school students. While Eva and Thayer (2017) reported high school staff members believed a stress reduction app would be helpful for students to use at home, further exploration into the feasibility of this method is warranted.

Summary

This literature review examined the rise in stress levels of emerging adults and provided evidence of the detrimental effects of chronic stress. It examined a theoretical framework underpinned by research related to coping with perceived stress. In addition, it detailed the history and psychometric properties of the PSS-10 as an instrument used to measure perceived stress levels. Moreover, this review provided evidence that showed promise for the use of app-based meditation interventions in a variety of educational settings that serve emerging adults.

The rationale for a stress management intervention for high school juniors and seniors is supported by the evidence of increased mental health issues cited by studies in college, and particularly evident in college freshman, due to both the developmental concerns with the emerging adult stage and the struggles involved in the transition to college life (Bamber & Kraenzle Schneider, 2016; Bland et al., 2012; Byrd & McKinney, 2012; Crowley & Munk, 2017; Dvořáková et al., 2017; Greeson et al., 2014; T. Miller et al., 2015; Oman et al., 2008). The research reported from studies of mindfulness meditation programs with high school students, while limited, also demonstrated that mindfulness meditation interventions showed promise in reducing perceived stress levels (Elder et al., 2011; Metz et al., 2013; Wisner et al., 2010).

In summary, this review of the literature highlighted the need for a stable measure of perceived stress that is suitable for use with adolescents. It examined the psychometric properties of three versions of the Perceived Stress Scale (Cohen et al., 1983; Cohen & Williamson, 1988), provided studies that emphasized the argument for interpreting scores as a state measure, discovered research that questioned the assumption of the PSS as a state measure and asserted that it was a trait measure, compared studies that examined and explained the rationale for reporting results with a one- or two-factor structure, and focused on studies of test-retest reliability.

The emphasis on test-retest reliability to determine instrument stability is particularly appropriate for this current project conducted with high school students. Although McCrae et al. (2011) referenced test-retest in personality scales, they raised an important point when they questioned if variations were more prevalent in adolescents than in adults. This study provided an exceptional opportunity to assess test-retest reliability of the PSS-10 since research indicated it has rarely been measured with adolescents. As Cronbach (1951) observed, “In practice,

psychologists and educators have often not had the opportunity to recapture their subjects for a second test" (p. 297). This study was designed to address this gap by examining the stability of the PSS-10, as a valid measure of perceived stress, with a high school population, and the results may contribute to the extant literature.

In addition, this review highlighted the need for a developmentally appropriate, easy-to-learn, effective, low-cost, and scalable mindfulness meditation intervention that can be implemented with high school students. This literature review has indicated that a smartphone app may meet this need. Recent studies (Adams et al., 2018; Bostock et al., 2019; Mani et al., 2015) showed that well-designed apps, which provide secular mindfulness meditation education with guided mindfulness meditation training that can be used at any time, have been developed and are available. The goal of this study was to expand on previous research to examine the effect of mindfulness mediation methods with a high school population. Specifically, as indicated in this review, the use of a mindfulness meditation app intervention may increase participant motivation and make it easier to deliver on a large scale in school or for independent use. In addition, this study may have provided information to inform programs and practices for feasible inclusion of mindfulness meditation instruction to benefit all students. This review of the literature did not discover any study that investigated the feasibility of using a mindfulness meditation app with high school students. Along with the need to examine the PSS-10, this gap in the research justified a need for the feasibility component of the study.

CHAPTER 3: METHODS AND RESEARCH DESIGN

This chapter is divided into two sections to describe the methods used for a test-retest reliability study as well as outline the methods employed in the original study. While researching the Perceived Stress Scale (PSS-10) (Cohen et al., 1983; Cohen & Williamson, 1988), I discovered that very little was known about the test-retest reliability of the instrument, and a significant gap existed since no test-retest reliability studies could be found that had been conducted exclusively with adolescents. As noted in Chapter 1, student compliance with independent use of the app in the original study did not occur as planned. However, I had amassed six sets of PSS-10 data from the original study. These six data sets consisted of two measures, each taken a day apart, from pretests, midtests, and posttests. The midtests were administered 6 weeks after the pretests, and the posttests were administered 3 weeks after the midtests (i.e., 9 weeks after the pretest).

The PSS-10 data sets obtained from the adolescent participants enrolled in the original study proved to be a valuable source of archival data. This enabled me to redesign the research project to incorporate a thorough comparative review of the test-retest reliability of the PSS-10 and include a study of test-retest reliability with adolescents by using subsets of PSS-10 scores obtained from the original study but now treated as archival data. It is important to note that the original study design had a treatment and control group. However, the archival PSS-10 data used for the test-retest reliability study was limited to subsets from the original study. Since the treatment began after the pretest, one of these subsets includes pretest scores obtained from all students (i.e., treatment and control groups from the original study) to obtain the largest possible data set, but the remaining subsets are derived solely from the pretest, midtest, and posttest scores from the original study's control group. Therefore, no subject included in the new test-

retest study was administered or directed to use the treatment that the original study was designed to assess.

The test-retest reliability study consisted of three parts. First, I combed the literature for prior studies of PSS-10 test-retest reliability and conducted a comprehensive review of the results. Second, I examined the test-retest reliability of the instrument over time intervals using the archival PSS-10 data obtained from the adolescents in the original study. Third, I compared the test-retest reliability results I obtained from the archival PSS-10 data to the results from prior studies. In addition, I augmented the original study to include an additional research question and a survey that sought to identify the barriers and motivations students experienced with using the meditation app.

Five research questions were addressed. The test-retest reliability study focused on investigating the first two research questions. Research question 1 sought to explore previous findings of test-retest reliability of the PSS-10 with consideration given to language, time interval, and age. Research question 2 was divided into three sections that focused on the test-retest reliability of the PSS-10 using the archival data from the adolescents in the original study. Research question 2A sought to determine 24-hour test-retest reliability; 2B explored how test-retest reliability changed over 24-hour, 3-week, 6-week, and 9-week intervals; and 2C compared the test-retest reliability results from the archival data of high school participants to the results of comparable test-retest reliability intervals from prior studies.

The original study focused on answering the last three research questions. Research question 3 investigated if students from the treatment group of the original study adhered to the baseline of compliance by independently meditating four times a week during the 8-week intervention. As noted, the students were not compliant. Therefore, research question 4, which

would have compared the PSS-10 scores of the treatment group to the scores of the control group to determine if app-based meditation had influenced stress levels, could not be answered. However, noncompliance enabled me to develop research question 5, which sought to investigate compliance by examining the barriers and motivations that students experienced with app-based meditation.

Participants and Setting for the Studies

The subjects came from an urban public technical school in the Northeast with 11th and 12th grade high school students in health classes during the 2019–2020 school year. The school principal and classroom teacher consented to this study and allowed access to the students. The high school has approximately 2,200 students and receives students from 10 different middle schools from the local city as well as three of the surrounding suburbs, one charter school, and various private schools. From a social justice perspective, this school serves many students who belong to one or more underserved populations. The local Department of Education school profile report indicated that many of the 2,200 students fell into categories that include marginalized populations. For example, approximate percentages of specific categories include: 26% First Language not English, 8% English Language Learners, 19% Students with Disabilities, 63% High Needs, and 43% Economically Disadvantaged. The students are ethnically diverse, and the Department of Education report noted that the ethnic breakdown was approximately 4% African American, 16% Asian, 34% Hispanic, 1% Native American, 0% Native Hawaiian, 41% White, and 4% Multi-race, Non-Hispanic. The diverse population of study participants was an asset for this research.

Research Design and Procedure for the Test-Retest Reliability Study

Comparative Review of Prior Studies

Research question 1 sought to examine the test-retest reliability of the PSS-10 with consideration to language, time interval, and age. To answer this, I combed the literature for studies that included an analysis of the test-retest reliability of the PSS-10 and prepared a comprehensive review. Since the PSS-10 did not exist until 1988, I included the test-retest reliability data of the original 14-item Perceived Stress Scale from the instrument development study (Cohen et al., 1983; Cohen & Williamson, 1988). The literature contained several PSS-10 test-retest reliability studies conducted predominantly with adults (i.e., > 18 years of age) that used versions of the PSS-10 translated into languages other than English. I organized the studies according to retest interval. In addition, I indicated which translated version was used, the population studied, the mean age of the participants, and specified if results were reported for composite scores only or included two-factor scores (see Appendix A).

Eligibility Criteria

The test-retest study included archival PSS-10 data from students who were enrolled in the original study and had completed at least two PSS-10 pretests, midtests, or posttests (i.e., within a 24-hour retest interval). The six administrations are identified as two pretest scores labeled T1A and T1B, two midtest scores labeled T2A and T2B, and two posttest scores labeled T3A and T3B. The original study is explained later in this chapter.

Subjects

The pool of subjects used for the archival data in the test-retest study was derived from the original study, which included 101 subjects from a control group and 110 subjects from a treatment group. The archival data used for each subgroup are described in the procedure section below.

Instrument

The 10-item PSS-10 used in the test-retest reliability study is based on an ordinal five-point Likert scale. Researchers who have studied the validity of the instrument reported the PSS-10 has exemplary internal consistency across many studies (Creswell & Creswell, 2018) and is considered generally valid (Taylor, 2015). Participants were asked to respond to each question by choosing an answer from 0 (never) to 4 (very often) reflecting how frequently they have felt or thought a certain way within the past month. The PSS-10 composite scores were calculated by reversing scores on the four positively worded items, questions 4, 5, 7, and 8 (e.g., 0 = 4, 1 = 3, 2 = 2, 3 = 1, and 4 = 0) and then adding all 10 items. Cohen and Williamson (1988) state scores can range from 0 to 40 and higher scores indicate greater perceived stress (see Appendix B).

Internal Validity

In this study, internal validity of testing participants within the research design was addressed as the PSS-10 was used for each pretest, midtest, and posttest to eliminate concern regarding instrument validity. As noted, although the PSS-10 has been used in studies with adolescents, this study provided a rare opportunity to specifically examine the test-retest reliability of the instrument using archival data from the original study that was conducted in a non-clinical setting exclusively with adolescents from a high school population.

PSS-10 Archival Database

The archival database used for the test-retest reliability study was derived from the pretests, midtests, and posttests of the paper and pencil version of the PSS-10 given during the original study. Each set of tests were given within a 24-hour interval. In the archival data, they are identified as two pretest scores labeled T1A and T1B, two midtest scores labeled T2A and T2B, and two posttest scores labeled T3A and T3B. The 24-hour interval measures used in the test-retest reliability study were culled from these three PSS-10 administrations. In addition, each

of the pairs of scores were averaged. They are identified as T1AVG (T1A and T1B), T2AVG (T2A and B), and T3AVG (T3A and B). These averages became data points for the analysis of longer test-retest reliability intervals. The T1AVG to T2AVG became the data points used for a 6-week interval. The T2AVG to T3AVG became the data points used for a 3-week interval, and the T1AVG to T3AVG became the data points used for a 9-week interval.

To obtain these data during the original study, I used the following process. I distributed the PSS-10 to each student. I read the directions aloud and answered any questions. I directed the students to carefully reflect on each of the 10 items and circle their response based on how they have felt for the last month. I emphasized that they should take their time and requested that they turn the PSS-10 over on their desk once they were finished responding to all questions. I explained that I would wait until all students had finished before I collected the instruments. When all students had completed the scale, I collected them. I placed a test number (e.g., T1A) and subject number on each PSS-10 for identification during data entry.

This process was repeated with students for the T1B pretest, for T2A and B midtests, and T3A and B posttests. The midtests were administered 6 weeks after the pretests, and the posttests were administered 3 weeks after the midtests (i.e., 9 weeks after the pretest). I entered individual item response numbers (i.e., 0–4) for each test (i.e., T1A–T3B) according to subject number into a database. These data were sorted according to treatment and control group from the original study and became archival data. The test-retest reliability study used subsets derived from this archival database.

Procedure

The following procedure is for the test-retest study that used archival data derived from the original study. Archival data were used from students who completed the PSS-10 at pretest

(T1), midtest (T2), and posttest (T3). Each pretest, midtest, and posttest was given twice within a 24-hour interval. It is important to mention that the data sets are unequal due to missing archival data from student absences but reflect the largest data sets available for each interval.

Different sample sizes were obtained from archival PSS-10 data and were used for four 24-hour test-retest intervals as well as for each of the 6-week, 3-week, and 9-week test-retest intervals. The largest 24-hour data set was gleaned from the archival data of all students who completed T1A and B (n = 190). The remaining 24-hour data sets were taken from subsets of archival data, using both T1A and B (n = 86), T2A and T2B (n = 67), and T3A and T3B (n = 74).

Test-retest reliability analyses were conducted on T1AVG and T2AVG scores obtained from participants who retook the PSS-10 after a 6-week interval (n = 62) and T2AVG and T3AVG scores obtained from participants who retook the PSS-10 after a 3-week interval (n = 55). From a purely exploratory standpoint, an analysis was conducted on the T1AVG and T3AVG scores obtained from participants who retook the PSS-10 after 9 weeks (n = 49). In addition, a one-way repeated-measures analysis of variance (ANOVA) and Friedman test was conducted with a subset of archival data that included all six of the PSS-10 measures of T1A and T1B, T2A and T2B, and T3A and T3B (n = 49).

Data Analysis Plan for Test-Retest Reliability Study

All data were analyzed using IBM SPSS Statistics (Version 27) predictive analytics software. Data were compiled in both SPSS and Excel, as Excel offered more flexibility in terms of data management and the Excel format was easily imported to and read by SPSS. The test-retest reliability study consisted of two research questions. The first question examined the test-retest reliability of the PSS-10 with respect to prior studies. The second question examined the

high school students' test-retest reliability results over different retest time intervals and compared the high school results to results with similar intervals from prior studies. Results of all analyses are reported in Chapter 4.

Test-Retest Reliability in Prior Studies

Research question 1 sought to explore previous findings of test-retest reliability of the PSS-10. I compared, analyzed, and summarized the results from prior studies of the test-retest reliability of the PSS-10 with consideration to language, time interval, and age. Specifically, I examined and compared the results across the following factors: English and non-English versions; time intervals (i.e., 1 week, 2–3 weeks, and 1 month); and populations.

Test-Retest Reliability Analyses

Six administrations of the PSS-10 are identified in the archival data as two pretest scores labeled T1A and T1B, two midtest scores labeled T2A and T2B, and two posttest scores labeled T3A and T3B. An average of the two composite scores of the PSS-10 was calculated at T1, T2, and T3 for each participant in the study. Therefore, the T1AVG score is the average of T1A and T1B, the T2AVG score is the average of T2A and T2B, and the T3AVG score is the average of T3A and T3B. Three sets of average scores were available from the archival data.

24-hours. Since the intervention from the original study had not begun and to obtain the largest possible data set, one subset of 24-hour interval data from the archival data of all students (i.e., the treatment and control groups from the original study) who completed both T1A and T1B was analyzed. A separate analysis of T1A and T1B data from a subset of archival data from the control group in the original study was conducted to eliminate the possibility of effect from participants who knew they had been assigned to the treatment group in the original study. In addition, 24-hour data was analyzed using the T2A and T2B data and the T3A and T3B data.

Therefore, I had four 24-hour data sets, a large data set comprised of T1A and T1B PSS-10 scores from both the treatment and control groups in the original study, and three smaller data sets of scores comprised of the T1A and B, T2A and B, and T3A and B from the original study's control group only. I used these four data sets to answer research question 2A, which sought to determine the 24-hour test-retest reliability of the PSS-10. An intraclass correlation coefficient (ICC) was calculated for each of the four separate 24-hour intervals using a two-way mixed-effect model based on absolute agreement of a single measure for each analysis (Koo & Li, 2016; Matheson, 2019; Qin et al., 2019). Mean estimations along with 95% and 99% confidence intervals (CI) are reported for each ICC.

Intervals of 6 weeks, 3 weeks, and 9 weeks. To answer research question 2B and facilitate analysis of test-retest reliability for intervals longer than 24 hours, 6-week, 3-week, and 9-week intervals were compared. Although I had a large archival data set for one of the three 24-hour comparisons, it is important to note that the 6-week, 3-week, and 9-week interval correlations used smaller archival data sets that were derived exclusively from the original study's control group. It is also important to clarify that the midtest occurred 6 weeks after the pretest and the posttest occurred 3 weeks after midpoint testing (9 weeks after pretesting). The 6-week interval data set included participants who had completed four PSS-10 measures, T1A and T1B (T1AVG) and T2A and T2B (T2AVG). The 3-week interval data set included participants who had completed four PSS-10 measures, T2A and T2B (T2AVG) and T3A and T3B (T3AVG). The 9-week interval data set included participants who had completed four PSS-10 measures, T1A and T1B (T1AVG) and T3A and T3B (T3AVG). Intraclass correlation coefficient analyses were conducted on T1AVG and T2AVG scores obtained from participants who retook the PSS-10 after a 6-week interval and T2AVG and T3AVG scores obtained from

participants who retook the PSS-10 after a 3-week interval. Mean estimations along with 95% and 99% CIs are reported for each ICC.

Analysis of variance due to time. To determine if time during the school year had any effect on the correlations, a one-way repeated-measures analysis of variance (ANOVA) was conducted with a subset of archival data from participants who had completed all six of the PSS-10 measures (i.e., T1A and B, T2A and B, and T3A and B). Although using this data set reduced sample size, the within-subjects design provided superior statistical power. The ANOVA was conducted to determine whether there was a statistically significant difference in PSS-10 average scores over point in time during the school year. The T1AVG, T2AVG, and T3AVG were used as the dependent variable and time as the independent variable. Although ordinal data has been historically analyzed using non-parametric tests, it has become increasingly prevalent for researchers to treat Likert scales as interval data and employ parametric tests (Jamieson, 2004). However, to rule out any concern, I also ran a Friedman test, the nonparametric equivalent of the one-way repeated-measures ANOVA.

Comparison of High School Results to Results From Prior Studies

To answer research question 2C, three ICC analyses were conducted to examine and compare the test-retest reliability of the PSS-10 at three separate time intervals to similar time intervals from prior studies. An ICC was calculated for each interval using a two-way mixed-effect model based on absolute agreement of a single measure for each analysis (Koo & Li, 2016; Matheson, 2019; Qin et al., 2019). One ICC was calculated with the largest archival data set at the interval of 24 hours from T1A to T1B. Using smaller archival data sets, one ICC was calculated for the 6-week interval from T1AVG to T2AVG and one ICC for the 3-week interval between the T2AVG and the T3AVG. Mean estimations along with 95% and 99% CIs are

reported for each ICC. In addition, Pearson product-moment correlations were calculated for each of the intervals to allow for comparison of the archival data from the current high school study's results to the test-retest reliability results reported from researchers who calculated the Pearson coefficient. I compared and analyzed the results from the archival data from the high school study to results obtained from similar time intervals in prior studies (i.e., 24 hours/2 days, 3 weeks, and 6 weeks).

The Revised a Priori Plan for Family-Wise Alpha and Statistical Power

The current study that focused on the test-retest reliability of the PSS-10 utilized a series of inferential statistical procedures with implications to the family-wise error rate (FWER) and statistical power. G*power (Faul et al., 2009) does not allow for power analyses on ICCs. However, as far as I could find, I believe the ICCs had reasonable power given the fact that the 95% and 99% confidence intervals do not pass through zero. Reductions in alpha levels were calculated for the first set of analyses and one to two follow-up inferential runs that might seem prudent. If needed, an additional analysis of inference would utilize an exploratory approach that exceeded alpha correction thresholds as outlined in Rubin (2017). The results of these exploratory analyses would be considered with more caution as they were obtained after the alpha levels had been assigned.

As Rubin (2017) and others have argued, such exploratory research is important to continue to investigate and report as long as it is done in a transparent way that allows researchers the ability to read such results with more hesitation and increased need for replication, as such results were determined after a reasonable FWER was obtained. At this stage of planning, the data set was archival, so I knew the sample sizes and the number of analyses for the first set of ICCs and added one more future test to determine the alpha unit. Moreover, given

the strong association between the variables in these sets of analyses, a "1/2 Bonferroni" could be reasonably utilized, as a full Bonferroni would be an over-correction for the FWER.

Table 1 summarizes the tests that were included within the FWER-corrected system, and any others outside this list should be viewed in an exploratory approach as outlined in Rubin (2017). With large enough samples and *p* values mostly testing whether ICCs are statistically larger than zero, alpha thresholds for each test and for the full FWER were calculated by dividing alpha of .05 by the number of tests and multiplying the answer by 2. For the 24-hour ICCs, they were given half the alpha of the other tests since they had much larger sample sizes, and the remaining ICCs were allocated the standard divided alpha as described above. This resulted in an alpha unit threshold of .0125, with 24-hour ICCs set at X/2 (.00625) for the remaining tests of the set listed in the table. For the reader who believes a full Bonferroni is not overly conservative, the unit alpha threshold would be .00625 with .003125 to be used for the 24-hour ICCs and .00625 to be used for the remaining four analyses. As depicted in Table 1, and reported in Chapter 4, in the findings all *p* values for the ICCs were well below the individual alpha thresholds and the FWER regardless of whether one used a more conservative correction for family-wide alpha inflation.

The *p* value for the Friedman should be considered outside the FWER aggregate process because it was simply a check to see if there was any difference. Here, the full alpha of .05 made more sense and was treated as an exploratory method, but the expectation was for there not to be a difference. Had the Friedman shown anything close to a *p* value of .05, this would have opened an additional complex discussion of how and whether this fits into the FWER issue. For now, I treat it separately, as it was clearly not a significant finding and was analyzed in the exploratory phase of the study. This also brings up a complicated issue of predicting the null hypothesis,

which technically raises difficult issues. However, for the purpose of this study and this dissertation, in consultation with the sponsoring chair of the dissertation, we determined this explanation and caveat were appropriate for how the Friedman was used as a basic extra check.

Table 1

Family-Wise Error Rate, Summary of Tests

Interval	**Test**	**Sample**	**[a]Aggregated FWER** (E signifies scientific notation)
24-hour	T1A and T1B	*n* = 190	[b]6.3506E -78
	T1A and T1B	*n* = 86	[b]5.5163E -34
	T2A and T2B	*n* = 67	[b]4.03E -27
	T3A and T3B	*n* = 74	[b]4.9316E -34
		Subtotal:	[c]4.03E -27
6-week	T1AVG and T2AVG	*n* = 62	[b]1.4906E -10
3-week	T2AVG and T3AVG	*n* = 55	[b]1.0679E -7
9-week	T1AVG and T3AVG	*n* = 49	[b]1.7E -5
		Grand Total:	[c]1.711E -5
1 Extra test	Additional Test not expected to see a difference (expecting large *p* value)		

Note. [a]Alpha per test: .0125, 24-hour: .00625

[b]All *p* values were well below their alpha allocated thresholds and the FWER total was below the alpha of .05 for the entire first set of ICCs of the study.

[c]Well below .05 Alpha FWER level.

Research Design and Procedure for the Original Study

Eligibility Criteria

All current 11th or 12th grade high school students enrolled in the selected teacher's health class were eligible for the original study. I recruited participants in each health class and explained details of the study, requested participation, and answered questions. Students were informed that they needed a smartphone and headphones to access the intervention. Students who did not have headphones were provided with a set for use during the study. Students and their parents/guardians received assent/consent forms with information regarding details of the study (see Appendix C). I obtained written assent and parent/guardian consent from all students who participated in the study, and all assent/consent forms were collected and secured by me. Students who did not return consent forms did not participate in the study but were given an alternate task by the classroom instructor.

Internal Validity

With issues of history, attrition, and demand characteristics inherent to the design, internal validity was a concern. I made a concerted attempt to minimize such threats, so they would not affect my ability to analyze and draw conclusions from the results. This effort would have aided me in determining if the results were due to the mindfulness meditation intervention and not due to other factors. Regarding history, no event occurred in one class but not another. Each of the groups were kept separate within the classroom setting. However, diffusion of treatment may have occurred as members of the two groups could communicate in other classes or outside of school. I requested that the students not discuss the details of the project with other students. A concerted effort was made to minimize compensatory demoralization or rivalry by emphasizing the waitlist condition as vital to the study. In addition, the waitlist group was

informed of their status, and they were provided with the same experience after Group 1's study concluded.

Sampling Strategy

The high school students were randomly assigned to Group 1 or Group 2 according to their scheduled health class period. Participation in this study was appropriate and consistent with the content and structure of the health course. The technical school followed a year-long calendar that placed students into alternating weeks of academic classes and technical shop classes designated as either A week or X week. Due to the organization of the school schedule, students attended health class during their designated shop week. The teacher taught each of the six sections of health on A week and X week in the same classroom, thus controlling for the effect of the classroom instructor and environment on the results. Group 1 (treatment group) was selected from the first period class, and Group 2 (control group) was selected from the second period class. The remaining treatment and control group classes were selected by alternating the class periods assigned through the school day to maximize consistency of time of day for the groups from morning to afternoon and control for the effect of time of day (see Appendix D).

Sample Size

An a priori sample size was calculated using G*Power version 3.1.9.4 (Faul, Erdfelder, Buchner, & Lang, 2009). Preliminary calculations indicated the sample size would need to be at least 162 students, with 71 students in each group, and may include as many as 200 students with 100 in each group. For this study, there was a range of 10–27 student participants in each health class. Due to the school's alternate week schedule, three sections of A week students were assigned to Group 1 and three sections of A week students to Group 2; three sections of X week students were assigned to Group 1 and three sections of X week students to Group 2. Group 1

consisted of 61 student participants in A week and 51 students in X week for a total of 112 students. Group 2 consisted of 54 student participants in A week and 47 in X week for a total of 101 students. In all, 213 students were enrolled in the study. During the second week of the study two students from Group 1 withdrew to focus on classwork. This resulted in a total of 211 students in the study, 110 students in the treatment Group 1, and 101 students in the control Group 2, which provided sufficient statistical power for the planned analyses.

Perceived Stress Scale Administration

For the original study design, the 10-item Perceived Stress Scale (PSS-10) (Cohen et al., 1983; Cohen & Williamson, 1988) was used to assess pretest, midtest, and posttest measures of stress before, during, and after the intervention with a treatment group and a waitlist control group (see Appendix E). Each pretest, midtest, and posttest was given twice within a 24-hour interval. At each testing session, I visited classrooms and administered the paper and pencil version of the PSS-10. At pretest, I distributed the PSS-10 to each student in the study. I read the directions aloud and answered any questions. I directed the students to carefully reflect on each of the 10 items and circle their response based on how they have felt for the last month. I emphasized that they should take their time and requested that they turn the PSS-10 over on their desk once they were finished responding to all questions. I explained that I would wait until all students had finished before I collected the instruments. When all students had completed the scale, I collected them. I placed a test number (e.g., T1A) and subject number on each PSS-10 for identification during data entry.

This process was repeated with all students for the T1B pretest and was used for midtests and posttests. The midtests were administered 6 weeks after the pretests, and the posttests were administered 3-weeks after the midtests (i.e., 9 weeks after the pretest). At the conclusion of the

study, I entered individual item response numbers (i.e., 0–4) for each test (i.e., T1A–T3B) according to subject number into a database for possible future analysis.

Intervention

The intervention in this study was mindfulness meditation with the Stop, Breathe & Think (SBT) smartphone app (Stop, Breathe & Think, 2019). An email from the Stop, Breathe & Think research team granted me permission to use the app in this study (see Appendix F). The app was available for download to both Apple and Android devices, was designed for daily mindfulness and meditation practice, and allowed the user to track meditation progress (Stop, Breathe & Think, 2019).

Procedure

Week 1

Day 1. On Monday of the first treatment week of the study, the SBT app intervention was introduced. Students received instruction on the process of downloading the app to their devices. The SBT app included *Get Started,* an introductory set of 10 lessons designed to provide a foundation to mindfulness and use of the program (see Appendix G). Students completed the first two sessions of *Get Started* in class. The students were shown how to create a Mindfulness Session electronic self-report log on Google Classroom to keep a record of the date, time of day, name of the meditation session selected, and length in minutes of the session.

Day 2. On Tuesday, students used the app on their devices to complete the third and fourth sessions of *Get Started*. Each of the 2-day classroom meditation sessions were designed to familiarize students with the process and allow for troubleshooting or any concern that may arise regarding downloading and access to the app. I showed students how to link their SBT app to their Apple Health (iOS) app or download and link the Google Fit (Android) app. The Apple

Health/Google Fit apps recorded and maintained a history of individual student meditation progress that was accessed by the researcher at the end of each week to confirm the Google Classroom self-report data.

Day 3. I asked students who were at the appropriate point in the *Get Started* program to use the app to meditate independently on Wednesday to complete the fifth session. Several students experienced technical issues and needed extra time to catch up and successfully complete the first five introduction sessions.

Day 4. On Thursday, I visited classes and requested that students complete the sixth session. I also answered questions regarding the app, the study, and/or the process. In addition, I showed the students how to set a reminder notification on the app to facilitate independent use and showed them how to create a Health/Fit App Screenshot log on Google Classroom. Students were asked to finish the remaining four lessons in the *Get Started* series during the second treatment week of the study.

Day 5. On Friday, I visited each class to answer questions and remind students to finish the remaining four lessons in the *Get Started* series during the second treatment week of the study. The full study schedule was provided to the students (see Appendix E). I asked students to continue to meditate independently for at least 4 days a week for the remaining weeks of the study. The meditation sessions varied in length from 1 to 10 minutes. Although a 1-minute session was accepted as a meditation, I encouraged students to use the longer sessions. I explained that students could meditate as much as they wished (i.e., more than 4 days a week or more than once a day). Students were reminded that their participation in meditation also included the weeks that they were not in health class, so they were asked to meditate independently during both A week and X week for a total of 8 weeks of active treatment time.

Students were reminded to complete a Mindfulness Session electronic self-report log on Google Classroom for each meditation. In addition, each time a student completed a meditation, it was automatically recorded in the linked Apple Health or Google Fit app. I thanked students for their participation and emphasized that this study may help provide information about stress and stress reduction in teenagers.

Weeks 2–8

I sent weekly reminders to students on Google Classroom to encourage them to use the app to meditate and record their sessions. I also thanked students for their continued participation. At the end of each week, I visited classes in the treatment group to collect meditation data.

Meditation Data Collection

At the end of each A and X class week, I reviewed students' self-report Mindfulness Session Google Classroom logs and asked each student to confirm the information from their phone by sending me a screenshot of progress data from the Apple Health or Google Fit app to their Google Classroom Health/Fit App Screenshot log. Due to the structure of the alternating week schedule, I collected 2 weeks of data during each data collection visit.

Survey Administration

Anonymous exit surveys were distributed in January 2021 to all students in the treatment group (see Appendix H). Surveys were given to X week students on January 9, 2021, and to A week students on January 16, 2021. Before the survey was distributed, I explained to students that the survey was anonymous and that no names were to be placed on it. I emphasized that I sought and valued honest opinions of their experience with meditation, the app, and the study. I reviewed the survey questions and answered all inquiries about the survey and the collection

process. I provided a sealed cardboard box with a long narrow opening in the top to allow for anonymous collection. I explained that I would leave the room while they answered the survey questions, that the cardboard box would remain sealed during survey collection, and the classroom teacher would notify me when all students had completed the survey. I asked the students to wait until I left the room to begin the survey and, once they were finished, to fold it in half and place it in the cardboard box. After instructions were given and questions answered, I distributed a survey to each student in the study and left the room.

Survey Data Collection

The classroom teacher alerted me once all surveys had been placed in the sealed box, and I returned to remove the box. The box remained sealed as these steps were repeated with each class in the treatment group for A week. The same process was used for X week. After the surveys were collected each week, I unsealed the box and marked each survey to indicate the week it was collected from. This was done to help identify any findings that may have varied from A to X week. When surveys from both weeks were collected, I assigned a respondent number to each survey and transcribed verbatim responses to each question for use in qualitative data analysis.

Data Analysis Plan for the Original Study

The original study focused on the last three of the five research questions. The third question was exploratory and assessed student compliance with the app-based intervention. If compliance was achieved, the fourth question was designed to analyze and compare PSS-10 scores from the treatment and control groups. Because results from the third question showed that participants did not comply, the PSS-10 scores of the control group and the treatment group were not analyzed so the fourth research question could not be answered. However, a fifth

research question was developed that used qualitative data gleaned from the anonymous exit survey. The survey answers were analyzed to explore potential reasons as to why adherence was poor and examine students' motivations and barriers to using the app and. Results of the frequency analysis and findings from the survey are reported in Chapter 4.

Meditation Frequency Analysis

To answer research question 3 and decide if research question 4 could be answered, information of date, time of day, and length in time for each meditation session was obtained for each student from their individual Mindfulness Session Google Classroom self-report logs and compiled by participant number into a summary log for the treatment group. For each participant in the treatment group, I crosschecked and confirmed their Mindfulness Session Google Classroom self-report data with their screenshot record from Apple Health/Google Fit. To facilitate analysis, data that was present on the Apple Health/Google Fit screenshot record but missing from the Mindfulness Session Google Classroom self-report log was added to the summary log as needed for each participant. At the end of the study, I created a pivot table in Excel that indicated the frequency of number of meditation sessions and length in time for each meditation session for all students in the treatment group for use in data analysis.

Survey Analysis

To answer research question 5, which sought to explore various aspects of the students' experience with using the app, I assigned each survey a respondent number and created a verbatim transcript in Microsoft Word for each question from all completed surveys. Using a process described by Bree and Gallagher (2016), I imported the raw data from participant responses to each question into Microsoft Excel. The data set was analyzed via thematic analysis

using an inductive, semantic procedure informed by the principles and steps delineated by Braun and Clarke (2006, 2019).

Ethical Issues

The scope of the research was fully explained, and written assent and consent was obtained from students and parents/guardians. Due to the structure and scheduling of the health classes, a fair representation of the student population was obtained. To ensure that student information was kept anonymous, a participant number was assigned to each student. The participant number was used to identify the student's PSS-10 scores and self-report meditation information so meditation duration and frequency could be linked. During the study, any electronic documents obtained by the researcher were secured in password-protected files stored on an encrypted hard drive. All paper documents related to the study including informed consent forms, PSS-10 tests, and anonymous surveys were secured in a locked file cabinet. One hard copy of a master list of students' names linked to participant number was maintained by the researcher, and the information was kept confidential and secured in a locked file cabinet separate from all other study information. All electronic study information gathered by the researcher was maintained on a password-protected file stored on a different encrypted hard drive than the master list of names. The data set had only participant numbers and therefore could not be linked back to the students' names. Anonymous surveys had only respondent numbers. All paper documents will eventually be destroyed with a crosscut shredder.

This study posed no physical risks to participants. Psychological risks were minimal. During the classroom meditation sessions, no student appeared distressed, and no student required a referral for additional support. Potential benefits of this study included a reduction in stress. Decisions concerning the stress measurement instrument and the app intervention were

made with careful attention to issues regarding gender, sexual orientation, race, ethnicity, and ability. The Stop, Breathe & Think app was selected because it was secular and accessible to all students in the study. The decision to use a smartphone app in this study was specifically intended to allow students independence, flexibility, freedom, and control over the intervention. In addition, participation in this study may have educated students on mindfulness meditation as a stress reduction method, allowed them access to specific stress reduction instruction, and provided them with a strategy that otherwise may not have been available to them due to cost, travel, or the time required to access a similar program.

Summary

In summary, this chapter detailed the research design for the test-retest reliability study and the original study. It described the methods used to determine the sample and setting; administer the PSS-10 pretests, midtests, and posttests; conduct the intervention; collect independent meditation data; administer the anonymous survey; and analyze the data. The test-test-retest reliability study focused on results from prior PSS-10 studies, examined results of the archival PSS-10 data obtained from high school students over different retest time intervals, and compared the results from the archival data derived from the original high school study to those from prior studies. The original study focused on the practical feasibility of using the SBT app to deliver mindfulness meditation instruction to high school students for classroom and independent use and investigated student attitudes about the experience. The results from the test-retest reliability study and the findings from the original study are presented in Chapter 4.

CHAPTER 4: RESULTS

Quantitative and qualitative findings of this study are presented in two sections. In the first section, results from prior studies of test-retest reliability analyses of the Perceived Stress Scale (PSS-10) (Cohen et al., 1983; Cohen & Williamson, 1988) are compared, and results from the current high school study are presented followed by a comparison of the current high school study to comparable prior test-retest reliability studies. The second section presents results of the feasibility research question of adherence to baseline compliance of meditation and the impact on research question 4, followed by a qualitative analysis of general themes generated from student responses to the anonymous exit survey that focused on their views of the Stop, Breathe & Think (SBT) app, meditation in general, and the structure of the study. The qualitative analysis provides further insight into the feasibility and effectiveness of running an app-based mindfulness meditation intervention with high school students.

Quantitative Analysis

Test-Retest Reliability in Prior Studies

Research question 1 sought to explore previous findings of test-retest reliability of the PSS-10 with consideration to language, time interval, and age. I examined and compared the results from prior studies across the following factors: English and non-English versions; time intervals (i.e., 1 week, 2–3 weeks, and 1 month); and with adult and adolescent participants. The literature contained several PSS-10 test-retest reliability studies conducted with adults (i.e., > 18 years of age) that used versions of the PSS-10 translated into languages other than English (see Appendix A for a full list of test-retest reliability studies).

English

Cohen, Kamarck, and Mermelstein (1983) developed the original English version of the PSS-14 and reported test-retest reliability of r = .85 for 82 freshman college students who retook the PSS-14 within a 2-day interval and r = .55 for a smoking cessation group of 64 adults who retook the PSS-14 after a 6-week interval. A thorough search of the literature revealed that the only study that could be found that investigated the test-retest reliability of the PSS-10, conducted by Miller, Medvedev, Hwang and Singh (2020), did not provide separate coefficients for each specific time interval so individual interval comparisons could not be made.

Non-English

Test-retest reliability studies with unidimensional results from non-English versions of the PSS-10 were compared. As noted in Chapter 2, studies have reported results for each of the two factors and/or the composite (i.e., unidimensional) scores. The research community has been divided as to whether the PSS-10 is comprised of two factors, so only studies that included composite scores were used for this comparison. The studies were grouped by time interval as to when the retests occurred. There are three separate time interval categories, (i.e., 7 day/1 week, 2–3 weeks, and 4 weeks/1 month). As needed, time interval labels were provided with more than one unit to reflect how the intervals were reported in the studies.

One-week interval. Results from five studies that used a 7-day (1-week) interval reported results that ranged from r_s = .79 to an intraclass correlation coefficient (ICC) of .95. The study that reported a result of r_s = .79 was conducted with a group of female university students with a mean age of 22.5 years. The female university students were part of a larger study that included a group of pregnant and a group of postpartum women who retested at a 2–3 weeks, which is included in the next section. Four studies reported results > .80. Two of these four reported ICCs in the .80s (i.e., .86, and .81). The study that reported an ICC = .86 was conducted

with university teachers (mean age = 45.5). The participants in the study that reported an ICC = .81 were female nurses (mean age = 48.3). The remaining two studies reported ICCs in the 90s (i.e., .95 and .91). The ICC of .95 was reported from a study of adults with lupus who ranged in age from 18–66 with a median of 49. The ICC of .91 was conducted with adults whose mean age was 32.2.

Two-to-three-week intervals. Seven studies used intervals that spanned from 2–3 weeks. Two studies were conducted with a group of pregnant women (mean age = 28.4) and a group of postpartum women (mean age = 29.7) with an interval of 2–3 weeks. Interestingly, each group reported r_s = .63. Five of the seven studies used a 2-week interval and reported test-retest reliability coefficients ranging from r_s = .68 to ICC= .93. The first study with university students (mean age = 26.9) reported r = .77. The second study of adult policewomen (mean age = 21.1) showed r_s = .68. A third study of university students (mean age = 18.3) resulted in r = .70. In the fourth study, researchers working with adults (mean age = 38) at a pain clinic for chronic headaches reported an ICC = .93. The fifth study was with adults (mean age = 44.32) and showed r = .88. Finally, the only study that used a 3-week interval, conducted with university students (mean age = 20.9), reported an ICC = .82.

One-month interval. It is important to note that these appeared to be the only two studies available in the literature that used a 4-week test-retest interval. One study of male and female medical students (mean age = 20.84) showed an ICC = .83. Another study, with female participants whose median age was 68, reported r_s = .43.

Test-Retest Reliability in Current High School Study

To answer research question 2 with the current high school study, archival test-retest reliability was examined at three different time intervals: 24 hours, 6 weeks, and 3 weeks. As

mentioned in Chapter 3, both the treatment group and the control group from the original study completed the PSS-10 at pretest (T1), midtest (T2), and posttest (T3). The pretest, midtest, and posttest were given twice within a 24-hour interval. The six administrations are identified as two pretest scores labeled T1A and T1B, two midtest scores labeled T2A and T2B, and two posttest scores labeled T3A and T3B. The pretest was given before the intervention with the treatment group had started, the midtest was given 6 weeks later, and the posttest was given 3 weeks after the midtest (9 weeks after the pretest). It is important to mention that the data sets for each of the 24-hour comparisons are unequal due to missing data from student absences but reflect the largest data sets available for each interval. These three sets of PSS-10 scores from archival data at pretest (T1A and B), at midtest (T2A and B), and posttest (T3A and B) allowed me to analyze and compare three separate 24-hour test-retest reliability correlations. Furthermore, each pair (i.e., A and B) of pretests, midtests, and posttests were averaged and provided data points for analysis of two longer test-retest reliability intervals of 6 weeks (pretest to midtest) and 3 weeks (midtest to posttest).

Analysis of 24-hour intervals. To answer research question 2A and provide data for research question 2B, an analysis was conducted for each testing period (i.e., pretest, midtest, and posttest) with a separation of 24 hours. To provide clarity and ease for the reader to interpret the data set, the two pretests are designated T1A and T1B, the two midtests are T2A and T2B, and the two posttests are T3A and T3B. To examine 24-hour test-retest reliability, the largest data sets available from participants who completed both A and B of each pretest, midtest, and posttest were examined. Since the intervention had not begun prior to pretest, archival data from students in both the treatment and control group who completed both T1A and T1B pretests (n = 190) were analyzed. However, the analysis of the midtests (T2A and T2B) and posttests (T3A

and T3B) was limited to archival data from the original study's control group only. An analysis of the data available from participants who completed both T2A and T2B midtests (n = 67) and who completed both T3A and T3B posttests (n = 74) was conducted. In addition, a separate analysis of T1A and T1B data from the original study's control group only (n = 86) was conducted to eliminate the possibility of effect from the original study's treatment group. This result (n = 86) allowed a comparison to the result from the larger group (n = 190) and provided three separate 24-hour test-retest reliability correlation coefficients from the original study's control group only.

An ICC was calculated for each of the three separate 24-hour intervals using a two-way mixed-effect model based on absolute agreement of a single measure for each analysis (Koo & Li, 2016; Matheson, 2019; Qin et al., 2019). Mean estimations along with 95% and 99% confidence intervals (CI) are reported for each ICC. Results show that the ICC = .92, 95% CI [0.89, 0.94], 99% CI [0.88, 0.94] using both treatment and control data for the 24-hour interval at pretest between T1A and B (n = 190); the ICC = .91, 95% CI [0.86, 0.94], 99% CI [0.84, 0.95] for the control group only for the 24-hour interval at pretest between T1A and B (n = 86); the ICC = .91, 95% CI [0.85, 0.94], 99% CI [0.82, 0.95] for the control group only for the 24-hour interval at midtest between T2A and T2B (n = 67); and the ICC= .93, 95% CI [0.89, 0.96], 99% CI [0.88, 0.96] for the control group only for the 24-hour interval at posttest between T3A and T3B (n = 74).

Analysis of 6-week, 3-week, and 9-week intervals. To answer research question 2B and facilitate analysis of test-retest reliability for intervals longer than 24 hours, T1A and B were averaged and are designated as T1AVG, T2A and B were averaged and are referred to as T2AVG, and T3A and B were averaged and are designated as T3AVG. As noted, the 24-hour

archival test-retest data for all participants (i.e., the original study's control and treatment groups) who completed both pretests were compared because the intervention had not started yet, and it provided a more robust sample with excellent statistical power (n = 190). Although combining the two groups provided a large data set, it is important to note that the 6-week and 3-week interval correlations used data sets obtained from the original study's control group only—since expectations of the treatment group may have influenced scores. It is also important to clarify the interval timeline and emphasize that the 6-week interval occurred before the 3-week interval. The 6-week interval occurred between the pretest and posttest and preceded the 3-week interval, which occurred between the midtest and posttest. The 6-week interval data set included participants who had completed all four PSS-10 pretests and midtests, T1A and T1B (T1AVG) and T2A and T2B (T2AVG). The 3-week interval data set included participants who had completed all four PSS-10 midtests and posttests, T2A and T2B (T2AVG) and T3A and T3B (T3AVG). Since archival PSS-10 data for both tests (i.e., A and B) at pretest, midtest, and posttest is missing from some participants, the sample sizes are not equal but reflect the largest samples available to provide the greatest amount of statistical power.

Correlational analyses were conducted on T1AVG and T2AVG scores obtained from 62 participants who retook the PSS-10 after a 6-week interval and T2AVG and T3AVG scores obtained from 55 participants who retook the PSS-10 after a 3-week interval. Moreover, as previously explained and noted in Chapter 3 (see Appendix E for the full study schedule), it is important to clarify that midpoint testing occurred 6 weeks after pretesting and that post-testing occurred 3 weeks after midpoint testing (9 weeks after pretesting). Results showed the ICC = .69, 95% CI [0.54, 0.80], 99% CI [0.48, 0.83] for the 6-week interval (n = 62), and the ICC = .63, 95% CI [0.45, 0.76], 99% CI [0.37, 0.80] for the 3-week interval (n = 55). From a purely

exploratory standpoint and not used for comparison with prior studies, an analysis of the 9-week interval occurring between T1AVG and T3AVG resulted in ICC = .55, 95% CI, [0.32, 0.72], 99% CI [0.24, 0.76] and Pearson product-moment correlations using a two-tailed test indicated $r = .55(47)$, $p < .001$ for the 9-week interval ($n = 49$).

In summary, four ICC analyses were available to examine and compare the test-retest reliability of the PSS-10 at three distinct time intervals to similar time intervals in other studies: two coefficients at the pretest interval of 24 hours ($n = 190$ with the original study's treatment and control, and $n = 86$ for the original study's control only), one coefficient for the 6-week interval between pretest and midtest ($n = 62$), and one coefficient for the 3-week interval between midtest and posttest ($n = 55$). As noted, an ICC was calculated for each interval using a two-way mixed-effect model based on absolute agreement of a single measure for each analysis (Koo & Li, 2016; Matheson, 2019; Qin et al., 2019). Mean estimations along with 95% and 99% CIs are reported for each ICC. In addition, Pearson product-moment correlations were also calculated for each of the three intervals to allow for comparison of the current high school study's results to the results of test-retest reliability results reported by researchers who calculated the Pearson coefficient.

Results show that the ICC = .92, 95% CI [0.89, 0.94], 99% CI [0.88, 0.94] for the large data set for the 24-hour interval ($n = 190$); the ICC = .91, 95% CI [0.86, 0.94], 99% CI [0.84, 0.95] for the smaller data set for the 24-hour interval ($n = 86$); the ICC = 0.69, 95% CI [0.54, 0.80], 99% CI [0.48, 0.83] for the 6-week interval ($n = 62$); and the ICC = 0.63, 95% CI [0.45, 0.76], 99% CI [0.37, 0.80] for the 3-week interval ($n = 55$). In addition, Pearson product-moment correlations using a two-tailed test were calculated and indicated $r = .92(188)$, $p < .001$ for the large data set for the 24-hour interval ($n = 190$); $r = .91(84)$, $p < .001$ for the smaller data set for

the 24-hour interval (n = 86); r = .69(60), p < .001 for the 6-week interval (n = 62); and r = .63(53), p < .001 for the 3-week interval (n = 55). Although test-retest reliability correlation data are often presented according to an increasing length of interval time, the correlations are presented chronologically according to the sequence in which the testing occurred (i.e., 24 hours, 6 weeks, and 3 weeks) and the correlation from the large 24-hour data set is used (see Appendix I).

Analysis of variance due to time. In order to determine if time during the school year had any effect on the correlations, a one-way repeated-measures analysis of variance (ANOVA) was conducted with a subset (n = 49) who completed all six of the PSS-10 measures (i.e., two pretests, two midtests, and two posttests). Although using this data set reduced sample size, the within-subjects design provided superior statistical power. The ANOVA was conducted to determine whether there was a statistically significant difference in PSS-10 average scores from pretest, midtest, and posttest. The PSS-10's T1AVG, T2AVG, and T3AVG were used as the dependent variable and time as the independent variable.

Results of the ANOVA conducted with the subset (n = 49) showed there were no outliers, and the data were normally distributed at each time point, as assessed by boxplot and Shapiro-Wilk test (p = .089, .219, .978; thus p > .05). The assumption of sphericity was met, as assessed by Mauchly's test of sphericity, $\chi^2(2) = 3.36$, p = .186. There were no statistically significant differences in PSS-10 average scores over time, $F(2, 96) = 23.78$, p = .284, with PSS-10 averages decreasing slightly from pretest (M = 19.04, SD = 7.05) to midtest (M = 17.97, SD = 7.09) to posttest (M = 17.74, SD = 6.90). Because the ANOVA indicated there was no statistically significant difference, post-hoc tests were not conducted.

As previously described in Chapter 2, the PSS-10 is a Likert scale that uses ordinal data. Historically, ordinal data has been analyzed using non-parametric tests, but it has become increasingly prevalent for researchers to treat Likert scales as interval data and employ parametric tests (Jamieson, 2004). While controversy has reigned regarding parametric versus non-parametric analysis of Likert data, Mircioiu and Atkinson (2017) argued that parametric analyses are more robust, and either type of analysis will lead to similar results. In order to rule out any concern regarding the use of parametric analyses with a Likert scale in this study, a nonparametric test was also employed. The nonparametric equivalent of the one-way repeated-measures ANOVA is the Friedman test. A Friedman test was run to determine if there were differences in PSS-10 average scores over time. Averages for the PSS-10 fluctuated from pretest (*Mdn* = 20.00) to midtest (*Mdn* = 16.00) to posttest (*Mdn* = 18.00), but the differences were not statistically significant, χ^2 (2) = 1.56, p = .46. Since the Friedman test indicated there was no statistically significant difference, post-hoc tests were not conducted.

Comparison to other studies. To answer research question 2C, the current high school study was compared to prior research with similar intervals of test-retest reliability. Prior studies were examined and compared to the current study's results (i.e., 2 days, 6 weeks, and 3 weeks). These comparisons are presented in Table 2. As previously mentioned in Chapter 2, Cohen et al. (1983) examined test-retest reliability during instrument development and confirmed a test-retest correlation of .85 for 82 college students who retook the PSS-14 after a 2-day interval and reported a lower test-retest correlation of .55 for 64 smoking-cessation participants who retook the PSS-14 after a 6-week interval. As seen in Table 2, the PSS-14 and PSS-10 are referenced as well as information from the studies such as language, population, age, interval, coefficient and factor.

Table 2

Test-Retest Reliability Studies

Reference	PSS version	Language/ Country	Population/ Age	Interval[a]	Coefficient[b]	Factor(s)
(Cohen et al., 1983)	PSS-14	English/ USA	Two college student groups Mean = 19.01 Mean = 20.75	2 days	r = .85	One
			Adults (smoking cessation group) Mean = 38.4	6 weeks	r = .55	One
Current Study	PSS-10	English/ USA	High school students Mean = 16.77	24 hours	ICC = .92 r = .92	One
				6 weeks	ICC=.69 r = .69	One
(Al-Dubai et al., 2012)	PSS-10	Malay/ Malaysia	Bachelor of Medical Science students Mean = 20.9	3 weeks	ICC = .82	One
Current Study	PSS-10	English/ USA	High school students Mean = 16.77	3 weeks	ICC = .63 *r* = .63	One

Note. [a]As indicated in study

[b]*r* = Pearson product-moment correlation

ICC = Intraclass correlation coefficient

Results of the quantitative analyses from the current study suggest that the test-retest reliability of the PSS-10, with a sample of adolescent participants in a high school setting, is high

for a 24-hour interval (ICC = .92) and drops sharply at 6-weeks (ICC = .69) and 3-weeks (ICC = .63). These results, of a decrease in the test-retest reliability coefficients of the current study's 24-hour interval (ICC = .92) compared to a 6-week interval (ICC = .69), are consistent with those of Cohen et al. (1983), who investigated test-retest reliability at a 2-day interval (r = .85) and a 6-week interval (r = .55). In contrast, the current study's 3-week interval resulted in an ICC = .63, which is a substantially lower correlation coefficient than was reported by Al-Dubai et al. (2012), who showed an ICC = .82 for an interval of 3 weeks.

As noted in Chapter 2, Al-Dubai et al. (2012) reported an ICC of .82 for a retest of the PSS-10, after a 3-week interval, with 74 students enrolled in a Bachelor of Medical Science program in Malaysia. The current study found a much lower ICC of .63 for 55 students who retook the PSS-10 within the 3-week interval that occurred between midtest and posttest. Due to study methodology, and in contrast to Al-Dubai et al. (2012), who administered two PSS-10 tests within 3 weeks, the 55 participants included in the 3-week interval of the current study took the PSS-10 six times, and the 3-week correlation is based on the T2AVG (third and fourth administrations) and T3AVG (fifth and sixth administrations), which may have affected the results (see Appendix A for a full report of all test-retest reliability studies reviewed). As noted, an analysis of the 9-week interval occurring between T1AVG and T3AVG resulted in ICC = .55, 95% CI, [0.32, 0.72], 99% [0.24, 0.76] and Pearson product-moment correlations using a two-tailed test indicated $r = .55(47)$, $p < .001$, (n = 49) but was not used for comparison.

As previously mentioned, the 3-week interval occurred after the 6-week interval and may have contributed to a lower correlation for the 3-week interval. Additionally, results of both the one-way repeated measures ANOVA and the Friedman test indicated that, while PSS-10 average scores declined and fluctuated slightly over time, the result was not statistically significant.

Although the ANOVA and Friedman test indicated that the point in time, during the fall semester, at which the participants took the PSS-10 did not affect student scores, it also raises questions as to why the 3-week interval was lower than the 6-week interval.

Adherence

To answer research question 3, an analysis of data from each student in the treatment group who created a self-report electronic meditation log (Group 1) (N = 104) was crosschecked with Apple Health or Google Fit data from each student's smartphone. As discussed in Chapter 3, students were requested to complete four sessions per week for 8 weeks. However, completing at least three sessions per week was considered as compliant. A frequency analysis of all data showed that one student, out of 104, met the baseline compliance of three meditation sessions for each of the 8 weeks with a total of 24 sessions (see Appendix J).

Although the lack of adherence was unfortunate, it allowed me to pursue a robust study of the PSS-10. While combing the literature to answer research question 1, I discovered that the instrument had not been thoroughly vetted with teenagers. Fortunately, I had access to a full archival data set of six PSS-10 composite scores from adolescents. This provided me with a rare opportunity to explore the test-retest reliability of the instrument with archival data from the current study's high school student population and augment research question 2 into research questions 2A, 2B, and 2C. The PSS-10 data was examined to investigate the test-retest reliability of the instrument over different time intervals. In addition, I compared the data from the adolescent students in the current high school study to the results of outcomes reported in prior studies conducted by researchers who tested similar time intervals with adults. Moreover, the lack of adherence meant that research question 4, designed to explore the effect of the meditation intervention on stress levels as measured by PSS-10 scores, could not be addressed. Therefore,

research question 5 was developed to examine barriers and motivations to compliance. The findings from research question 5 are reported in the qualitative analysis section of this chapter.

Qualitative Analysis

As noted in Chapter 3, an anonymous exit survey (see Appendix H) was developed to explore the feasibility aspect of the research goal, gain insight into the lived experience of the participants, examine potential reasons why adherence was poor, and derive possible strategies to increase student motivation to use the app. Specifically, the questions were designed to glean likes and dislikes of meditation, the app, their personal experience, and request student recommendations. Again, this information was sought to answer research question 5 to gain information about the feasibility of introducing a meditation app in the classroom and investigate student experience and feelings about independent use. Furthermore, I hoped the anonymous design of the survey allowed students to feel comfortable enough to honestly express their views and provide empirical information that might help future researchers.

This methodological decision is consistent with a *pragmatic* worldview, and Patton (2002) asserted that a pragmatic stance gives the researcher flexibility to determine *methodological appropriateness* (p. 72), which allows situational adjustments to be made to a study. Creswell and Creswell (2018) stated that a pragmatic worldview is not limited to a specific research philosophy and allows the researcher the ability to choose methods of data collection and analysis to best address the research question. Moreover, this decision to employ a *pragmatic* strategy was underpinned by the research of Levitt, Motulsky, Wertz, Morrow, and Ponterotto (2017) in an effort to facilitate a "meaningful contribution in relation to the study goals" (p. 16) and increase the integrity of the study. Levitt et al. (2017) proposed integrity as composed of fidelity and utility. In this instance, fidelity was construed as an attempt to obtain

the students' thoughts, feelings, perceptions, and experiences during this study; and utility was interpreted as an attempt to assess the feasibility, through student report, of introducing an app-based stress reduction method in the classroom for independent use by students. Furthermore, as advocated by Braun and Clarke (2019), adding this qualitative survey to the study provided me with a fortunate opportunity to experiment with my research method and conduct an analysis to better understand the experiences and attitudes of the participants.

Positionality and Reflexivity

Qualitative analysis is inherently personal, and the researcher must account for their own personal views (Wertz, 2011). Personally, I value meditation as a method of stress reduction. I believe it has contributed to my own emotional regulation and think it has the potential to help others. I assumed that students would be interested in meditation and would particularly enjoy an app-based approach since it allowed them freedom and control over their practice. As a former high school teacher and counselor, I believe that it is important to inform students that stress can be harmful and can have a significant negative effect on their physical and mental health. I believe that the school curriculum should provide strategies to increase student well-being. I also wanted to introduce students to a stress reduction method that does not require specialized equipment, is easy to learn, and could be done almost anywhere whenever it was convenient for them. Finally, heeding the advice of Braun and Clarke (2021a), I am fully aware that I may be biased and have worked to remain neutral in analysis, but I acknowledge that I am analyzing this data set from the perspective of a veteran teacher and counselor.

Data Analysis

Analysis of the initial survey prompt, which asked students to indicate how many times they had used the app outside of health class (i.e., independently), showed that 35 (36%) students

circled A. Never, 37 (38%) students circled B. Once or twice, 26 (26%) students circled C. A few times, and 0 students circled D. Many times. As previously mentioned, the questions in the survey were designed to ascertain student views of meditation and the app. Although 36% of the respondents stated that they never used the app independently, 64% used the app independently, and 100% of the students used it to meditate in the classroom during the introduction period. Therefore, all the respondents had some experience with using the app. The data set provided valuable insights from all students regarding their experience of meditation in class and perceptions of the app as well as from two-thirds of the students who used the app independently. The responses detailed student-perceived barriers and motivations to use, likes, dislikes, and recommendations for improving the experience.

The data set was analyzed via thematic analysis using an inductive, semantic procedure informed by the principles and steps delineated by Braun and Clarke (2006, 2019). However, as recommended by Braun and Clarke (2019), I did not approach the analysis solely as a rigid step-by-step process but rather adopted a reflexive, recursive stance and spent months poring over the data set looking for patterns and commonalities that best explained what the participants were trying to say and derive meaning from the responses. Specifically, I used *reflexive* thematic analysis (V. Braun & Clarke, 2021a), a version of thematic analysis that values subjectivity and advises the researcher to assume an "active role in coding and theme generation" (p. 6). Moreover, this approach also encourages researchers to move beyond surface interpretation and delve further into the data to generate hidden (latent) meaning (V. Braun & Clarke, 2021b).

This reflexive approach was an iterative, fluid, and nonlinear process. First, I assigned each survey a number and created a verbatim transcript in Microsoft Word for each question from all completed surveys ($N = 98$). This initial process allowed me to become thoroughly

immersed in the data and familiar with each response. Second, to identify preliminary codes, I color-coded participant responses in Microsoft Excel using a process described by Bree and Gallagher (2016). Based on the authors' instructions, I imported all the raw data of participant responses into Excel, sorted responses according to each question, and assigned colors as each code was generated. If more than one code was identified, the response was copied and placed with the appropriate category and assigned the respective color(s) that applied to the response. Third, responses were iteratively reviewed; similar codes were combined, and codes were revised to eliminate redundancy. Fourth, tentative themes were generated, and themes were broken into sub-themes or combined under a common theme when necessary or possible. Fifth, the data set was reviewed to ensure the themes and sub-themes correctly represented the data. Finally, I developed a table of themes and sub-themes with illustrative participant responses extracted from the data set (see Appendix K).

Heeding the advice of Braun, Clarke, Boulton, Davey, and McEvoy (2020), I searched for patterns to generate themes across the entire data set and resisted the temptation to simply summarize responses to each of the nine survey questions. In addition, as emphasized by Braun and Clarke (2021a), who cautioned researchers to consider their own position relative to both subject matter and participants, I was aware that I was viewing student responses through my eyes and it was impossible to completely separate my view from the analysis. In addition, as stated by Varpio, Ajjawi, Monrouxe, O'Brien, and Rees (2017), while reading responses required active interpretation, analysis was significantly more interpretive. During this process, I continually returned to my exploratory question to ensure that I was examining and interpreting the data set through a feasibility lens. Initially, I attempted to uncover barriers and motivations students reported that they experienced to use the app. Furthermore, I sought information that

would help me glean strategies that a school could use to introduce and support an app-based intervention and encourage future student use.

Initial Analysis of Use and Codes to Identify Motivations and Barriers

Analysis of the first survey prompt, which asked students to indicate how many times they had used the app independently, showed that 35 (36%) students circled A. Never, 37 (38%) students circled B. Once or twice, 26 (26%) students circled C. A few times, and 0 students circled D. Many times. As previously mentioned, the questions in the survey were designed to ascertain student views of meditation and the app, and it is important to note that 100% of the students used it to meditate in the classroom during the introduction period. While information from the self-report anonymous survey showed that 36% of the respondents stated they never used the app independently, electronic data collected directly from the app was in sharp contrast and revealed a much lower number that indicated 13.5% of students never used the app independently. In addition, electronic data showed that 53.8% used it 1–6 times and 32.7% used the app > 6 times during the study (see Appendix J).

Analysis of responses to Questions 1–4 revealed that students often listed more than one reason and generated different codes. As previously mentioned, if more than one code was identified, the response was copied and placed with the appropriate category and assigned the respective color(s) that applied to the response. To provide specific description, data extracts from the anonymous survey responses are identified by respondent numbers in brackets.

Barriers

Analysis of the 72 responses to the first survey question—"If you DID NOT use the app outside of Health class, please list a few reasons why"—generated codes that revealed barriers to use. Students reported that they did not use the meditation app because they "were too busy or

had no time" (*n* = 38), "forgot" (*n* = 23) "didn't need, want, or like it" (*n* = 20), or felt that "it didn't work" (*n* = 12). In addition, the 72 responses to the follow-up question—"If you DID NOT use the app outside of Health class, what would have helped you use it?"—indicated that many students believed they would use the app if they had "more time or better time management" (*n* = 16), "notifications" (*n* = 13), "if I needed help" (*n* = 11), if the app had "more variety, (e.g., games, activities, or different voices)" (*n* = 11), or if they were required to use it because it was "mandatory or assigned" (*n* = 8). Students also indicated that "nothing would motivate use" (*n* = 11), that they may have used the app if they had "access to the premium version" (*n* = 2), "if I felt that it worked" (*n* = 2), or if "it was easier to use" (*n* = 2).

Motivations

Analysis of the 26 responses to Question 3—"If you DID use the app outside of Health class, please list a few reasons why"— generated codes that students were motivated to use the app to "relax, calm down, and pause to think" (*n* = 14), "aid in sleep" (*n* = 8), and "relieve stress, depression, or anxiety" (*n* = 8) or due to "boredom" (*n* = 4) or "curiosity" (*n* = 4). The 26 responses to Question 4—"If you DID use the app outside of Health class, how did you feel after a meditation session?"—revealed that most students reported they felt "calm and/or relaxed" (*n* = 22), while some students also felt "in control or focused" (*n* = 6) and "less stressed" (*n* = 2). Also, two students reported fluctuating changes after sessions. For example, "After each session, I would feel even slightly calmer [sic] or I would feel indifferent. It really depended on specific sessions" [90] or "Sometimes I would be in a good mood after doing a session, and sometimes I would be still in a bad mood after a session" [98]. Finally, two students reported "no change" (*n* = 2).

Qualitative Findings

During analysis of the full data set of 98 surveys (anonymous respondent numbers are indicated in brackets), three main themes of *resistance*, *hesitance*, and *acceptance* were generated, and *recommendations* were examined (see Appendix K for the complete analysis of themes, sub-themes, codes, and relevant data extracts). These themes were constructed to address the research question and provide insight that may help future researchers. Specifically, I attempted to look beyond simply compiling students' perceived barriers and motivations. I sought to understand the experience from a student viewpoint to assess feasibility since identifying the main themes might allow me to derive strategies to address barriers, build on existing motivations, and increase student interest for app-based meditation.

In general, a few students were resistant or averse to meditation, and some students were accepting and receptive to meditation. However, most students appeared to fall within the theme of hesitant or reluctant since their responses to each of the questions in the same survey varied (i.e., some responses indicated resistance and some responses indicated acceptance). For example, the response regarding independent meditation from respondent [34] stated, "I didn't feel like I needed it. I was fine without it," but the response to meditation in class stated, "I liked the breathing techniques & setting imagery." Additionally, respondent [58] stated, "I only use it during class and I don't need the app" but also indicated that during class "The meditation app helped me feel calm and relaxed." Furthermore, Question 5—"What did you like and/or dislike about practicing meditation with the app during Health class?"—allowed participants to include both likes and dislikes in a single response. Examples of respondent answers to that question included, "I like how easy it was to use and how it recommended different sessions based on your emotions, but I hated how repetitive they were. All the sessions sound alike and they sometimes recommended the same ones multiple times" [23], "Liked having some quiet time to

myself. Dislike having to make a log for every session" [29], and "I've never practiced meditation before so it was cool to try but it was also hard for me to stay focused on what the speaker was saying to me" [89].

Resistance

Out of 98 surveys, four students maintained a theme of resistance in their response to each question. The overarching theme of resistance included sub-themes that they did not value or like independent use, would not attempt independent use, and did not enjoy meditation with the app in the classroom. For example, respondent [13] stated, "I didn't use the app because it wasn't useful for me outside of school. Wasn't helpful," explained "Nothing would have helped me [use it]," added "I like how we were able to go on our phones," and indicated that if the app was allowed in school, "I wouldn't have used it." Similarly, respondent [65] stated, "I get bored using the app" and suggested that a motivation to use it would be "Maybe make the app play siege" (a shooter type video game). In addition, [65] emphasized, "I did not like meditation, I found it boring and a waste of time for me. I did like that the app ran in the background so I could use Reddit" (an internet forum site) and "I can't give a good recomodation [sic] since I found this app useless."

Interestingly, some students were personally resistant but believed the app could be helpful for someone else. For instance, respondent [5] stated, "It's not the thing for me I just don't do that type of stuff" and "Nothing that you could do to the app would make me use it" but added, "It's a good app and it would help people who want to be helped" and further noted, "Great idea for people that want to use it everything you need is right in front of you."

Hesitance

Most of the students surveyed reported hesitance. The predominant sub-themes of hesitance were that it was difficult to schedule, it was not a primary concern, they would only use the app if they needed to for their mental health, and that they experienced fluctuating results. Examples of difficulty with scheduling included statements such as “Throughout my day, I usually am doing chores or working around the house, or on my car. To be honest, I always was busy and when I did, I didn’t think to go on my phone” [38], and “Because I was too busy with work and sports. Had no time to stop and do this” [6]. Respondents who indicated that using the app was not a primary concern noted, “Most of the time it was just me forgetting. The couple times I used it was when I felt like my mental health was falling apart” [33] and “I didn’t have time and it wasn’t really a priority for me. I have a job and I leave at 7 PM so I’m usually exhausted when I get home” [75]. Data extracts from those who claimed that they would only use it if they were experiencing mental health issues clarified, “If I was more stressed in my life I might of [sic] tried it and if it was made manditory [sic] for students” [36] and “I probably would have used it if I felt sad or depressed" [76]. Respondents also provided evidence of fluctuating experiences in class—“I didn’t like that I was with people. I’d feel more relaxed if I was by myself” [42] and “I like how it made me feel calm. I disliked it because it’s not really my thing” [91]—as well as when using the app independently: “After each session, I would feel even slightly calmer or I would feel indifferent. It really depended on specific sessions” [90]. Interestingly, some students reported that they did not use the app independently but seemed to value it in school. For example, respondent [20] stated

> One reason I did not use the app outside health class [was] because it seems more of a hassle. Another reason why is because I do not have the time to use it. Another reason is because the app doesn’t work for me.

However, respondent [20] reported that during meditation in class, “It gave me the ability to relax and lower my sword” and stated they would use the app “Right before [or] after a test. During stressful moments” and recommended that “More time given” would improve the study. Some students were candid about their lack of compliance. For example, when answering Question 7—“What changes have you noticed in yourself since the beginning of this study?”—respondent [89] stated, “Honestly none but I didn’t give the app a fair enough shot.”

Acceptance

The theme of acceptance included a sub-theme that students used meditation with the app as a strategy to induce serenity. Additional sub-themes indicated that they had a positive independent experience, felt that it was a tranquil classroom experience, would use the app in school, and experienced an increase in well-being. Data extracts of the first sub-theme included:

> It calmed me down when I was paranoid. It grounded me when I didn’t think I could make it through something. It helped me improve my mental health a little bit and it changed my perspective of myself, so I hate myself less. [55]

Other data extracts included, “I mostly use it when I was trying to relax or sleep just to quiet my mind a bit” [51] and “To calm down any excess anxiety. To take me time to chill out. To take extra stress off my mind” [90]. Examples of the second and third sub-themes suggested that students had a positive independent experience—“I felt a little calm. Made me feel I was in control of my emotions. I really think it can help” [42]—and a tranquil classroom experience—“I liked how it calms me down and relaxes me. I like how it kind of makes me escape reality and focus on myself. I did not really dislike anything” [74]. In addition, respondents that supported the sub-theme that students would use the app during the school day stated, “I would 100% use it before I took a test. It would help so much” [27] and “I would maybe use it at the times that I

was anxious or nervous or when I just needed to calm down" [21]. Finally, the sub-theme of an increase in overall well-being was supported by statements such as "I'm more calm with situations that I used to get mad at" [65], "I've been more observant [sic] with myself and how I feel during certain situations" [42], "I ended up having a more positive outlook in life" [19], "I've been happier, less worried about everything, and I'm trying to speak up for myself more" [55], "I have noticed that I am quieter and able to listen and focus better" [51], and "Feeling miserable → not so miserable" [97].

Recommendations

Question 9 in the survey asked, "What recommendations do you have to improve this study experience for future classes?" Several students proposed ideas to enhance the experience or improve the study. Sub-themes included suggestions such as providing a variety of activities either on the app or in class, requiring or increasing app use in class and allowing it in other classes, purchasing access to the premium version, and providing strategies to facilitate independent use. To increase variety, students suggested adding content such as "Put meditation games on the app" [2], "The app was cool, the app should add more talks and some interaction" [40], "A music session for those who want a type of music genre to listen to [rather] than listen to someone talk" [60], and classroom activities: "I recommend that you try to get everyone envolved [sic] and after you share your thoughts and feelings with each other" [24].

A surprising finding was that some respondents believed use of the meditation app should be a requirement. Statements such as "I think this study would've been better if some of the meditation was mandatory, because this would increase student motivation" [81], "I would have used it more if it was assigned and not a voluntary thing" [23], and "If it was a homework grade or classwork grade" [69] were unexpected. In addition, some students felt more time in class and

allowing use in other classes would be helpful: “I would focus on doing it more in class & having more features because kids (like myself) get distracted by fun things to do at home” [61], “Designate more time to use app in class” [64], “I would recommend asking teachers if they would allow/offer group meditation before test: teacher would pull up a meditation video that the students could listen to” [51], and “Give out permission slips for their teachers to sign and give them permission to use the app before tests, during stressful assignments, etc.…” [89]. A few students expressed a desire for the paid version of the app: “I wish that we could access more sessions (the locked ones)” [37]. Students appeared to seek strategies that would help increase independent use. Statements related to time management—“If I was a little more organized at home, it definitely would’ve helped in using the app at home” [62]—an improved reminder system—“If I had a reminder. There was a notification from the app, but I would see it and maybe forget or was busy” [33]—and education to reinforce adherence—“Maybe talk about how its [sic] successful” [57] and “Explain the importance of practicing at home” [49].

Finally, 27 students simply indicated that they had no recommendations or wrote statements such as “I wouldn’t know what to improve” [1] and “Keep it the same” [35], and 7 respondents left the answer line blank. However, students also left comments such as, “No recommendations!! It was a good experience” [50], “Nothing. Since the app has what it needs to help someone relax” [30], and “Nothing yall [sic] doing great” [88].

Summary

This chapter reported quantitative results and qualitative findings to answer the five research questions investigated in this study. To address research question 1, this chapter reported the results of a comparative review of prior studies of the test-retest reliability of the Perceived Stress Scale with a focus on the 10-item PSS-10. Results of the test-retest reliability

analyses of time intervals (i.e., 24 hours, 3 weeks, 6 weeks, and 9 weeks) conducted with the current high school participants were reported to answer research questions 2A and 2B. Prior studies were examined and compared to results from the current high school study over similar intervals (i.e., 24 hours/2 days, 3 weeks, and 6 weeks) to answer question 2C. Results from a frequency table, created to answer research question 3, which sought to discover if students demonstrated compliance to the study protocol, showed that question 4, designed to explore the effect of the meditation intervention on stress levels as measured by PSS-10 scores, could not be addressed. Qualitative findings from the anonymous exit survey were presented and used to answer research question 5, which was developed to examine barriers and motivations students experienced with compliance to meditation, the app, and the study. Quantitative results and qualitative findings are compared, synthesized, and discussed in the next chapter.

CHAPTER 5: DISCUSSION

This study sought to examine the Perceived Stress Scale (PSS-10) (Cohen et al., 1983; Cohen & Williamson, 1988) as a stress measurement instrument and a meditation app as a stress reduction method with high school students. It consisted of five research questions and used quantitative and qualitative methods. The quantitative results and qualitative findings were presented in the previous chapter. This discussion is divided into four sections. The first section synthesizes the results of a comparative review of prior research of the test-retest reliability of the Perceived Stress Scale (PSS-14 and PSS-10). The second section discusses the results of PSS-10 test-retest reliability over different time intervals with the current study's high school participants, compares the results to prior research, and presents a rationale for use of the instrument with adolescents. The third section discusses implications and directions for future research. The fourth section explores the feasibility of using an app-based stress reduction method with high school students, discusses student compliance with independent use, and synthesizes qualitative data that may provide information to aid schools that wish to implement app-based meditation with students. Taken together, information from this study may enhance our understanding of stress measurement and app-based meditation for future stress reduction programs.

Discussion of Quantitative Results

Test-Retest Reliability in Prior Studies

Research question 1 sought to compare, analyze, and summarize previous findings of test-retest reliability of the PSS-10 with consideration to language, time interval, and age. I examined and compared the results from prior studies across the following factors: English and non-English versions; time intervals (i.e., 1 week, 2–3 weeks, and 1 month); and with adult and

adolescent participants. The literature contained several PSS-10 test-retest reliability studies conducted with adults (i.e., > 18 years of age) that used versions of the PSS-10 translated into languages other than English (see Appendix A for a full list of test-retest reliability studies). It is common to find psychometric properties included as a validity check for translated versions of instruments. However, given the depth of research that has used the PSS-10 with both English and non-English based populations, there is a surprising lack of studies that included an assessment of the test-retest reliability of the instrument.

It is important to mention that any comparison of results from prior studies must be approached with caution since the studies included disparate factors. For example, test-retest intervals ranged from 1 week to 1 month; the average age of participants in each study varied from 20 to greater than 60; and both clinical and non-clinical populations were included. In addition, results were reported using three different correlation coefficients for the metric. Researchers used the intraclass correlation (ICC), Pearson product-moment correlation, and Spearman rank correlation. This lack of metric consistency is discussed later but did not preclude comparison. In addition, researchers have used a wide variety of numerical ranges and subjective terms in an attempt to express a degree of quality for test-retest reliability results (Koo & Li, 2016; Matheson, 2019). For example, Koo and Li (2016) described 0.50 as poor, values between 0.50–0.75 as moderate, between 0.75–0.90 as good, and > 0.90 as excellent. Additionally, many researchers only consider instruments with values > .90 as acceptable for clinical use (Matheson, 2019). The inconsistency in reporting test-retest reliability results, as it pertains to this review of Perceived Stress Scale studies, will also be discussed.

English

Cohen, Kamarck, and Mermelstein (1983) developed the original English version of the PSS-14 and reported test-retest reliability of r = .85 for 82 freshman college students who retook the PSS-14 within a 2-day interval and r = .55 for a smoking cessation group of 64 adults who retook the PSS-14 after a 6-week interval. A thorough search of the literature revealed that no studies appear to have replicated Cohen et al.'s test-retest reliability research with the English PSS-14. The PSS-10 is a refined version of the PSS-14. Due to its superior internal reliability (Cohen & Williamson, 1988), the PSS-10 is the version most commonly used in research. Prior to the work done by Miller, Medvedev, Hwang, and Singh (2020), no research could be found that investigated the test-retest reliability of the English PSS-10. Miller et al. reported summed ICCs of .66, .65, and .69, but separate results for each interval were not provided and therefore comparisons could not be made.

Non-English

Test-retest reliability studies with unidimensional results from non-English versions of the PSS-10 were compared. As noted in Chapter 2, studies have reported results for each of the two factors and/or the composite (i.e., unidimensional) scores. The research community has been divided as to whether the PSS-10 is comprised of two factors, so only studies that included composite scores were used for this comparison. The studies were grouped by time interval as to when the retests occurred. There are three separate time interval categories (i.e., 7 day/1 week, 2–3 weeks, and 4 weeks/1 month). As needed, time interval labels were provided with more than one unit to reflect how the intervals were reported in the studies.

One-week interval. Some researchers may consider the results from studies that used a 1-week interval as adequate since they are > .75. However, only two of the studies reported results > .90, which, researchers have argued, may be considered excellent or acceptable for

clinical use (Koo & Li, 2016; Matheson, 2019). In addition, while the difference between .79 and .81 is very small, the difference between .79 and .95 is noteworthy. This discrepancy may be due to a difference in medical status, age, and/or language. The r_s = .79 result was from a population of healthy female university students (mean age of 22.5 with a standard deviation of ± 3.1) who completed an Arabic version of the PSS-10, and the ICC = .95 was from a group of predominantly female adults with lupus (aged 18–66 with a median age of 49) who completed a Simplified Chinese version of the instrument. In addition, two different correlation coefficients (i.e., Spearman and ICC) were compared, which may have contributed to the difference. It is important to note that two of the five studies showed test-retest reliability coefficients that were > .90 (i.e., ICC = .95 and .91), and the other three studies showed results that were lower but close to the .80 range (i.e., ICC =.86, ICC = .81, and r_s = .79). These results may indicate that the PSS-10 could possibly be stable at 1 week, but there is clearly not enough information to draw that conclusion with confidence.

Two-to-three-week intervals. As reported in Chapter 4, with an interval of 2–3 weeks, the test-retest reliability coefficients ranged from .63 to .93 and represented a wide range of ages and populations. However, these studies used Spearman, Pearson, and ICC coefficients, so an inconsistency of the metric used may have contributed to this disparity. In addition, the only study that reported a coefficient > .90, specifically ICC = .93, retested the participants at 2 weeks, and they were adults with a mean age of 38 who were in a clinical setting (i.e., a pain clinic held in an Educational and Therapeutic Center). It is possible that the center provided a supportive environment, which may have affected results. The coefficient of .93 may provide tentative evidence that the instrument could possibly be stable at a 2-week interval. However, it is the only study available for this time frame that showed a clinically acceptable correlation (i.e.,

> .90) and will need to be replicated several times before any determination can be made for PSS-10 test-retest reliability at 2 weeks.

One-month interval. Two studies used a 1-month (4-week) interval. A study of students (mean age = 20.84) showed an ICC = .83, while a study of older participants (median age = 68) reported an extremely low result of r_s = .43. This discrepancy is based on the only two studies available in the literature that used a 4-week test-retest interval. It is possible that the difference in coefficients may be attributed to mean age (i.e., 20.84 compared to 68) and/or that the first study included male and female participants while the second study was conducted with females only. However, the lack of studies with adults at a 1-month interval, which is within the range that Cohen et al. (1983) hypothesized that the English Perceived Stress Scale would remain stable, prevents researchers from drawing any conclusions about the 1-month stability of the instrument. Most importantly, the coefficient of .83 is still much lower than a threshold of .90, which many researchers deem necessary to consider a psychometric instrument appropriate for clinical use (Matheson, 2019).

Adults

The above review of non-English PSS-10 studies showed that the instrument appeared to be stable for clinical use (i.e., ICC > .90) in two studies with a 1-week interval and one study with a 2-week interval. The three studies that reported results > .90 were conducted with adults who ranged in age from 18 to 66. It is important to note that two studies, which reported results < .90, may have included some participants who were younger than 18. In one of these studies, a group of 129 Chinese university students that retested at a 2-week interval were selected from a group of 1,096 students with a mean age of 18.3 and a standard deviation 0.7 years. Additionally, 70 students from a private university in Malaysia, who retested at 2-weeks, were selected from a

cohort of 242 students with a mean age of 20.9 and a standard deviation of 6 years. Therefore, it is possible that some of the students who participated in the retest portion of each of these two studies may have been less than 18 years old. However, the specific ages of the retest groups were not reported, so it is unclear if students < 18 were included. Age notwithstanding, the most important point is that there is simply not enough data to draw a definite conclusion about the stability of the non-English PSS-10.

While there are too few studies with the non-English PSS-10, there is a significant gap in the literature and complete lack of data on the test-retest reliability of the English version of the PSS-10. As Lee (2012) emphasized, this research needs to be conducted with adults from diverse populations and cultures and should include a wide span of ages. Moreover, Cohen et al. (1983) stated that the PSS-14 was intended for use with populations who possessed a minimum of a junior high school education. However, Cohen et al. (1983) did not specifically include this age group during PSS-14 test development. It is important to note that the sample of 82 freshman university students who participated in the test-retest reliability portion during the original instrument development study of the PSS-14 were culled from a larger sample of 332 freshman with a mean age of M = 19.01 with a SD of 2.75 (Cohen et al., 1983). Although some of the test-retest participants may have been less than 18 years old, the mean age and standard deviation of the 82 freshman who retested were not reported. Furthermore, the data used when Cohen and Williamson (1988) refined the PSS-14 to the PSS-10 were from participants > 18. Therefore, it did not appear that the adolescent population was considered during the development of the PSS-10.

Adolescents

A thorough review of the literature showed that there appeared to be a total deficit of studies, in any language, of PSS-10 test-retest reliability conducted exclusively with adolescents. Although two of the prior studies with adults may have included some participants who were less than 18 years old, both studies were conducted with a population of young adults in university settings. Those who attended university may have experienced a different environment than students who attended high school. The university students may have faced new situations without their previous support systems of family, friends, or teachers (Bland et al., 2012). Additionally, university students may have contended with unfamiliar challenges (e.g., transition to a new setting, increased academic requirements, managing finances, and navigating new relationships) (Huberty et al., 2019; Ramasubramanian, 2017). Due to these differences, university participants may have been developmentally more mature than adolescents still attending secondary school. Therefore, the results from these studies do not accurately reflect the adolescent population. This is a notable oversight since the PSS-10 has been used in several studies with adolescents (Bluth et al., 2015; M. Braun et al., 2014; Foret et al., 2012; Kohn & Milrose, 1993; Lemon & Watson, 2011; Siqueira et al., 2000). A thorough review of the literature did not result in any published article that addressed the test-retest reliability of the PSS-10 with an adolescent population. Moreover, no study, with adults or adolescents, could be found that replicated the original 48-hour and 6-week interval test-retest reliability research conducted by Cohen et al. (1983). In an attempt to fill these significant gaps, the current high school study used the English version of the PSS-10 with adolescents and conducted test-retest reliability analyses at intervals of 24 hours, 3 weeks, 6 weeks, and 9 weeks. The 24-hour and 6-week intervals from the high school study are comparable to the 2-day and 6-week intervals used by Cohen et al. (1983).

Test-Retest Reliability in the Current High School Study

24-hours. The 24-hour test-retest reliability of the PSS-10 in the high school study's sample of adolescent students was examined to answer research question 2A. In the high school study, the consistently high ICCs > .91 for 24-hour intervals showed excellent PSS-10 test-retest reliability. Taking into consideration that the participants were adolescents, these results were surprising. Researchers have reported that teenagers are still developing emotional regulation skills, can be impulsive, and may be more vulnerable to stress than adults (Arain et al., 2013; Eagleman, 2015; Giedd, 2008; Siegel, 2014; Wu et al., 2021). Indeed, those who have worked or lived with adolescents have often witnessed the wide range of emotions that some teens have exhibited within a very short time frame (i.e., hours, minutes, or seconds). Therefore, these excellent 24-hour ICCs were unexpected. However, the results indicated that the instrument may be a dependable 24-hour measure with adolescents since it was highly stable from one test to the next. In addition, this may signify that the PSS-10 meets the threshold for clinical use with an adolescent population at a 24-hour interval. This implies that the PSS-10 could be used to accurately measure the self-perceived stress levels of high school students. However, intervals longer than 24 hours were also examined to determine the magnitude and duration of the PSS-10's stability over time. In the current high school study, data analysis of the control group provided additional test-retest reliability coefficients for 3-week, 6-week, and 9-week intervals.

Comparison of intervals. Research question 2B sought to investigate how test-retest reliability changed as a function of time and specifically focused on the 24-hour, 3-week, 6-week, and 9-week intervals available from the high school study. As stated above, the ICCs for four separate 24-hour intervals were excellent at > .91. A 6-week interval resulted in an ICC = .69 (n = 62), and a 3-week interval showed an ICC = .63 (n = 55). In addition, the 9-week

interval from pretest to posttest showed an ICC = .55 (n = 49). While some researchers may believe that the 9-week result could be interpreted as moderate, it is very close to .50, which most researchers would consider to be poor. As noted in Chapter 4, the 6-week interval occurred between pretest and midtest and therefore preceded the 3-week interval that occurred between midtest and posttest. The ICC correlations for the 6-week interval are slightly higher than those for the 3-week interval (i.e., .69 vs. .63), which appears counter to the belief asserted by Cohen et al. (1983) that shorter intervals will result in higher test-retest correlation coefficients. This significant drop at a 3-week interval was of particular interest. Additionally, since the 6-week interval occurred between September and October and the 3-week interval occurred between October and November, the point in time in the school year during which students completed the PSS-10 may have been a factor. For example, November may have been more stressful for students as they transitioned from the first marking term to the second. Academically, they may have been concerned about their grades and increasing workloads. On top of their academic work, students who participated in athletics, extracurricular activities, and/or held jobs after school may have started to experience fatigue. In addition, the climate in the Northeast could have been a factor since the temperature drops and the amount of sunlight decreases. Moreover, the upcoming holidays may have been a source of stress for some students. To consider the point in time during the school year, the data was analyzed with an ANOVA and Friedman test.

Point in time during school year. Results of both the one-way repeated measures ANOVA and the Friedman test indicated that, while PSS-10 average scores declined and fluctuated slightly over time, the results were not statistically significant. Although the ANOVA and Friedman tests determined that the point in time, during the fall semester, when the participants completed the PSS-10 did not affect student scores, it also raised a question as to

why the coefficient at the 3-week interval was lower than the 6-week interval (i.e., .63 vs. .69). Since the 3-week interval occurred after the 6-week interval, one possible explanation is test fatigue since the 3-week correlation was based on the midtest and posttest. Students may not have answered the items on the posttest with as much care and attention because it was the fifth and sixth time that they had taken the PSS-10 during that semester. Future researchers will need to control for this possible confounding factor in their research design. One possible solution may be to counterbalance the order of the different time interval and month of administration across the academic calendar to help control for these potential confounding variables.

Comparison of High School Study to Prior Studies

Due to the absence of studies on the PSS-10 with adolescents, studies with adults were used to answer research question 2C that compared the test-retest reliability of the PSS-10 in the high school study to the test-retest reliability of the PSS-10 in prior studies. The high school study's higher correlation coefficient of .92 (24-hour interval) compared to Cohen et al.'s .85 (2-day interval) may be attributed to a difference in age and developmental level, the additional day between testing in the latter study, or another yet unidentified confounding variable. It is important to mention that the mean age of the students in the high school study was 16.77, while the college student sample in Cohen et al.'s study was gleaned from a population with a mean age of 19.01. However, the specific mean age of the college student test-retest reliability sample is unknown. In addition, the correlations may differ because, as previously noted, college students could be dealing with ongoing challenges related to the college experience that high school students are not faced with. Furthermore, Cohen et al. (1983) did not specify the testing environment as to time of day, setting, or test administrator. In addition to a shorter interval, in the current high school study, all PSS-10 tests occurred at the same time of day, in the same

classroom, and I administered all tests, eliminating and controlling for any halo or interactive effects based on different test administrators. In addition to these factors, Cohen et al. (1983) used the PSS-14 for their test-retest reliability analysis since the PSS-10 was not developed until 1988 (Cohen & Williamson, 1988). The four additional questions on the PSS-14 may have contributed to the lower test-retest reliability coefficient correlation of Cohen et al.'s (1983) study. Students in the current high school study took the slightly shorter PSS-10, which may have affected the results.

Additionally, the current high school study showed a higher coefficient of $r = .69$ for 62 students who retook the PSS-10 at a 6-week interval compared to Cohen et al. (1983), who reported $r = .55$ for the smoking cessation group of 64 adults who retook the PSS-14 after a 6-week interval. The high school study's higher correlation coefficient may have been due to a difference in age (i.e., adolescents $M = 16.7$ versus adults $M = 38.4$), but it is important to note that a smoking cessation group is not a general population group, and the participants who were trying to stop smoking may have experienced greater fluctuations in stress during the 6-week interval. Moreover, similar to the shorter interval (i.e., 24 hours/2 days) comparison, the four additional questions on the PSS-14 may have been responsible for a difference in test-retest reliability for a 6-week time interval compared to the high school study that retested at a 6-week interval using the PSS-10.

The only prior study available for a 3-week interval reported an ICC = .82, which is much higher than the current high school study's ICC = .63 within the same 3-week interval. As stated above, the current high school study's 3-week interval was measured from midtest to posttest, and it is possible that test fatigue may have affected the results. In addition, the much higher coefficient of .82 may be due to a difference in language and/or culture (i.e., Malay versus

English) or a difference in age and developmental level since the mean age of the young adult university students in the prior study was 20.9 and the mean age of the current high school study's adolescent population was 16.77. Furthermore, this striking difference of a much lower correlation coefficient of .63 with the current high school study may suggest that 3 weeks is much too long of an interval for the PSS-10 to remain reliable with an adolescent population.

Practical Utility of the PSS-10 with Adolescents

Research has shown that adolescents may behave differently than adults, and therefore their ability to perceive stress may differ. Areas in the teenage brain responsible for thinking, judgement, self-regulation, and impulse control are still under development, and adolescents are still learning how to self-regulate and self-reflect (Eagleman, 2015; Kass, 2017). Research has revealed that there is an increase in neuroplasticity during late adolescence—the developmental period when a teenager's identity is capable of being molded and habits and choices can be influenced (Roeser & Pinela, 2014). Therefore, the high school years may be an optimal time to measure stress levels and provide strategies for stress management.

The results from the current high school study suggest that the PSS-10 possesses excellent 24-hour stability and supports the use of the instrument as a screening tool to identify teens who may be experiencing high levels of stress within the context of a short time period. Identifying stress levels of students could begin with entry to high school. This may allow school staff to provide students who score high on the PSS-10 with support or intervention. The PSS-10 is free for educational use, brief, and easy to score, which makes it ideal for quick assessment. In addition, it can be administered to large groups in a short period of time. Therefore, it may not significantly impact time on learning if administered during class. Baseline stress levels of freshman students could be collected, and students could be reassessed as they progress

throughout the year and/or year to year. In addition, stress levels could be compared between groups such as those preparing to go to college, work, the military, or other post-secondary pursuits. The PSS-10 test-retest reliability, investigated during this study, may help guide future research and provide valuable information to inform practice for stress management programs with adolescents.

The current high school study appears to have been one of the first attempts to thoroughly examine the test-retest reliability of the PSS-10 with an adolescent population and may contribute to the extant literature on the psychometric properties of the scale. Since test-retest reliability studies conducted with adolescents could not be found in the literature, the current high school study results were compared to the results from prior studies with adult populations. Analysis of 24-hour, 3-week, 6-week, and 9-week intervals provided tentative evidence that the PSS-10 is an excellent short-term measure of stress in adolescents but may not be stable after 24 hours. Furthermore, the significantly lower test-retest result of .63 at 3 weeks in the high school study, compared to .82 with an identical interval from a study with adults, is interesting. However, for an instrument that has been used in as much research as the PSS-10 has, there are too few studies with adults and no prior studies with adolescents to allow meaningful comparisons and draw conclusions about the instrument's test-retest reliability and stability.

Implication of Findings

There is a documented increase in stress that can impact work, school, relationships, and many aspects of life (O'Connor et al., 2021). High levels of stress may result in unhealthy physical conditions (O'Connor et al., 2021) and contribute to an increased risk of developing mental health conditions such as anxiety and depression (Sapolsky, 2004). The stress levels of all age groups are on the rise but are at unprecedented levels in teenagers (APA, 2014). This is

particularly concerning since over 50% of mental health conditions arise during the teenage years (Konaszewski et al., 2021). Additionally, stress can have a negative effect on learning by decreasing attention, memory, and focus (Metz et al., 2013). Researchers have argued that stress in teenagers has been neglected, adolescent health measurement is lacking (Guthold et al., 2021), and adult stress instruments may not be appropriate for use with adolescents (Byrne et al., 2007). Therefore, it is vital for researchers to include adolescents in studies of stress measurement and treatment.

The Perceived Stress Scale was originally developed to measure stress for research, but it has often been used in clinical settings (Taylor, 2015). To date, the PSS-10 is one of the most popular instruments used by researchers to measure psychological distress and determine the effect of interventions in treatment outcome studies (Galante et al., 2021; Makhubela, 2020). These findings draw attention to the importance of considering four factors that may benefit the research community. First**,** the significant lack of test-retest reliability in the psychometric research of the English PSS-10 is concerning, and it is important to continue to assess the instrument. As noted, while test-retest reliability studies of the non-English versions exist, they are limited and difficult to compare due to differences in variables such as age of population, language, and culture. In addition, the results from non-English PSS-10 versions may not be comparable to results from the English PSS-10, but until more research is conducted this cannot be known. Moreover, this research should be conducted with groups of adults and adolescents from diverse populations regarding gender, sexuality, race, ethnicity, socioeconomic status, and ability. This increase in research of the English PSS-10 would provide continued critical examination and may allow robust comparisons to be made.

Second, the results from prior studies were reported using the ICC, Pearson, and Spearman correlations, which may have affected comparisons. Researchers have advised that the ICC should be used for test-retest reliability research with a self-report instrument such as the PSS-10 (Koo & Li, 2016; Perinetti, 2018; Vetter & Schober, 2018; Yen & Lo, 2002). Furthermore, researchers have recommended that authors should specify which ICC form was used in analysis and cautioned readers to determine if it was the correct form before relying on the results (Koo & Li, 2016; Sainani, 2017). Therefore, it seems practical to suggest that future researchers who investigate PSS-10 test-retest reliability use the ICC and specify which form was used.

Third, in addition to coefficient inconsistency, the research community has not established standard thresholds for reporting test-retest reliability correlation results. Indeed, explanation of test-retest reliability for psychometric instruments results varied considerably in the literature. Some researchers reported > 0.75 as excellent while others considered > 0.90 as "acceptable" (Matheson, 2019). Koo and Li (2016) described 0.50 as poor, 0.50–0.75 as moderate, 0.75–0.90 as good, and > 0.90 as excellent. While it appeared that researchers concurred that < 0.10 is a weak correlation and > 0.90 a strong one, the lack of firmly established standards requires cautious interpretation (Schober et al., 2018). Due to these inconsistencies, it may be worthwhile for the research community to agree to eliminate descriptive labels and simply report test-retest reliability correlation coefficient numbers. The numbers are clear and straightforward. A decision to rely on coefficient numbers may help eliminate confusion and facilitate interpreting results or comparing studies. In addition, reporting coefficient numbers only may allow future researchers to bypass subjective labels and select an instrument based on the reported ICCs for a specific time interval.

Finally, and most importantly, if the research community continues to use the PSS-10 to detect a change in stress level to assess the effect of interventions, it will be essential to determine temporal stability. It may be necessary to assess test-retest reliability with several different time intervals to establish when the stability of the PSS-10 begins to decline and if the coefficient is high enough for the time interval to accurately determine if a treatment is successful. Unfortunately, at the present time, there are too few studies available to make this determination. For example, three prior studies reported ICCs appropriate for clinical use (i.e., > .90), two at a 1-week interval and one at a 2-week interval. However, these results were with adults using the non-English PSS-10 and were conducted in different languages. Not only are there too few studies to be able to draw a conclusion about test-retest reliability, but several studies will need to be conducted in the same language to allow for comparison. In addition, such studies in each language should account for diversity within identity (e.g., gender, sexuality, race, ethnicity, socioeconomic status, and ability). The documented inconsistency of results from prior research of test-retest reliability of the PSS-10, coupled with the limited research with the non-English versions and a total absence of research with the English version, justifies a need for extensive investigation with adult populations. Most importantly, except for the current high school study, there is a glaring deficit of research on this topic for or with adolescents.

The current high school study showed a test-retest reliability coefficient of .63 at 3 weeks. Undeniably, .63 is markedly less than the > .90 results obtained after a 24-hour interval in the high school study. Therefore, the findings from the current high school study strongly suggest that the PSS-10 may not be stable for more than a day with adolescents. More research is needed to see if the current high school study's excellent 24-hour test-retest reliability results can be replicated with other adolescent populations. As suggested for adults, such studies should

include representation from a variety of populations. Moreover, it is worth considering that stable test-retest reliability intervals may vary between adults and adolescents. Research has shown that adolescents are particularly vulnerable to stress because the prefrontal cortex (PFC) does not fully mature until the mid-20s (Eagleman, 2015; Wu et al., 2021). This lack of PFC development can result in impulsive behavior (Siegel, 2014). When faced with stressful situations, adolescents often respond more rapidly and acutely than adults (Erbe & Lohrmann, 2015; Jensen & Nutt, 2015; G. C. Patton et al., 2016; Siegel, 2014). These differences, due to a still developing brain, may account for differences in the results of test-retest reliability studies with adolescent populations. Therefore, it is imperative that researchers make it a priority to include this age group in future PSS-10 test-retest reliability studies to test for possible differences in stability with adolescents.

The test-retest reliability of the PSS-10 needs to be pursued through rigorous investigation with both adolescents and adults. These concerns are consistent with researchers who argued that systematic longitudinal studies are needed to clarify the degree to which PSS-10 scores remain stable over time (Lee, 2012; Roberti et al., 2006). The current high school study showed a significant drop after a 3-week interval. Therefore, it will be essential for future research to focus on incremental time intervals (e.g., 24 hours, 72 hours, 1 week, 2 weeks, etc.). It may be necessary to conduct several studies at each interval to determine when the metric begins to decline.

Suggestions for Future Research

Test-retest reliability coefficients have often been used to determine if a psychometric scale measures a construct as a state or trait. Generally, coefficients < .60 indicate that the instrument measured a state, and coefficients > .70 indicate the instrument measured a trait

(Medvedev et al., 2017). Although McCrae et al. (2011) referenced test-retest reliability in personality scales, they raised an important point when they questioned if variations were more prevalent in adolescents than in adults. The high school study produced clinically acceptable 24-hour test-retest reliability coefficients > .91, two coefficients of .63 and .69 for intervals of 3 weeks and 6 weeks respectively, and a coefficient of .55 for a 9-week interval. The lower coefficients at 3-weeks and 6-weeks may indicate that the participant experienced a true change in their stress level. In addition, the decrease in coefficient as the time interval increased appeared to be consistent with Cohen et al. (1983) who asserted that test-retest reliability coefficients should be much higher for short retest intervals compared to longer ones. However, the differences in test-retest reliability results make it difficult to attempt to determine if the instrument measures a fleeting state or an enduring trait.

The binary concept of state and trait is complex, and the documented variation in test-retest reliability results over different time intervals for the PSS-10 may indicate that the instrument will transcend such labels. Similar to the argument presented by Allen and Potkay (1981), who claimed that the distinctions between state and trait in personality tests were not completely rigid and declared them to be arbitrary, such binary labels may not capture the complexity and contextual aspect of perceived stress. For example, an air temperature of 72°F may be perceived as warm to some and cool to others. However, this perception may depend on variables that include, but are not limited to, humidity, wind speed, level of cloud cover, prior acclimation, and type of clothing worn. Therefore, it may be important for future researchers to bypass the issue of determining if the PSS-10 measures a state or a trait and simply report the test-retest reliability coefficient result for each time interval. This may avoid the subjective labels of state or trait and allow researchers to draw their own conclusions when they interpret the

results. For example, they may find that the PSS-10 will retain excellent stability for 1 week but not for longer intervals. Moreover, future researchers could use the test-retest reliability coefficient numbers reported for a given interval of time to determine the utility of the instrument and decide if it meets the needs and demands of the time sensitivity required for their research design.

Historically, researchers have used the dichotomous terms of state and trait to describe the Perceived Stress Scale. Indeed, Cohen et al.'s (1983) assertion that the instrument is a state measure conflicts with Miller et al. (2020), who applied generalizability theory (G theory) and argued that the PSS-10 is a trait measure. While the high school study did not employ G theory, it is interesting to note that Miller et al. reported similar ICCs of .65, .66, and .69, with an intervention group, full sample, and control group respectively prior to applying G theory to their data sets. Miller et al. claimed these results spanned three testing sessions (i.e., pretest, 8-week post-intervention, and 6-week follow-up), but specific results for each interval were not provided and therefore cannot be compared to other studies. In addition, the participants in Miller et al.'s study were adults (M = 47.07 years). However, it is unclear how Miller et al. concluded that ICCs of .65–.69 represented trait-oriented metrics since > .70 has historically been indicative of a trait measure. Future studies may benefit from replicating the work of Miller et al. On the other hand, although the dichotomous terms of state and trait have been used in the past, the bigger issue may be to replicate prior studies, without employing G theory, to concentrate on finding the optimal retest interval(s) that will provide consistently stable results for adults and for adolescents.

While it is critical to demonstrate that the instrument is a sound and clinically acceptable measure in the short term (i.e., 24 hours), it may also be possible for future researchers to attempt

to pinpoint when the phenomenon of perceived stress changes and how often the metric may vary. Future research could involve changing the directions to the participant on the PSS-10. For example, instead of instructing respondents to provide answers based on the past month, they could test out various time prompts (i.e., 24 hours, 1 week, 2 weeks, etc.) to see if such changes had an impact on the results. The original instructions, which direct the respondent to reflect on their feelings and thoughts during the last month, may serve to diminish the peaks and valleys of perceived stress that they have experienced over the previous 30 days. For example, as the respondent considers the events of the last month, highly stressful situations may have occurred and then resolved within the same time frame. This may result in a situation where their perception of the event is altered, and it may be rendered less stressful. The process of reflecting over the past month may have helped to solidify and stabilize the measure.

On the other hand, some respondents may unconsciously answer the questions based on their most recent (i.e., past day or week) feelings and thoughts and not reflect on a full 30-day retrospective. This situation may serve to either exacerbate or reduce their memory of stressful events, and their responses may not accurately represent their thoughts and feelings over the last month. Given these concerns, using shorter time prompts may allow the respondent to fully focus on recent experiences and may subsequently provide different results. It may be worth investigating what the results would be if the directions on the instrument matched the time frame of the retest (e.g., 24 hours, 48 hours, or 1 week). Such results could be compared to results attained when the directions on the instrument did not match the retest time frame, like those of the current high school study, which reported excellent 24-hour results but directed students to recall the events of the last month. For example, instead of asking respondents to reflect on the last month and retesting 24 hours later, the directions on the instrument could

prompt them to reflect on the last 24 hours and then retest 24 hours later. It is interesting to note that the excellent 24-hour results from the current high school study may have indicated that the respondent had a stable memory of the events that occurred over the last month from the first administration of the PSS-10 to the second administration 24 hours later. Therefore, the 1-month recall was stable for 24 hours. In essence, the instrument assessed the adolescents' memories of the events of the past 30 days twice with a 1-day interval between the first and second administration. In the future, it may be important to further unpack the 24-hour interval with adults and compare the results to the current high school study.

Moreover, it would be interesting to investigate daily stress levels and compare the results to a monthly score. For example, the PSS-10 could be used to measure stress levels each day for 1 month. The range of scores could then be examined to investigate how accurately daily stress measures may predict the results of a one-time score that instructs participants to review their thoughts and feelings over the last month. Daily measures may serve to highlight the peaks and valleys of perceived stress. An average of 30 daily scores may or may not be equal to a one-time score that instructs respondents to reflect on the last month. Additionally, this research could be conducted with different age groups of adults and adolescents to explore if developmental differences impact the results.

One interesting observational theory related to developmental levels arose during analysis of the prior non-English test-retest reliability studies. In the future, it may be worth examining studies to see if test-retest reliability coefficients increase or decrease with the age of the population tested. Exploring a possible relationship between age and test-retest reliability coefficients could investigate how perceived stress may vary in populations that represent different stages of human development (i.e., adolescents, young adults, adults, and aging adults).

However, this observation occurred during an examination of the very limited research conducted with non-English versions of the PSS-10. The documented lack of information with non-English versions serves to highlight the total absence of test-retest reliability research with the English PSS-10 with both adult and adolescent populations. The current lack of research of PSS-10 test-retest reliability prevents a comparison of results to investigate if there may be any possible pattern related to the stages of development and changes in stress level.

Additionally, daily stress levels may rise and fall within seconds as an individual perceives a threat and then it vanishes—for example, momentarily thinking that one has lost their cell phone only to find it in a different pocket. Such fleeting states may occur more often for some personality types, who tend to fear the worst, while others may be more equanimous and perceive the experience differently. Indeed, personality traits have been shown to affect an individual's perception, coping methods, and recovery (Childs et al., 2014; Luo et al., 2017). Specifically, researchers have asserted that neurotic individuals may tend to experience a greater degree of perceived stress and possess less effective coping strategies (Barron & Gore, 2020; Carver & Connor-Smith, 2010; Ebstrup et al., 2011; Piekarska, 2020; You et al., 2020). Future research could continue to investigate the possible correlation between personality traits and an individual's perception of stress as measured on the PSS-10.

Cohen et al. (1983) posited that perceived stress was "dependent on personal and contextual factors" (p. 386) and not just a compilation of life events. Indeed, the construct of perceived stress may be expected to vary over time and may be influenced by ongoing events and different perceptions. In fact, some researchers have argued that the scale is divided into two factors and claimed that Factor 1 measured perceived distress while Factor 2 measured the perceived ability to cope (Makhubela, 2020; Sun et al., 2019). Makhubela (2020) reported that

many researchers have supported the two-factor model over the unidimensional one but conceded that the differences were minor and, in some instances, conflicting. In addition, Michaelides, Christodoulou, Kkeli, Karekla, and Panayiotou (2016) found very small distinctions between the two factors, which may support Cohen and Williamson (1988), who originally suggested a composite score for the instrument should be used. However, it may be worthwhile to conduct additional studies with the English PSS-10 to further explore factor structure and its possible influence on other psychometric properties such as test-retest reliability.

Summary of Quantitative Discussion

In this study, research questions 1 and 2 sought to examine the test-retest reliability of the PSS-10 from prior studies, investigate the test-retest reliability with an adolescent population over different time intervals, and compare the results from adolescents to those of the prior studies. This discussion has acknowledged that the PSS-10 is a popular, practical, and useful tool for stress research. However, this study has also laid the groundwork and identified that further investigation is warranted. Future research of both the English and non-English versions of the PSS-10 is needed to address four areas. First, additional test-retest reliability studies should be conducted with a variety of populations and age groups of adults and adolescents. Second, future test-retest reliability research should choose a form of the ICC for analysis to provide consistency and aid in comparison to other studies. Third, it may be practical to establish standard values and terms to describe test-retest reliability results. Fourth, and most importantly, the research community should make every effort to determine a finite interval (e.g., 24 hours, 1 week, or 1 month) of temporal stability for the PSS-10 with both adults and adolescents. This last point is particularly critical for researchers who intend to conduct treatment outcome studies to

ensure that a change in PSS-10 scores signifies that a true change has occurred, and the result is not due to the instability of the instrument.

In particular, there is a need to determine finite test-retest reliability coefficients for specific time intervals to determine the stability of the PSS-10 and pinpoint when that stability begins to decline. Cohen et al. (1983) reported a test-retest reliability of .85 with a 2-day interval. However, that result was from a study of the English PSS-14, which may not be comparable to the English PSS-10. No other studies could be found that reported test-retest reliability results for specific intervals with the English PSS-10. The limited studies of unrelated (i.e., conducted in different languages and with different populations) non-English versions of the PSS-10 showed that two studies reported results that indicated the instrument may meet the test-retest reliability requirements needed for clinical use with adults at a 1-week interval, but two studies are not sufficient to draw such an important conclusion with conviction.

The results from the current high school study showed that the English PSS-10 was highly stable for 24-hours with adolescents but dropped sharply at a 3-week interval. Future research will be needed to see if these results can be replicated. In addition, it will be necessary to conduct future test-retest reliability studies, over different time intervals, with both the English and non-English versions of the PSS-10 and with participants who represent a wide variety of ages and populations. Beyond the current high school study's 24-hour results, and with consideration given to prior studies of different retest intervals, conducted with adults or adolescents, using either the English or non-English versions, there was simply not enough evidence that the PSS-10 demonstrated consistent test-retest reliability. Moreover, the single overarching question that arose from this portion of the study was if it may be possible to

determine if any version of the PSS-10 will be sufficiently stable for longer intervals (e.g., 1 month) to be used in research or treatment outcome studies that require such a time frame.

While future research into the above-mentioned four areas is paramount, the popularity of the PSS-10 may warrant further investigations. These may include determining if the commonly used, but possibly outdated, terms of state or trait are appropriate descriptions for the complex construct of perceived stress; exploring any possible correlation to personality traits; testing the variance of daily scores; and continuing to investigate the instrument's factor structure. Prior to this study, little was known about the test-retest reliability of the English PSS-10 with adults, and there was no data on adolescents. Taken together, these recommendations are provided to help ensure that the PSS-10 will continue to be considered a valuable and reliable tool for future research with all populations but utilized only in the most appropriate research designs based on the findings of the aforementioned future directions section.

Discussion of Exploratory Results

Compliance

Research question 3 investigated if participants adhered to the baseline compliance of using the app to meditate independently at least four times a week during the 8-week intervention. An examination of the frequency table (see Appendix J) showed that students were not compliant. These findings appeared to be consistent with findings from other researchers who reported that non-adherence to digital interventions was a concern (Gál et al., 2020; Mrazek et al., 2019; Psihogios et al., 2020; Weber et al., 2019). There are several possible reasons as to why compliance to the app-based meditation in this study was poor. One possible explanation for lack of compliance could be due to the developmental phase of the participants.

The age of students in this study (M = 16.7 years) encompasses the developmental phase of adolescence and the beginning of emerging adulthood (Arnett, 2000). Adolescents are still growing and becoming more mature. Research has shown that the teenage brain is still evolving and undergoing a complex process of neural pruning and refinement (Giedd, 2008). These changes are particularly evident in the prefrontal cortex, an area of the brain responsible for the executive functions that control organization and planning, which will continue to be refined until the mid-20s (Arain et al., 2013; Giedd, 2008). During this developmental stage, teens are vulnerable to poor decision-making, but they also possess an increased ability to learn and adapt (Blakemore & Choudhury, 2006; Dunning et al., 2019; Giedd, 2008). A lack of organization and planning, coupled with poor decision-making, may have contributed to the noncompliance exhibited by this study's population of adolescents. Undeniably, quantitative data from the frequency table demonstrated the students' lack of adherence to meditation. This rendered research question 4, which would have compared Perceived Stress Scale (PSS-10) scores from the control and waitlist groups, as irrelevant. However, the increased adaptability of this phase may have provided an opportune time to capitalize on school as an ideal place to learn, practice, and derive support for app-based meditation. Therefore, research question 5 was developed and designed to use qualitative methods. Responses from the anonymous exit survey were analyzed to explore the barriers and motivations students experienced with meditation, the app, and the study. Without question, data from the qualitative analysis provided a granular and comprehensive view of the student experience.

Discussion of Qualitative Findings

Responses from the survey were used to answer research question 5, which sought to identify the barriers and motivations students experienced with independent app use for

meditation. Findings from the qualitative data were viewed with substantial consideration of the students' developmental phase and my perspective as a veteran teacher and counselor. These lenses were critical and provided me with a valuable perspective and insight into the student experience. The themes and sub-themes generated from the data provided information that helped determine the feasibility of using app-based meditation with high school students. These findings may prove useful to those who wish to develop, introduce, and support app-based meditation as a classroom initiative and establish a schoolwide stress management program.

The sub-themes generated under resistance, hesitance, and acceptance provided data that may help inform a framework to foster and execute a schoolwide app-based meditation program to help reduce student stress and increase well-being. The sub-themes identified under acceptance supplied motivations that could be built on. Most importantly, the sub-themes identified under resistance and hesitance presented evidence that could be used to address barriers to help support app-based meditation and encourage student use. The findings from this study are preliminary but provided some promising initial evidence that it may be feasible to develop a stress management program that uses app-based meditation as a stress reduction method for high school students. This evidence prompted me to recommend a potential protocol for implementation.

Feasibility and Practical Utility

As seen in data analysis, students reported barriers to compliance and revealed motivations to use the app. The findings showed that most students believed they could benefit from meditation if they could find time or remember to do it. Several students reported that they experienced increased emotional regulation and a decrease in reactivity to stressful situations. These findings are consistent with the Transactional Model of Stress and Coping (Lazarus &

Folkman, 1984) that formed a theoretical basis for this study. Moreover, one of the most promising findings is that most students who were resistant or hesitant to engage in independent meditation reported that they had a positive experience with classroom meditation. An unexpected sub-theme was generated from students who reported that they would have been more motivated to meditate if it was a part of the curriculum or a required assignment. This surprising finding was particularly helpful and provided a basis for my recommendations.

The findings appeared to show that app-based meditation could be successful if it was introduced, implemented, and supported in the classroom with additional support provided for both independent and schoolwide use. A synthesis of themes and recommendations suggested that a three-prong approach may be the most feasible way to develop a comprehensive program to address and support student needs. The aim of this three-part approach would be to educate staff and students, build meditation into the curriculum, and encourage, promote, and provide schoolwide support for the practice. First, provide training on stress management and app-based meditation for school staff and foster stress awareness for students via classroom curriculum. Second, slowly integrate a variety of meditation activities into the classroom and then require independent practice. Third, promote and support schoolwide use. Committing to action in the three overarching areas of education, practice, and support may set the stage for a systemic program.

Education

Staff. An important first step may be to provide the school staff with information about the possible effects of chronic stress. This study showed the PSS-10 possessed excellent short-term stability. The PSS-10 could be distributed to staff so they could measure their own stress level. This may raise awareness and provide motivation for staff members to engage in

meditation as a stress management activity. Perhaps the best foundation to introduce this as a schoolwide initiative would be to allow the staff time to download and use the app and enable them to become familiar with it. This may allow teachers who elect to deliver direct app instruction to feel more comfortable before they decide to implement the method with students or support its use in other school settings. One advantage of using an app is that meditation instruction is provided in the sessions and therefore eliminates the need for "teaching" the practice. The only resources needed to implement app-based meditation are a supportive teacher and classroom time. In addition, all staff could be equipped with this information and therefore be capable of supporting app use in class, during lunch, or other times during the school day. In this study, the classroom teacher was highly supportive of meditation with the app, and this attitude may have directly impacted student acceptance. Moreover, providing school staff with a self-care strategy that may reduce their own stress could have long-term positive effects on both the individual and the school community.

Students. Administering the PSS-10 to each student to measure and assess their own stress level may motivate and encourage students to learn more about stress and stress management. Students recommended that the study would have been better if it included more information about the effectiveness of meditation (i.e., "proof that it works") and the importance of routine home practice. As with the staff, it would be important to provide students with a wide variety of information about the negative physical and psychological effects of stress, importance of stress management, and the potentially promising use of meditation to reduce or prevent stress. Moreover, education may help address barriers for students who reported that they did not need, want, or like meditation or felt that it did not work. Yeager, Dahl, and Dweck (2018) suggested interventions for teens were more effective if they demonstrated respect and did not

threaten a student's social status. Allowing students to measure their own stress level and identify the sources of their stress could be a potential solution to respectfully engage students and may appeal to their individuality in a manner that would not threaten status.

Student interest may be increased by providing information from respected athletes, musicians, artists, leaders, etc., as well as those who represent the study's population (i.e., similar in age, gender, race, and ethnicity) who have benefited from a meditation practice. Role models such as these may lend credibility to the practice of meditation as a stress reduction method and motivate students who are resistant or hesitant. In addition, education about stress prevention may be helpful to students who appeared to believe that meditation, or any other stress reduction method, is only necessary if someone is highly stressed or suffering from mental health issues like anxiety or depression. Furthermore, it may benefit students who claimed that they did not need, want, or like meditation since they may have felt that admitting to needing or seeking help was a threat to their status.

Practice

Classroom. Students who reported that they were motivated to use the app cited reasons consistent with the sub-themes of acceptance. They claimed their intention was to relax, calm down, pause to think, aid sleep, and relieve stress, depression, or anxiety. Except for sleep, routine classroom meditation could accomplish similar goals. Survey findings showed that a vast majority of participants reported that classroom meditation was a positive experience and claimed they experienced feelings such as being calm, relaxed, in control, focused, and less stressed. Several students indicated that scheduled meditation would be a welcome addition to the class and a pleasant break in the school day. In addition, students indicated that they would be more likely to meditate if it was a mandatory part of the curriculum. Anecdotally, the health

teacher and I observed a distinct change in the classroom environment after a meditation session. There seemed to be a collective serenity, and students appeared to be composed and ready to face the next task. In addition, classroom meditation may support students who were already motivated to meditate and encourage students who claimed to be resistant or hesitant.

Students who identified sub-themes under hesitance may benefit from classroom meditation because it would allow them time to become more comfortable with the practice. Additionally, it would address barriers that many students identified of needing more time or better time management and eliminate the need for notifications or reminders. Some of these barriers may point to the fact that executive functions are still developing, and the skills to manage time, organize, prioritize, and/or set long range goals are still being formed (Arain et al., 2013; Giedd, 2008). Classroom practice may facilitate and support the development of these skills.

As with teaching any new skill, it would be important to have a protocol in place to introduce the skill slowly and reinforce progress. This may help address concerns expressed from students who were reluctant or hesitant to meditate. A possible solution to this could be to implement classroom meditation through a tiered three-phase approach of whole class app meditation, directed classroom practice with the app, and then independent app use in the classroom. The length of time needed to complete each phase may vary and depend on the level of engagement students show. Teachers could assess progress and adjust each of the three phases as needed.

I propose that the teacher first start with a brief introduction to mindfulness meditation and spend a few weeks using the app for whole class meditation. I recommend that the instructor have the class listen to the first session, “What to Expect,” from the *Get Started* series to

familiarize students with the foundations of mindful meditation. The app can easily be amplified through a Bluetooth speaker. After this session, I suggest that the teacher start class once or twice a week by listening to the “One Minute to Mindfulness” session on the app. If possible, I encourage the instructor to direct students to face in the same direction (i.e., not face each other) to eliminate distraction and afford students privacy during this activity. Additionally, the instructor may elect to model the practice and meditate with the class. Since some students suggested that they would like to participate in collaborative activities and share their experience, the teacher could choose to follow this up with a short classroom discussion or written assignment. This activity may have an added benefit of allowing the teacher to gauge student reactions to meditation and provide support as needed.

Second, introduce the app for use in the classroom. An important lesson learned from this study is that it would be prudent to give students ample time to download the app and work out any technical problems prior to use in class. Once all students had access to the app, the instructor could direct students to listen to the same short (i.e., 5 minutes or less) session and write about or discuss their reactions. Classroom practice would support students who reported that they enjoyed and benefited from the experience as well as give time to those who found it difficult to schedule. Classroom meditation could be done once a week for a few weeks. If time allowed, this practice could increase to two or three times a week.

Continuing to include collaborative classroom discussions may elicit the sharing of positive experiences (i.e., increased calm, relaxation, or better focus) that students reported in the survey. This may provide their hesitant or resistant peers with a reason to reconsider their position on meditation and allow them a forum to express and process their own views. In addition, several students indicated that they would like to explore different apps. One

suggestion is that the teacher could assign a class project to investigate and evaluate other apps that offer a free version. Such an assignment may further display respect for student opinions and contribute to engagement for meditation as a stress management strategy. If possible, the teacher could approve appropriate apps for class use.

Third, move from an assigned meditation session to allowing students to choose their own session from the app to complete in the classroom. Allowing students to choose a meditation would display respect for the student. In addition, such activities may not threaten status since they "have" to do it for class and therefore do not need to take direct responsibility for the task. If possible, it would be helpful to allow meditation a few times a week since students may need to use the app for several sessions to experience results, if any. However, the instructor may need to place a time limit (i.e., < 10 minutes) on individual sessions. Once students became accustomed to using the app in class, and possibly realize benefits from the practice, independent use could be assigned as homework, which would eliminate the need for a time limit.

Independent. To introduce independent meditation, I recommend that the instructor have the class listen to the second session "Routine Maker" from the *Get Started* series to provide students with advice on how to incorporate the practice into their daily routine. The information from this session may help students who reported challenges with scheduling and remembering to complete the activity. In the current study, independent meditation with the app allowed students the freedom to choose the meditation session, time, and place. However, participation in the study did require voluntarily meditating four times a week for 8 weeks and completion of a self-report log.

A few students indicated that they did not like filling out the self-report log. This step could be eliminated. This study showed that self-report data was not always accurate and data from the app could be submitted electronically as proof of completion. This modification may result in more accurate reporting since students would not have to be concerned about completing a log post-meditation when they may be in a relaxed state and forget about recording a session. Electronic data would be particularly convenient for students who reported that they used meditation to aid sleep. Additionally, although no student in the present study reported that they experienced adverse effects, the possibility needs to be considered. Students who may not like meditation could be allowed to select other contemplative sessions from the app library and perform a body scan or listen to nature sounds to complete the assignment.

Support

Findings revealed that the majority of students would welcome the opportunity to practice meditation during the school day and recommended that staff should allow app-based meditation. Students reported that they would use the app before a test, quiz, presentation, or during other stressful times. As indicated, education would provide the entire school staff with information to support app-based meditation. Teachers could elect to allow time for students to use the app before stressful assessments, challenging assignments, or as needed. Furthermore, some teachers could decide to allow a brief period of app use at the start of class or as a transition period at the end. The data showed that several students would practice meditation during free time in the school day if it were allowed. Lunchtime was most often cited. As recommended, schoolwide education would equip all staff who worked during lunch periods with information to support use of the app. In addition, there are other times during the school day (e.g., homeroom or break periods) when app-based meditation could be used.

Limitations

This study had several limitations, and these findings are only generalizable to populations the sample represents. This study was approved in the spring of 2019 and student enrollment occurred in the fall of 2019. A new school cell phone policy was adopted over the summer and implemented at the start of the 2019–2020 school year. The prior policy provided vague limits for student use of cell phones and gave instructors the authority to decide on cell phone use in their class or shop. The new policy forbade any use of cell phones during the official school day except for the student's 22-minute lunch period. Fortunately, the administration honored their prior approval, and students who enrolled in this study were allowed to use their phones. While I greatly appreciated being able to conduct the study as designed, the first limitation of this study was that students may have enrolled to gain access to their phones and not because they were interested in stress reduction and/or app-based meditation. The robust enrollment provided a great advantage since I was able to obtain a large data set for analysis of the PSS-10. However, enrolling students who may not have been interested in meditation could have contributed to the low percentage of independent app use by students.

The second limitation was that the implementation of the Stop, Breathe & Think app in the classroom presented several technical challenges. Issues varied from the app company sending duplicate invitations, students receiving broken links to the invitation, individual differences in Apple and Android cell phones, working with differences in operating systems, Wi-Fi dead zones in the classroom, schoolwide Wi-Fi interruptions, student difficulty in accessing Google classroom, and pairing the SBT app with Apple Health or Google Fit. These

issues may have been distracting and prevented students from becoming fully engaged in the app and the meditation sessions.

A third limitation was student absence. Health class was held during the students' shop week, and those who teach health are often made aware of planned absences. For example, students are required to attend OSHA trainings, Skills/USA meetings, PSAT registration, field trips, and/or remain in their shop to finish a task. In addition, there were some occasions when absences were unexpected or due to student illness. Although I attempted to schedule PSS-10 administration and app training around absences, it did affect the PSS-10 data sets, and some students had to "catch-up" during introduction of the app.

Conclusion

By investigating research questions 1 through 5, this study demonstrated that high school students may benefit from reliable stress measurement, and it may be feasible to integrate app-based meditation into the classroom via the suggestions provided. Quantitative results showed that the PSS-10 appears to be a highly reliable instrument for a 24-hour assessment of stress. A major advantage of the PSS-10 is the ease and brevity of administration and scoring. The PSS-10 may be appropriate to implement in high school both as a baseline stress measure, periodic assessment, and to increase awareness of stress level. This may motivate students to use app-based meditation as a stress reduction strategy.

This study was conducted with the hope that any knowledge obtained could potentially contribute to the extant literature on adolescent stress measurement and stress management, benefit student well-being, and provide information that may be useful to a high school community. Although the SBT app was used in this study, the findings may be applied and adapted to other meditation apps that offer free sessions. For example, the Healthy Minds

Program (HMP) is a free app that includes mindfulness meditation (Healthy Minds Innovations, 2022). Interestingly, the PSS-14 was one of the instruments used to evaluate the efficacy of the app (Goldberg, Imhoff-Smith, et al., 2020), and the HMP app uses PSS-10 questions as an optional stress assessment.

Qualitative findings provided evidence that most students benefited from app-based meditation in the classroom. They asserted that it provided a tranquil experience, a welcome break in the school day, and it contributed to their well-being. Most importantly, no student reported that they were harmed. Analysis indicated that many students believed that app-based meditation should be included in the curriculum and allowed at suitable times during the school day. These findings infer that it may be feasible to implement app-based meditation into a high school setting as a beneficial method of stress reduction. Access to accurate instruments that can measure stress levels as well as methods to help manage stress are vital for everyone. Providing high school students with a reliable stress measure, like the PSS-10, and a viable stress reduction method, such as app-based meditation, may benefit them during their development as emerging adults.

References

Adams, Z. W., Sieverdes, J. C., Brunner-Jackson, B., Mueller, M., Chandler, J., Diaz, V., … Treiber, F. A. (2018). Meditation smartphone application effects on prehypertensive adults' blood pressure: Dose-response feasibility trial. *Health Psychology*, *37*(9), 850–860. https://doi.org/10.1037/hea0000584

Alderman, E. M., & Breuner, C. C. (2019). Unique needs of the adolescent. *Pediatrics*, *144*(6), e20193150. https://doi.org/10.1542/peds.2019-3150

Aldridge, V. K., Dovey, T. M., & Wade, A. (2017). Assessing test-retest reliability of psychological measures: Persistent methodological problems. *European Psychologist*, *22*(4), 207–218. https://doi.org/10.1027/1016-9040/a000298

Al-Dubai, S. A. R., Alshagga, M. A., Rampal, K. G., & Sulaiman, N. A. (2012). Factor structure and reliability of the Malay version of the perceived stress scale among Malaysian medical students. *The Malaysian Journal of Medical Sciences: MJMS*, *19*(3), 43–49.

Ali, A. M., Hendawy, A. O., Ahmad, O., Al Sabbah, H., Smail, L., & Kunugi, H. (2021). The Arabic version of the Cohen Perceived Stress Scale: Factorial validity and measurement invariance. *Brain Sciences*, *11*(4), 419. https://doi.org/10.3390/brainsci11040419

Allen, B. P., & Potkay, C. R. (1981). On the arbitrary distinction between states and traits. *Journal of Personality and Social Psychology*, *41*(5), 916–928. https://doi.org/10.1037/0022-3514.41.5.916

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs*, *47*(1), i–171. https://doi.org/10.1037/h0093360

Almadi, T., Cathers, I., Hamdan Mansour, A. M., & Chow, C. M. (2012). An Arabic version of the perceived stress scale: Translation and validation study. *International Journal of Nursing Studies*, *49*(1), 84–89. https://doi.org/10.1016/j.ijnurstu.2011.07.012

American Psychological Association [APA]. (2014). *Stress in America: Are teens adopting adults' stress habits?* Retrieved from https://www.apa.org/news/press/releases/stress/2013/stress-report.pdf

Antonson, C., Thorsen, F., Sundquist, J., & Sundquist, K. (2018). Upper secondary school students' compliance with two Internet-based self-help programmes: A randomised controlled trial. *European Child & Adolescent Psychiatry*, *27*(2), 191–200. https://doi.org/10.1007/s00787-017-1035-6

Arain, M., Haque, M., Johal, L., Mathur, P., Nel, W., Rais, A., … Sharma, S. (2013). Maturation of the adolescent brain. *Neuropsychiatric Disease and Treatment*, *9*, 449–461. http://dx.doi.org/10.2147/NDT.S39776

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, *55*(5), 469–480. https://doi.org/10.1037/0003-066X.55.5.469

Arnett, J. J., Žukauskienė, R., & Sugimura, K. (2014). The new life stage of emerging adulthood at ages 18–29 years: Implications for mental health. *The Lancet Psychiatry*, *1*(7), 569–576. https://doi.org/10.1016/S2215-0366(14)00080-7

Arza, A., Garzón-Rey, J. M., Lázaro, J., Gil, E., Lopez-Anton, R., de la Camara, C., … Aguiló, J. (2019). Measuring acute stress response through physiological signals: Towards a quantitative assessment of stress. *Medical & Biological Engineering & Computing*, *57*(1), 271–287. https://doi.org/10.1007/s11517-018-1879-z

Azzopardi, P., Kennedy, E., & Patton, G. (2017). *Data and indicators to measure adolescent health, social development and well-being* (Innocenti Research Briefs No. 2017/04). https://doi.org/10.18356/4a435ae9-en

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

Barron, R., & Gore, J. S. (2020). The many faces of stress: Three factors that predict physical illness in young adults. *Psychological Reports*, 0033294120936195. https://doi.org/10.1177/0033294120936195

Beauchemin, J., Hutchins, T. L., & Patterson, F. (2008). Mindfulness meditation may lessen anxiety, promote social skills, and improve academic performance among adolescents with learning disabilities. *Complementary Health Practice Review, 13*(1), 34–45. https://doi.org/10.1177/1533210107311624

Ben Loubir, D., Serhier, Z., Battas, O., Agoub, M., & Bennani Othmani, M. (2014). Evaluation of psychometric properties of the Arabic version of PSS stress measuring scale in the Moroccan population. *SAGE Open, 4*(4), 2158244014564353. https://doi.org/10.1177/2158244014564353

Benson, H., & Klipper, M. Z. (2000). *The relaxation response*. New York, NY: Harper Torch.

Benson-Henry Institute. (n.d.). Retrieved January 21, 2021, from https://bensonhenryinstitute.org/

Blakemore, S.-J., & Choudhury, S. (2006). Development of the adolescent brain: Implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry, 47*(3–4), 296–312. https://doi.org/10.1111/j.1469-7610.2006.01611.x

Bland, H. W., Melton, B. F., Welle, P., & Bigham, L. (2012). Stress tolerance: New challenges for millennial college students. *College Student Journal*, *46*(2), 362–375.

Bluth, K., Roberson, P. N. E., & Gaylord, S. A. (2015). A pilot study of a mindfulness intervention for adolescents and the potential role of self-compassion in reducing stress. *EXPLORE*, *11*(4), 292–295. https://doi.org/10.1016/j.explore.2015.04.005

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, *6*. https://doi.org/10.3389/fpubh.2018.00149

Bostock, S., Crosswell, A. D., Prather, A. A., & Steptoe, A. (2019). Mindfulness on-the-go: Effects of a mindfulness meditation app on work stress and well-being. *Journal of Occupational Health Psychology*, *24*(1), 127–138. https://doi.org/10.1037/ocp0000118

Braun, M., Levy, B., Collins, G., & Mogilner, L. (2014). Off the mat: Piloting a mindfulness-based curriculum with adolescents in East Harlem. *International Journal of Child Health and Human Development; Hauppauge*, *7*(3), 325–330.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, *11*(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

Braun, V., & Clarke, V. (2021a). Conceptual and design thinking for thematic analysis. *Qualitative Psychology*, *9*(1), 3–26. https://doi.org/10.1037/qup0000196

Braun, V., & Clarke, V. (2021b). To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qualitative Research in Sport, Exercise and Health, 13*(2), 201–216. https://doi.org/10.1080/2159676X.2019.1704846

Braun, V., Clarke, V., Boulton, E., Davey, L., & McEvoy, C. (2020). The online survey as a qualitative research tool. *International Journal of Social Research Methodology*, *24*(6), 1–14. https://doi.org/10.1080/13645579.2020.1805550

Bree, R. T., & Gallagher, G. (2016). Using Microsoft Excel to code and thematically analyse qualitative data: A simple, cost-effective approach. *All Ireland Journal of Higher Education*, *8*(2). Retrieved from https://ojs.aishe.org/index.php/aishe-j/article/view/281

Byrd, D. R., & McKinney, K. J. (2012). Individual, interpersonal, and institutional level factors associated with the mental health of college students. *Journal of American College Health*, *60*(3), 185–193. https://doi.org/10.1080/07448481.2011.584334

Byrne, D. G., Davenport, S. C., & Mazanov, J. (2007). Profiles of adolescent stress: The development of the adolescent stress questionnaire (ASQ). *Journal of Adolescence*, *30*(3), 393–416. https://doi.org/10.1016/j.adolescence.2006.04.004

Carnegie Mellon University. (2015). *Dr. Cohen's Scales: Laboratory for the study of stress, immunity, and disease*. Retrieved December 19, 2021, from https://www.cmu.edu/dietrich/psychology/stress-immunity-disease-lab/scales/index.html

Carver, C. S., & Connor-Smith, J. (2010). Personality and Coping. *Annual Review of Psychology, 61*(1), 679–704. https://doi.org/10.1146/annurev.psych.093008.100352

Chaaya, M., Osman, H., Naassan, G., & Mahfoud, Z. (2010). Validation of the Arabic version of the Cohen perceived stress scale (PSS-10) among pregnant and postpartum women. *BMC Psychiatry*, *10*. https://doi.org/10.1186/1471-244X-10-111

Champion, L., Economides, M., & Chandler, C. (2018). The efficacy of a brief app-based mindfulness intervention on psychosocial outcomes in healthy adults: A pilot randomised controlled trial. *PLOS ONE*, *13*(12), e0209482. https://doi.org/10.1371/journal.pone.0209482

Childs, E., White, T. L., & de Wit, H. (2014). Personality traits modulate emotional and physiological responses to stress. *Behavioural Pharmacology*, *25*(5–6), 493–502. https://doi.org/10.1097/FBP.0000000000000064

Chiu, Y.-H., Lu, F. J.-H., Lin, J.-H., Nien, C.-L., Hsu, Y.-W., & Liu, H.-Y. (2016). Psychometric properties of the Perceived Stress Scale (PSS): Measurement invariance between athletes and non-athletes and construct validity. *PeerJ*, *4*, e2790. https://doi.org/10.7717/peerj.2790

Christensen, D. S., Dich, N., Flensborg-Madsen, T., Garde, E., Hansen, Å. M., & Mortensen, E. L. (2019). Objective and subjective stress, personality, and allostatic load. *Brain and Behavior*, *9*(9), e01386. https://doi.org/10.1002/brb3.1386

Cohen, S., Doyle, W. J., Skoner, D. P., Fireman, P., Gwaltney, J. M. Jr., & Newsom, J. T. (1995). State and trait negative affect as predictors of objective and subjective symptoms of respiratory viral infections. *Journal of Personality and Social Psychology*, *68*(1), 159–169. https://doi.org/10.1037/0022-3514.68.1.159

Cohen, S., Gianaros, P. J., & Manuck, S. B. (2016). A stage model of stress and disease. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *11*(4), 456–463. https://doi.org/10.1177/1745691616646305

Cohen, S., & Janicki-Deverts, D. (2012). Who's stressed? Distributions of psychological stress in the United States in probability samples from 1983, 2006, and 2009. *Journal of Applied Social Psychology*, *42*(6), 1320–1334. https://doi.org/10.1111/j.1559-1816.2012.00900.x

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, *24*(4), 385–396. https://doi.org/10.2307/2136404

Cohen, S., Tyrrell, D. A., & Smith, A. P. (1991). Psychological stress and susceptibility to the common cold. *The New England Journal of Medicine; Boston*, *325*(9), 606–612. http://dx.doi.org.ezproxyles.flo.org/10.1056/NEJM199108293250903

Cohen, S., & Williamson, G. (1988). Perceived stress in a probability sample of the United States. In S. Spacapan & S. Oskamp (Eds.), *The Social Psychology of Health* (pp. 31–67). Thousand Oaks, CA: SAGE.

Colbert, R. D. (2013). Effect of the transcendental meditation program on graduation, college acceptance and dropout rates for students attending an urban public high school. *Education*, *133*(4), 495–501.

Cole, S. R. (1999). Assessment of differential item functioning in the Perceived Stress Scale-10. *Journal of Epidemiology and Community Health*, *53*(5), 319–320. http://dx.doi.org.ezproxyles.flo.org/10.1136/jech.53.5.319

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Los Angeles, CA: SAGE.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–354. https://doi.org/10.1007/bf02310555

Crosswell, A. D., & Lockwood, K. G. (2020). Best practices for stress measurement: How to measure psychological stress in health research. *Health Psychology Open*, *7*(2), 1–12. https://doi.org/10.1177/2055102920933072

Crowley, C., & Munk, D. (2017). An examination of the impact of a college level meditation course on college student well being. *College Student Journal*, *51*(1), 91–98.

Dahl, R. E. (2004). Adolescent brain development: A period of vulnerabilities and opportunities. Keynote address. *Annals of the New York Academy of Sciences*, *1021*(1), 1–22. https://doi.org/10.1196/annals.1308.001

Dao-Tran, T.-H., Anderson, D., & Seib, C. (2017). The Vietnamese version of the Perceived Stress Scale (PSS-10): Translation equivalence and psychometric properties among older women. *BMC Psychiatry*, *17*. https://doi.org/10.1186/s12888-017-1221-6

Dunning, D. L., Griffiths, K., Kuyken, W., Crane, C., Foulkes, L., Parker, J., & Dalgleish, T. (2019). Research review: The effects of mindfulness-based interventions on cognition and mental health in children and adolescents – a meta-analysis of randomized controlled trials. *Journal of Child Psychology and Psychiatry*, *60*(3), 244–258. https://doi.org/10.1111/jcpp.12980

Dvořáková, K., Kishida, M., Li, J., Elavsky, S., Broderick, P. C., Agrusti, M. R., & Greenberg, M. T. (2017). Promoting healthy transition to college through mindfulness training with first-year college students: Pilot randomized controlled trial. *Journal of American College Health*, *65*(4), 259–267. https://doi.org/10.1080/07448481.2017.1278605

Eagleman, D. (2015). *The brain* (First American edition). New York, NY: Pantheon Books.

Ebstrup, J., Eplov, L., Pisinger, C., & Jørgensen, T. (2011). Association between the Five Factor personality traits and perceived stress: Is the effect mediated by general self-efficacy? *Anxiety, Stress & Coping*, *24*(4), 407–419. https://doi.org/10.1080/10615806.2010.540012

Elder, C., Nidich, S., Colbert, R., Hagelin, J., Grayshield, L., Oviedo-Lim, D., … Gerace, D. (2011). Reduced psychological distress in racial and ethnic minority students practicing the transcendental meditation program. *Journal of Instructional Psychology*, *38*(2), 109–116.

Epel, E. S., Crosswell, A. D., Mayer, S. E., Prather, A. A., Slavich, G. M., Puterman, E., & Mendes, W. B. (2018). More than a feeling: A unified view of stress measurement for population science. *Frontiers in Neuroendocrinology*, *49*, 146–169. https://doi.org/10.1016/j.yfrne.2018.03.001

Erbe, R., & Lohrmann, D. (2015). Mindfulness meditation for adolescent stress and well-being: A systematic review of the literature with implications for school health programs. *Health Educator*, *47*(2), 12–19.

Eva, A. L., & Thayer, N. M. (2017). Learning to BREATHE: A pilot study of a mindfulness-based intervention to support marginalized youth. *Journal of Evidence-Based Complementary & Alternative Medicine*, *22*(4), 580–591. https://doi.org/10.1177/2156587217696928

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Feiss, R., Dolinger, S. B., Merritt, M., Reiche, E., Martin, K., Yanes, J. A., … Pangelinan, M. (2019). A systematic review and meta-analysis of school-based stress, anxiety, and depression prevention programs for adolescents. *Journal of Youth and Adolescence*, *48*(9), 1668–1685. https://doi.org/10.1007/s10964-019-01085-0

Feld, L. D., & Shusterman, A. (2015). Into the pressure cooker: Student stress in college preparatory high schools. *Journal of Adolescence*, *41*, 31–42. https://doi.org/10.1016/j.adolescence.2015.02.003

Figalová, N., & Charvát, M. (2021). The perceived stress scale: Reliability and validity study in the Czech Republic. *Ceskoslovenska Psychologie*, *65*(1), 46–59. https://doi.org/10.51561/cspsych.65.1.46

Flett, J. A. M., Fletcher, B. D., Riordan, B. C., Patterson, T., Hayne, H., & Conner, T. S. (2019). The peril of self-reported adherence in digital interventions: A brief example. *Internet Interventions*, *18*. https://doi.org/10.1016/j.invent.2019.100267

Flett, J. A. M., Hayne, H., Riordan, B. C., Thompson, L. M., & Conner, T. S. (2019). Mobile mindfulness meditation: A randomised controlled trial of the effect of two popular apps on mental health. *Mindfulness*, *10*(5), 863–876. https://doi.org/10.1007/s12671-018-1050-9

Foret, M. M., Scult, M., Wilcher, M., Chudnofsky, R., Malloy, L., Hasheminejad, N., & Park, E. R. (2012). Integrating a relaxation response-based curriculum into a public high school in Massachusetts. *Journal of Adolescence*, *35*(2), 325–332. https://doi.org/10.1016/j.adolescence.2011.08.008

Gál, É., Ștefan, S., & Cristea, I. A. (2020). The efficacy of mindfulness meditation apps in enhancing users' well-being and mental health related outcomes: A meta-analysis of

randomized controlled trials. *Journal of Affective Disorders, 279*, 131–142. https://doi.org/10.1016/j.jad.2020.09.134

Galante, J., Friedrich, C., Dawson, A. F., Modrego-Alarcón, M., Gebbing, P., Delgado-Suárez, I., … Jones, P. B. (2021). Mindfulness-based programmes for mental health promotion in adults in nonclinical settings: A systematic review and meta-analysis of randomised controlled trials. *PLOS Medicine, 18*(1), e1003481. https://doi.org/10.1371/journal.pmed.1003481

Garland, E. L., Gaylord, S. A., & Fredrickson, B. L. (2011). Positive reappraisal mediates the stress-reductive effects of mindfulness: An upward spiral process. *Mindfulness, 2*(1), 59–67. https://doi.org/10.1007/s12671-011-0043-8

Geiser, C., Götz, T., Preckel, F., & Freund, P. A. (2017). States and traits: Theories, models, and assessment. *European Journal of Psychological Assessment, 33*(4), 219–223. https://doi.org/10.1027/1015-5759/a000413

Giedd, J. N. (2008). The teen brain: Insights from neuroimaging. *Journal of Adolescent Health, 42*(4), 335–343. https://doi.org/10.1016/j.jadohealth.2008.01.007

Goldberg, S. B., Imhoff-Smith, T., Bolt, D. M., Wilson-Mendenhall, C. D., Dahl, C. J., Davidson, R. J., & Rosenkranz, M. A. (2020). Testing the efficacy of a multicomponent, self-guided, smartphone-based meditation app: Three-armed randomized controlled trial. *JMIR Mental Health, 7*(11), e23825. https://doi.org/10.2196/23825

Goldberg, S. B., Knoeppel, C., Davidson, R. J., & Flook, L. (2020). Does practice quality mediate the relationship between practice time and outcome in mindfulness-based stress reduction? *Journal of Counseling Psychology, 67*(1), 115–122.

Goleman, D., & Davidson, R. J. (2017). *Altered traits: Science reveals how meditation changes your mind, brain, and body*. New York, NY: Avery.

Goyal, M., Singh, S., Sibinga, E., Gould, N., Rowland-Seymour, A., Sharma, R., … Cramer, H. (2014). Meditation programs for psychological stress and well-being: A systematic review and meta-analysis. *JAMA Internal Medicine*, *174*(3), 357–368. https://doi.org/10.1016/j.dza.2014.07.007

Greeson, J. M., Juberg, M. K., Maytan, M., James, K., & Rogers, H. (2014). A randomized controlled trial of Koru: A mindfulness program for college students and other emerging adults. *Journal of American College Health*, *62*(4), 222–233. https://doi.org/10.1080/07448481.2014.887571

Guthold, R., Carvajal, L., Adebayo, E., Azzopardi, P., Baltag, V., Dastgiri, S., … Requejo, J. (2021). The Importance of Mental Health Measurement to Improve Global Adolescent Health. *Journal of Adolescent Health*. https://doi.org/10.1016/j.jadohealth.2021.03.030

Hansen, T. I., Lehn, H., Evensmoen, H. R., & Håberg, A. K. (2016). Initial assessment of reliability of a self-administered web-based neuropsychological test battery. *Computers in Human Behavior, 63*, 91–97. https://doi.org/10.1016/j.chb.2016.05.025

Healthy Minds Innovations. (2022). *Healthy minds program app*. Retrieved January 20, 2022, from https://hminnovations.org/meditation-app

Helms, J. E., Henze, K. T., Sass, T. L., & Mifsud, V. A. (2006). Treating Cronbach's alpha reliability coefficients as data in counseling research. *The Counseling Psychologist, 34*(5), 630–660. https://doi.org/10.1177/0011000006288308

Holden, R. R. (2000). Psychometrics. In *Encyclopedia of psychology, Vol. 6.* (pp. 417–419). Washington, DC; New York, NY: American Psychological Association. (2004-12704-133). https://doi.org/10.1037/10521-133

Hölzel, B. K., Lazar, S. W., Gard, T., Schuman-Olivier, Z., Vago, D. R., & Ott, U. (2011). How does mindfulness meditation work? Proposing mechanisms of action from a conceptual and neural perspective. *Perspectives on Psychological Science*, *6*(6), 537–559. https://doi.org/10.1177/1745691611419671

Howells, A., Ivtzan, I., & Eiroa-Orosa, F. J. (2016). Putting the 'app' in happiness: A randomised controlled trial of a smartphone-based mindfulness intervention to enhance wellbeing. *Journal of Happiness Studies*, *17*(1), 163–185. https://doi.org/10.1007/s10902-014-9589-1

Huberty, J., Green, J., Glissmann, C., Larkey, L., Puzia, M., & Lee, C. (2019). Efficacy of the mindfulness meditation mobile app "Calm" to reduce stress among college students: Randomized controlled trial. *JMIR MHealth and UHealth*, *7*(6). https://doi.org/10.2196/14273

Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education*, *38*(12), 1217–1218. https://doi.org/10.1111/j.1365-2929.2004.02012.x

Jensen, F. E., & Nutt, A. E. (2015). *The teenage brain: A neuroscientist's survival guide to raising adolescents and young adults* (1st edition.). New York, NY: Harper.

Jiang, J. M., Seng, E. K., Zimmerman, M. E., Sliwinski, M., Kim, M., & Lipton, R. B. (2017). Evaluation of the reliability, validity, and predictive validity of the subscales of the Perceived Stress Scale in older adults. *Journal of Alzheimer's Disease*, *59*(3), 987–996. https://doi.org/10.3233/JAD-170289

Kabat-Zinn, J. (2006). Mindfulness-based interventions in context: Past, present, and future. *Clinical Psychology: Science and Practice, 10*(2), 144–156. https://doi.org/10.1093/clipsy.bpg016

Kabat-Zinn, J. (2013). *Full catastrophe living: Using the wisdom of your body and mind to face stress, pain, and illness* (Revised, Updated Edition). New York, NY: Bantam.

Kadison, R., & DiGeronimo, T. F. (2005). *College of the overwhelmed: The campus mental health crisis and what we must do about it*. San Francisco, CA: Jossey-Bass.

Kang, Y. S., Choi, S. Y., & Ryu, E. (2009). The effectiveness of a stress coping program based on mindfulness meditation on the stress, anxiety, and depression experienced by nursing students in Korea. *Nurse Education Today*, *29*(5), 538–543. https://doi.org/10.1016/j.nedt.2008.12.003

Kass, J. D. (2017). *A person-centered approach to psychospiritual maturation: Mentoring psychological resilience and inclusive community in higher education*. Cham, Switzerland: Palgrave Macmillan.

Kass, J. D., & Trantham, S. M. (2013). Perspectives from clinical neuroscience: Mindfulness and therapeutic use of the arts. In L. Rappaport (Ed.), *Mindfulness and the arts therapies: Theory and practice* (pp. 288–315). London, UK: Jessica Kingsley Publishers.

Khalili, R., Sirati nir, M., Ebadi, A., Tavallai, A., & Habibi, M. (2017). Validity and reliability of the Cohen 10-item Perceived Stress Scale in patients with chronic headache: Persian version. *Asian Journal of Psychiatry*, *26*, 136–140. https://doi.org/10.1016/j.ajp.2017.01.010

Kohn, P. M., & Milrose, J. A. (1993). The inventory of high-school students' recent life experience. *Journal of Youth and Adolescence; New York*, *22*(1), 43–55.

Konaszewski, K., Niesiobędzka, M., & Surzykiewicz, J. (2021). Resilience and mental health among juveniles: Role of strategies for coping with stress. *Health and Quality of Life Outcomes, 19,* 58. https://doi.org/10.1186/s12955-021-01701-3

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Krapić, N., Hudek-Knežević, J., & Kardum, I. (2015). Stress in adolescence: Effects on development. In J. D. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 562–569). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.23031-6

Lau, N., O'Daffer, A., Colt, S., Yi-Frazier, J. P., Palermo, T. M., McCauley, E., & Rosenberg, A. R. (2020). Android and iPhone mobile apps for psychosocial wellness and stress management: Systematic search in app stores and literature review. *JMIR MHealth and UHealth, 8*(5), e17798. https://doi.org/10.2196/17798

Lazarus, R. S. (2006). *Stress and emotion: A new synthesis*. New York, NY: Springer.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York, NY: Springer.

Lee, E.-H. (2012). Review of the psychometric evidence of the Perceived Stress Scale. *Asian Nursing Research; Seoul, 6*(4), 121–127. http://dx.doi.org.ezproxyles.flo.org/10.1016/j.anr.2012.08.004

Lemon, J. C., & Watson, J. C. (2011). Early identification of potential high school dropouts: An investigation of the relationship among at-risk status, wellness, perceived stress, and mattering. *Journal of At-Risk Issues, 16*(2), 7.

Leung, D. Y., Lam, T., & Chan, S. S. (2010). Three versions of Perceived Stress Scale: Validation in a sample of Chinese cardiac patients who smoke. *BMC Public Health*, *10*(1). https://doi.org/10.1186/1471-2458-10-513

Levitt, H. M., Motulsky, S. L., Wertz, F. J., Morrow, S. L., & Ponterotto, J. G. (2017). Recommendations for designing and reviewing qualitative research in psychology: Promoting methodological integrity. *Qualitative Psychology*, *4*(1), 2–22. https://doi.org/10.1037/qup0000082

Li, P., & Ramirez, D. R. (2017). Zen and psychotherapy. In A. Masuda & W. T. O'Donohue (Eds.), *Handbook of Zen, mindfulness, and behavioral health* (pp. 169–194). Cham, Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-319-54595-0_15

Linardon, J., & Fuller-Tyszkiewicz, M. (2020). Attrition and adherence in smartphone-delivered interventions for mental health problems: A systematic and meta-analytic review. *Journal of Consulting and Clinical Psychology*, *88*(1), 1–13. https://doi.org/10.1037/ccp0000459

Liston, C., McEwen, B. S., & Casey, B. J. (2009). Psychosocial stress reversibly disrupts prefrontal processing and attentional control. *Proceedings of the National Academy of Sciences*, *106*(3), 912–917. https://doi.org/10.1073/pnas.0807041106

Liu, J., Tang, W., Chen, G., Lu, Y., Feng, C., & Tu, X. M. (2016). Correlation and agreement: Overview and clarification of competing concepts and measures. *Shanghai Archives of Psychiatry*, *28*(2), 115–120. https://doi.org/10.11919/j.issn.1002-0829.216045

Liu, X., Zhao, Y., Li, J., Dai, J., Wang, X., & Wang, S. (2020). Factor structure of the 10-item perceived stress scale and measurement invariance across genders among Chinese adolescents. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.00537

Loi, C. X. A., Spencer, N., & Willam, A. (2008). Stress management with university students in Canada, US, and Germany: A multinational health education analysis. *HEE-Journal - The Journal of Health, Environment, & Education*, *1*, 60–67.

Lu, W., Bian, Q., Wang, W., Wu, X., Wang, Z., & Zhao, M. (2017). Chinese version of the Perceived Stress Scale-10: A psychometric study in Chinese university students. *PLoS ONE*, *12*(12). https://doi.org/10.1371/journal.pone.0189543

Luo, J., Derringer, J., Briley, D. A., Roberts, B. W., & Mõttus, R. (2017). Genetic and Environmental Pathways Underlying Personality Traits and Perceived Stress: Concurrent and Longitudinal Twin Studies. *European Journal of Personality*, *31*(6), 614–629. https://doi.org/10.1002/per.2127

Makhubela, M. (2020). Assessing psychological stress in South African university students: Measurement validity of the Perceived Stress Scale (PSS-10) in diverse populations. *Current Psychology*. https://doi.org/10.1007/s12144-020-00784-3

Mani, M., Kavanagh, D. J., Hides, L., & Stoyanov, S. R. (2015). Review and evaluation of mindfulness-based iPhone apps. *JMIR MHealth and UHealth*, *3*(3), 1–10. https://doi.org/10.2196/mhealth.4328

Massachusetts Department of Elementary and Secondary Education. (2018). *Strategic Plan*. Massachusetts Department of Elementary and Secondary Education. Retrieved from http://www.doe.mass.edu/research/strategicplan.docx

Matheson, G. J. (2019). We need to talk about reliability: Making better use of test-retest studies for study design and interpretation. *PeerJ*, *7*, e6918. https://doi.org/10.7717/peerj.6918

Mayor, E. (2015). Gender roles and traits in stress and health. *Frontiers in Psychology*, *6*, 1–7. https://doi.org/10.3389/fpsyg.2015.00779

McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review*, *15*(1), 28–50. https://doi.org/10.1177/1088868310366253

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30–46. https://doi.org/10.1037/1082-989X.1.1.30

McKay, M., Andretta, J., & Perry, J. (2019). The shortened version of the Adolescent Stress Questionnaire (ASQ-S; Sweden): A validation study in United Kingdom adolescents. *Scandinavian Journal of Child and Adolescent Psychiatry and Psychology*, *7*, 81–87. https://doi.org/10.21307/sjcapp-2019-011

Medvedev, O. N., Krägeloh, C. U., Narayanan, A., & Siegert, R. J. (2017). Measuring mindfulness: Applying generalizability theory to distinguish between state and trait. *Mindfulness*, *8*(4), 1036–1046. https://doi.org/10.1007/s12671-017-0679-0

Mendelson, T., Greenberg, M. T., Dariotis, J. K., Gould, L. F., Rhoades, B. L., & Leaf, P. J. (2010). Feasibility and preliminary outcomes of a school-based mindfulness intervention for urban youth. *Journal of Abnormal Child Psychology*, *38*(7), 985–994. http://dx.doi.org.ezproxyles.flo.org/10.1007/s10802-010-9418-x

Metz, S. M., Frank, J. L., Reibel, D., Cantrell, T., Sanders, R., & Broderick, P. C. (2013). The effectiveness of the learning to BREATHE program on adolescent emotion regulation. *Research in Human Development*, *10*(3), 252–272. https://doi.org/10.1080/15427609.2013.818488

Michaelides, M. P., Christodoulou, A., Kkeli, N., Karekla, M., & Panayiotou, G. (2016). Factorial structure of the perceived stress scale and implications for scoring. *European*

Review of Applied Psychology, *66*(6), 309–316. https://doi.org/10.1016/j.erap.2016.07.002

Miller, T., Chandler, L., & Mouttapa, M. (2015). A needs assessment, development, and formative evaluation of a health promotion smartphone application for college students. *American Journal of Health Education*, *46*(4), 207–215. https://doi.org/10.1080/19325037.2015.1044138

Miller, Y. R., Medvedev, O. N., Hwang, Y.-S., & Singh, N. N. (2020). Applying generalizability theory to the Perceived Stress Scale to evaluate stable and dynamic aspects of educators' stress. *International Journal of Stress Management*. https://doi.org/10.1037/str0000207

Mircioiu, C., & Atkinson, J. (2017). A Comparison of Parametric and Non-Parametric Methods Applied to a Likert Scale. *Pharmacy*, *5*(4), 26. https://doi.org/10.3390/pharmacy5020026

Mozumder, M. K. (2017). Validation of Bengali perceived stress scale among LGBT population. *BMC Psychiatry*, *17*(1), 314. https://doi.org/10.1186/s12888-017-1482-0

Mrazek, A. J., Mrazek, M. D., Cherolini, C. M., Cloughesy, J. N., Cynman, D. J., Gougis, L. J., … Schooler, J. W. (2019). The future of mindfulness training is digital, and the future is now. *Current Opinion in Psychology*, *28*, 81–86. https://doi.org/10.1016/j.copsyc.2018.11.012

O'Connor, D. B., Thayer, J. F., & Vedhara, K. (2021). Stress and health: A review of psychobiological processes. *Annual Review of Psychology*, *72*(1), 663–688. https://doi.org/10.1146/annurev-psych-062520-122331

Oman, D., Shapiro, S. L., Thoresen, C. E., Plante, T. G., & Flinders, T. (2008). Meditation lowers stress and supports forgiveness among college students: A randomized controlled

trial. *Journal of American College Health*, *56*(5), 569–578. https://doi.org/10.3200/JACH.56.5.569-578

Patton, G. C., Sawyer, S. M., Santelli, J. S., Ross, D. A., Afifi, R., Allen, N. B., … Viner, R. M. (2016). Our future: A Lancet commission on adolescent health and wellbeing. *The Lancet; London*, *387*(10036), 2423–2478. http://dx.doi.org.ezproxyles.flo.org/10.1016/S0140-6736(16)00579-1

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed). Thousand Oaks, Calif: Sage Publications.

Perinetti, G. (2018). StaTips Part IV: Selection, interpretation and reporting of the intraclass correlation coefficient. *South European Journal of Orthodontics and Dentofacial Research*, *5*. https://doi.org/10.5937/sejodr5-17434

Piekarska, J. (2020). Determinants of perceived stress in adolescence: The role of personality traits, emotional abilities, trait emotional intelligence, self-efficacy, and self-esteem. *Advances in Cognitive Psychology*, *16*(4), 309–320. https://doi.org/10.5709/acp-0305-z

Psihogios, A. M., Stiles-Shields, C., & Neary, M. (2020). The needle in the haystack: Identifying credible mobile health apps for pediatric populations during a pandemic and beyond. *Journal of Pediatric Psychology*, *45*(10), 1106–1113. https://doi.org/10.1093/jpepsy/jsaa094

Qin, S., Nelson, L., Mcleod, L., Eremenco, S., & Coons, S. J. (2019). Assessing test–retest reliability of patient-reported outcome measures using intraclass correlation coefficients: Recommendations for selecting and documenting the analytical formula. *Quality of Life Research*, *28*(4), 1029–1033. https://doi.org/10.1007/s11136-018-2076-0

Quach, D., Gibler, R. C., & Jastrowski Mano, K. E. (2017). Does home practice compliance make a difference in the effectiveness of mindfulness interventions for adolescents? *Mindfulness*, *8*(2), 495–504. https://doi.org/10.1007/s12671-016-0624-7

Ramasubramanian, S. (2017). Mindfulness, stress coping and everyday resilience among emerging youth in a university setting: A mixed methods approach. *International Journal of Adolescence and Youth*, *22*(3), 308–321. https://doi.org/10.1080/02673843.2016.1175361

Remor, E. (2006). Psychometric properties of a European Spanish version of the Perceived Stress Scale (PSS). *The Spanish Journal of Psychology*, *9*(1), 86–93. https://doi.org/10.1017/S1138741600006004

Rhode, P. C., Froehlich-Grobe, K., Hockemeyer, J. R., Carlson, J. A., & Lee, J. (2012). Assessing stress in disability: Developing and piloting the disability related stress scale. *Disability and Health Journal*, *5*(3), 168–176. https://doi.org/10.1016/j.dhjo.2012.03.002

Rideout, V., Fox, S., & Trust, W. B. (2018). Digital health practices, social media use, and mental well-being among teens and young adults in the U.S. *Articles, Abstracts, and Reports*, *1093*, 96.

Riley, T. J., Kirsch, A. C., Shapiro, J. B., & Conley, C. S. (2016). Examining stress and coping as a mediator for internalizing symptomatology: A comparison between sexual minority and majority first-year college students. *Journal of Adolescence*, *49*, 124–133. https://doi.org/10.1016/j.adolescence.2016.03.005

Roberti, J. W., Harrington, L. N., & Storch, E. A. (2006). Further psychometric support for the 10-item version of the perceived stress scale. *Journal of College Counseling*, *9*(2). Retrieved from

http://ezproxyles.flo.org/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ750042&site=eds-live&scope=site

Roberts, B. W. (2018). A revised sociogenomic model of personality traits. *Journal of Personality*, *86*(1), 23–35. https://doi.org/10.1111/jopy.12323

Robinson, A. M. (2018). Let's talk about stress: History of stress research. *Review of General Psychology*, *22*(3), 334–342. https://doi.org/10.1037/gpr0000137

Roeser, R. W., & Pinela, C. (2014). Mindfulness and compassion training in adolescence: A developmental contemplative science perspective. *New Directions for Youth Development*, *2014*(142), 9–30. https://doi.org/10.1002/yd.20094

Rogers, H. B. (2013). Mindfulness meditation for increasing resilience in college students. *Psychiatric Annals; Thorofare*, *43*(12), 545–548. http://dx.doi.org.ezproxyles.flo.org/10.3928/00485713-20131206-06

Rubin, M. (2017). Do *p* values lose their meaning in exploratory analyses? It depends how you define the familywise error rate. *Review of General Psychology*, *21*(3), 269–275. https://doi.org/10.1037/gpr0000123

Sainani, K. L. (2017). Reliability Statistics. *PM&R*, *9*(6), 622–628. https://doi.org/10.1016/j.pmrj.2017.05.001

Sandhu, S. S., Ismail, N. H., & Rampal, K. G. (2015). The Malay version of the Perceived Stress Scale (PSS)-10 is a reliable and valid measure for stress among nurses in Malaysia. *The Malaysian Journal of Medical Sciences*, *22*(6), 26–31.

Sapolsky, R. M. (2004). *Why zebras don't get ulcers.* New York, NY: Owl Book/Henry Holt and Co.

Sapolsky, R. M. (2017). *Behave: The biology of humans at our best and worst*. New York, NY: Penguin Press.

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, *44*(4), 695–729. https://doi.org/10.1177/0539018405058216

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia*, *126*(5), 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Sedlmeier, P., Eberth, J., Schwarz, M., Zimmermann, D., Haarig, F., Jaeger, S., & Kunze, S. (2012). The psychological effects of meditation: A meta-analysis. *Psychological Bulletin*, *138*, 1139–1171. https://doi.org/10.1037/a0028168

Seiffge-Krenke, I., Aunola, K., & Nurmi, J.-E. (2009). Changes in stress perception and coping during adolescence: The role of situational and personal factors. *Child Development, 80*(1), 259–279. (2009-01908-020). https://doi.org/10.1111/j.1467-8624.2008.01258.x

Selye, H. (1978). *The stress of life* (Rev. ed.). New York, NY: McGraw-Hill Education.

Shapiro, S. L., Brown, K. W., & Astin, J. (2011). Toward the integration of meditation into higher education: A review of research evidence. *Teachers College Record*, *113*(3), 493–528.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Siegel, D. J. (2012). *Pocket guide to interpersonal neurobiology: An integrative handbook of the mind* (1st. ed.). New York, NY: W.W. Norton & Co.

Siegel, D. J. (2014). *Brainstorm: The power and purpose of the teenage brain*. New York, NY: Penguin Publishing Group. Retrieved from http://ebook.3m.com/library/neworleanspubliclibrary-document_id-dxz3ug9

Singh, A. A., & McKleroy, V. S. (2011). "Just getting out of bed is a revolutionary act": The resilience of transgender people of color who have survived traumatic life events. *Traumatology*, *17*(2), 34–44. https://doi.org/10.1177/1534765610369261

Siqueira, L., Diab, M., Bodian, C., & Rolnitzky, L. (2000). Adolescents becoming smokers: The roles of stress and coping methods. *Journal of Adolescent Health*, *27*(6), 399–408. https://doi.org/10.1016/S1054-139X(00)00167-1

Siqueira Reis, R., Ferreira Hino, A. A., & Romélio Rodriguez Añez, C. (2010). Perceived Stress Scale: Reliability and validity study in Brazil. *Journal of Health Psychology*, *15*(1), 107–114. https://doi.org/10.1177/1359105309346343

Stop, Breathe & Think. (2019). Stop, Breathe & Think: Guided meditations & mindfulness. Retrieved April 17, 2019, from Stop, Breathe & Think website: https://www.stopbreathethink.com/

Sun, Y., Gao, L., Kan, Y., & Shi, B.-X. (2019). The Perceived Stress Scale-10 (PSS-10) is reliable and has construct validity in Chinese patients with systemic lupus erythematosus. *Lupus*, *28*(2), 149–155. https://doi.org/10.1177/0961203318815595

Tan, S. Y., & Yip, A. (2018). Hans Selye (1907–1982): Founder of the stress theory. *Singapore Medical Journal*, *59*(4), 170–171. https://doi.org/10.11622/smedj.2018043

Tang, Y.-Y., Hölzel, B. K., & Posner, M. I. (2015). The neuroscience of mindfulness meditation. *Nature Reviews Neuroscience*, *16*(4), 213–225. https://doi.org/10.1038/nrn3916

Taylor, J. M. (2015). Psychometric analysis of the Ten-Item Perceived Stress Scale. *Psychological Assessment*, *27*(1), 90–101. https://doi.org/10.1037/a0038100

Twenge, J. M., Cooper, A. B., Joiner, T. E., Duffy, M. E., & Binau, S. G. (2019). Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005–2017. *Journal of Abnormal Psychology*, *128*(3), 185–199. https://doi.org/10.1037/abn0000410

van Loon, A. W. G., Creemers, H. E., Beumer, W. Y., Okorn, A., Vogelaar, S., Saab, N., … Asscher, J. J. (2020). Can schools reduce adolescent psychological stress? A multilevel meta-analysis of the effectiveness of school-based intervention programs. *Journal of Youth and Adolescence*, *49*(6), 1127–1145. https://doi.org/10.1007/s10964-020-01201-5

Varpio, L., Ajjawi, R., Monrouxe, L. V., O'Brien, B. C., & Rees, C. E. (2017). Shedding the cobra effect: Problematising thematic emergence, triangulation, saturation and member checking. *Medical Education*, *51*(1), 40–50. https://doi.org/10.1111/medu.13124

Vaz, S., Falkmer, T., Passmore, A. E., Parsons, R., & Andreou, P. (2013). The case for using the repeatability coefficient when calculating test–retest reliability. *PLoS ONE*, *8*(9), e73990. https://doi.org/10.1371/journal.pone.0073990

Vetter, T. R., & Schober, P. (2018). Agreement analysis: What he said, she said versus you said. *Anesthesia & Analgesia*, *126*(6), 2123–2128. https://doi.org/10.1213/ANE.0000000000002924

Vettese, L. C., Toneatto, T., Stea, J. N., Nguyen, L., & Wang, J. J. (2009). Do mindfulness meditation participants do their homework? And does it make a difference? A review of the empirical evidence. *Journal of Cognitive Psychotherapy, 23*(3), 198–225. https://doi.org/10.1891/0889-8391.23.3.198

Wang, Z., Chen, J., Boyd, J. E., Zhang, H., Jia, X., Qiu, J., & Xiao, Z. (2011). Psychometric properties of the Chinese version of the Perceived Stress Scale in policewomen. *PLoS ONE*, *6*(12). Retrieved from https://ezproxyles.flo.org/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2011-29231-001&site=ehost-live&scope=site

Warnecke, E., Quinn, S., Ogden, K., Towle, N., & Nelson, M. R. (2011). A randomised controlled trial of the effects of mindfulness practice on medical student stress levels: Effects of mindfulness practice on student stress levels. *Medical Education*, *45*(4), 381–388. https://doi.org/10.1111/j.1365-2923.2010.03877.x

Watson, D. (1988). Intraindividual and interindividual analyses of positive and negative affect: Their relation to health complaints, perceived stress, and daily activities. *Journal of Personality & Social Psychology*, *54*(6), 1020–1030. https://doi.org/10.1037/0022-3514.54.6.1020

Weber, S., Lorenz, C., & Hemmings, N. (2019). Improving stress and positive mental health at work via an app-based intervention: A large-scale multi-center randomized control trial. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02745

Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength and Conditioning Research*, *19*(1), 231–240. http://dx.doi.org.ezproxyles.flo.org/10.1519/15184.1

Wertz, F. J. (Ed.). (2011). *Five ways of doing qualitative analysis: Phenomenological psychology, grounded theory, discourse analysis, narrative research, and intuitive inquiry*. New York: Guilford Press.

Wisner, B. L., Jones, B., & Gwin, D. (2010). School-based meditation practices for adolescents: A resource for strengthening self-regulation, emotional coping, and self-esteem. *Children & Schools*, *32*(3), 150–159. https://doi.org/10.1093/cs/32.3.150

Wongpakaran, N., & Wongpakaran, T. (2010). The Thai version of the PSS-10: An Investigation of its psychometric properties. *BioPsychoSocial Medicine*, 6. https://doi.org/10.1186/1751-0759-4-6

Wu, J., Tong, H., Liu, Z., Tao, J., Chen, L., Chan, C. C. H., & Lee, T. M. C. (2021). Neurobiological effects of perceived stress are different between adolescents and middle-aged adults. *Brain Imaging and Behavior*, *15*(2), 846–854. https://doi.org/10.1007/s11682-020-00294-7

Yeager, D. S., Dahl, R. E., & Dweck, C. S. (2018). Why interventions to influence adolescent behavior often fail but could succeed. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *13*(1), 101–122. https://doi.org/10.1177/1745691617722620

Yen, M., & Lo, L.-H. (2002). Examining test-retest reliability. *Nursing Research*, *51*(1), 59–62.

You, M., Laborde, S., Dosseville, F., Salinas, A., & Allen, M. S. (2020). Associations of chronotype, Big Five, and emotional competences with perceived stress in university students. *Chronobiology International*, *37*(7), 1090–1098. https://doi.org/10.1080/07420528.2020.1752705

Appendix A

Test-Retest Reliability Studies

Reference	PSS version	Language/ Country	Interval[a]	Population/Age	Coefficient[b]	Factor(s)
(Cohen et al., 1983)	PSS-14	English/ USA	2-days	Two student groups Mean =19.01 Mean = 20.75	*r* = .85	One
			6-weeks	Adults (smoking cessation group) Mean =38.4	*r* = .55	One
Current Study	PSS-10	English/ USA	1-day	Students Mean =16.77	ICC = .92 *r* = .92	One
			6-weeks		ICC = .69 *r* = .69	One
(Siqueira Reis et al., 2010)	PSS-10	Portuguese/ Brazil	7-days	University teachers Mean = 45.5	ICC = .83	Factor 1
					ICC = .68	Factor 2
					ICC = .86	Composite
(Sandhu et al., 2015)	PSS-10	Malay/ Malaysia	7-days	Female Nurses Mean = 48.3	ICC = .81	One
(Sun et al., 2019)	PSS-10	Simplified Chinese/ China	7-days	Adults with lupus Median = 49	ICC = .82	Factor 1
					ICC = .99	Factor 2
					ICC = .95	Composite

Reference	PSS version	Language/ Country	Interval[a]	Population/Age	Coefficient[b]	Factor(s)
(Ben Loubir et al., 2014)	PSS-10	Arabic/ Morocco	1-week	Adults (Moroccan locals who speak Arabic) Mean = 32.2	ICC = .91	One
(Chaaya et al., 2010)	PSS-10	Arabic/ Lebanon	1-week	Female students Mean = 22.5	[c]r_s = .79	One
			2-3 weeks	Pregnant women Mean = 28.4	r_s = .63	One
			2-3 weeks	Postpartum women Mean = 29.7	r_s = .63	One
(Chiu et al., 2016)	PSS-10	Chinese/ Taiwan	8-9 days	Student athletes Mean = 20.08	r = .66 r = .50	Factor 1 Factor 2
(Remor, 2006)	PSS-10	European Spanish/ Spain	2-weeks	University students Mean = 26.9	r = .77	One

Reference	PSS version	Language/ Country	Interval[a]	Population/Age	Coefficient[b]	Factor(s)
(Wang et al., 2011)	PSS-10	Simplified Chinese/ China	2-weeks	Adult policewomen Mean = 21.1	r_s = .72 r_s = .63 r_s = .68	Factor 1 Factor 2 Composite
(Lu et al., 2017)	PSS-10	Simplified Chinese/ China	2-weeks	Chinese University students Mean = 18.3	*r* = .70	One
(Khalili et al., 2017)	PSS-10	Persian/ Iran	[d]2-weeks	Adults (patients admitted to pain clinic for chronic headache) Mean = 38	ICC = .95 ICC = .90 ICC = .93	Factor 1 Factor 2 Composite
(Figalova & Charvat 2021)	PSS-10	Czech/ Czech Republic	2-weeks	Adults (general adult population) Mean = 44.32	*r* = .88	One

Reference	PSS version	Language/ Country	Interval[a]	Population/Age	Coefficient[b]	Factor(s)
(Al-Dubai et al., 2012)	PSS-10	Malay/ Malaysia	3-weeks	Bachelor of Medical Science students Mean = 20.9	ICC = .82	One
Current Study	PSS-10	English/ USA	3-weeks	Students Mean=16.77	ICC = .63 r = .63	One
(Wongpakaran & Wongpakaran, 2010	PSS-10	Thai/ Thailand	4-weeks	Medical students (years 1 – 5) Mean = 20.84	ICC = .83	One
Dao-Tran et al., 2017)	PSS-10	Vietnamese/ Vietnam	1-month	Vietnamese women > 60 Median = 68	r_s = .43	One

Note. [a] As indicated in study
[b] r = Pearson product-moment correlation
ICC = Intraclass correlation coefficient
r_s = Spearman's rank-order correlation
[c] Chaaya et al. (2010) noted reliability may have been higher among students compared to the other two groups due to the difference in time to retest.
[d] Not reported in study, but through personal correspondence, Dr. Khalili confirmed a 2-week interval

Appendix B

Perceived Stress Scale

PERCEIVED STRESS SCALE

The questions in this scale ask you about your feelings and thoughts during the last month. **In each case, you will be asked to indicate by circling *how often* you felt or thought a certain way.**

Name _____ Date _____

Age _____ Gender (*Circle*): **M F** Other _____

	0 = Never	**1 = Almost Never**	**2 = Sometimes**	**3 = Fairly Often**	**4 = Very Often**
1. In the last month, how often have you been upset because of something that happened unexpectedly?	0	1	2	3	4
2. In the last month, how often have you felt that you were unable to control the important things in your life?	0	1	2	3	4
3. In the last month, how often have you felt nervous and "stressed"?	0	1	2	3	4
4. In the last month, how often have you felt confident about your ability to handle your personal problems?	0	1	2	3	4
5. In the last month, how often have you felt that things were going your way?	0	1	2	3	4
6. In the last month, how often have you found that you could not cope with all the things that you had to do?	0	1	2	3	4
7. In the last month, how often have you been able to control irritations in your life?	0	1	2	3	4
8. In the last month, how often have you felt that you were on top of things?	0	1	2	3	4
9. In the last month, how often have you been angered because of things that were outside of your control?	0	1	2	3	4
10. In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?	0	1	2	3	4

info@mindgarden.com
www.mindgarden.com

References

The PSS Scale is reprinted with permission of the American Sociological Association, from Cohen, S., Kamarck, T., and Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24,* 386-396.
Cohen, S. and Williamson, G. Perceived Stress in a Probability Sample of the United States. Spacapan, S. and Oskamp, S. (Eds.) *The Social Psychology of Health.* Newbury Park, CA: Sage, 1988.

Appendix C

Student Informed Assent - Group 1

Title: The Effect of Mindfulness Meditation Using a Smartphone app on Stress in High School Students

Researcher: Mrs. Eileen Kaskons, a doctoral student at Lesley University

Description and Purpose: You are being asked to volunteer for this research study during your Health class. The purpose of this study is to investigate the effect of the Stop, Breathe & Think (SB&T) mindfulness app on student stress levels. SB&T (https://www.stopbreathethink.com/) includes lessons in nonreligious mindfulness meditation.

Procedures: This study will be conducted at Greater Lowell Technical High School from September to November 2019 with the researcher and health teacher as part of your Health class. The entire study will take ten weeks. By signing this form, you agree to:

The First Week of the Study:

- Complete a short ten-question Perceived Stress Scale (PSS-10) assessment two times.

During the Study:

In class:

- Bring smartphone and headphones to class. (If you do not have equipment, an iPad or headphones will be provided.)
- Download the Stop, Breathe & Think (SB&T) app to your smartphone.
- Listen to the first four of ten introductory sessions.

Outside of class:

- Finish the remaining six introductory sessions on your own.
- Listen to one mindfulness session four times a week for eight weeks. You can choose from sessions that are one to nine minutes long.
- Log the session name and length on an electronic document. Show your "My Progress" screen on the SB&T app to the researcher at the end of each week in Health class.

The Last Week of the Study:

- After using the app for eight weeks, you will repeat the PSS-10 assessment.

Risks: Participation in this research is voluntary. Even if your parent/guardian has given consent, **you have the right to refuse to be in this study**. If you decide to be in the study and change your mind, you have the right to withdraw at any time. You may skip questions and/or choose not to complete mindfulness sessions. Your decision to participate, not participate, or withdraw will not affect your grade in Health class or result in any negative consequences.

Participation in this research may lead to a reduction in stress and poses minimal risk but may not provide any benefit to you. In the event of any discomfort during the study, you will be referred to your school counselor.

Confidentiality, Privacy and Anonymity: All student information and records will be kept private and confidential **to the extent allowed by law.** The researcher will use a subject number instead of your name on all forms, documents, and future presentations or publications. Once the study is over, all electronic data and paper documents will be destroyed. The researcher may present the outcomes of this study for academic purposes but will never reveal your name or any facts that might identify you.

If at any time you have questions or problems about this study, you can contact the researcher, Mrs. Eileen Kaskons (978-244-0791 or ekaskons@lesley.edu), or the Lesley University sponsoring faculty, Dr. Adam Meiselman (617-349-8477 or ameiselm@lesley.edu). We will give you a copy of this assent form to keep.

Signatures and Names:

_____	_____	_____
Date	Researcher's Signature	Print Name
_____	_____	_____
Date	Your Signature	Print Name

There is a standing committee for Human Subjects in Research at Lesley University to which complaints or problems concerning any research project may, and should, be reported if they arise. Contact the Committee Chairpersons at irb@lesley.edu.

Parent/Guardian Informed Consent - Group 1

Title: The Effect of Mindfulness Meditation Using a Smartphone app on Stress in High School Students

Researcher: Eileen Kaskons, a doctoral student at Lesley University.

Description and Purpose: Your child is being asked to volunteer for this research study during their Health class. The purpose of this study is to investigate the effect of the Stop, Breathe & Think (SB&T) mindfulness app on student stress levels. SB&T (https://www.stopbreathethink.com/) includes lessons in nonreligious mindfulness meditation.

Procedures: This study will be conducted at Greater Lowell Technical High School from September to November 2019 with the researcher and health teacher as part of your child's Health class. The entire study will take ten weeks. By signing this form, you agree that your child can participate in the following activities:

The First Week of the Study:

- Complete a short ten-question Perceived Stress Scale (PSS-10) assessment two times.

During the Study:

In class:

- Bring smartphone and headphones to class. (If your child does not have equipment, an iPad or headphones will be provided.)
- Download the Stop, Breathe & Think (SB&T) app to their smartphone.
- Listen to the first four of ten introductory sessions.

Outside of class:

- Finish the remaining six introductory sessions on their own.
- Listen to one mindfulness session four times a week for eight weeks. Your child can choose from sessions that are one to nine minutes long.
- Log the session name and length on an electronic document. Show their "My Progress" screen on the SB&T app to the researcher at the end of each week in Health class.

The Last Week of the Study:

- After using the app for eight weeks, your child will repeat the PSS-10 assessment.

Risks: Participation in this research is voluntary. Your child has the right to refuse to be in this study. If your child decides to be in the study, your child has the right to change their mind and withdraw at any time. Your child may skip questions and/or choose not to complete mindfulness sessions. All of your child's questions will be answered at any time. Your child's decision to participate, not participate, or withdraw will not affect their grade in Health class or result in any negative consequences.

Participation in this research may lead to a reduction in stress and poses minimal risk. This research may not provide any benefit to your child. Parents and guardians are encouraged to discuss this study with their child and report any concerns to the researcher. In the event of any discomfort during the study, your child will be referred to their school counselor.

Confidentiality, Privacy and Anonymity: We will use a subject number instead of your child's name on all forms and documents. Electronic data will be kept confidential and secured on a password protected computer and paper documents will be secured in a locked file cabinet. The researcher will keep one confidential hard copy and electronic back-up of a master list of students' names linked to student number secured in a locked file cabinet separate from other study information. Once the study is complete, electronic data will be destroyed and hard copies of any documents will be destroyed by a crosscut shredder.

All student information and records will be kept private and confidential **to the extent allowed by law**. The researcher may present the outcomes of this study for academic purposes such as articles, teaching, or conference presentations. The researcher may use the data and results of the study for future analysis or publications connected to the study. A subject number instead of your child's name will be used on study records and any facts that might identify your child will not appear when the results of this study are presented or published.

If at any time you have questions or problems about this study, you can contact the researcher, Eileen Kaskons (978-244-0791 or ekaskons@lesley.edu), or the Lesley University sponsoring faculty, Dr. Adam Meiselman (617-349-8477 or ameiselm@lesley.edu).
We will give you a copy of this consent form to keep.

Signatures and Names:

_____	_____	_____
Date	Researcher's Signature	Print Name
_____	_____	
Date	Print Your Child's Name	
_____	_____	_____
Date	Parent/Guardian Signature or Legally Authorized Representative	Print Name

There is a standing committee for Human Subjects in Research at Lesley University to which complaints or problems concerning any research project may, and should, be reported if they arise. Contact the Committee Chairpersons at irb@lesley.edu.

Student Informed Assent – Group 2

Title: The Effect of Mindfulness Meditation Using a Smartphone app on Stress in High School Students.

Researcher: Mrs. Eileen Kaskons, a doctoral student at Lesley University.

Description and Purpose: You are being asked to volunteer for this research study during your Health class. The purpose of this study is to investigate the effect of the Stop, Breathe & Think (SB&T) mindfulness app on student stress levels. SB&T (https://www.stopbreathethink.com/) includes lessons in nonreligious mindfulness meditation.

Procedures: This study will be conducted at Greater Lowell Technical High School from September to November 2019 with the researcher and health teacher as part of your Health class. Your class has been selected to be on a waitlist and is a very important part of the study because you will help establish pre-study stress levels. You will receive the same app used by the non-waitlist group after that group's study has been completed. Since you are waiting until the second half of the semester to participate, we will measure your stress level six times during the first ten weeks before you start using the app. By signing this form, you agree to:

<u>During the First 10 Weeks of the Semester:</u>

- Complete a short ten-question Perceived Stress Scale (PSS-10) assessment six times.

<u>After the Non-waitlist Group has Completed the Study:</u>
In class:

- Bring smartphone and headphones to class. (If you do not have equipment, an iPad or headphones will be provided.)
- Download the Stop, Breathe & Think (SB&T) app to your smartphone.
- Listen to the first four of ten introductory sessions.

Outside of class:

- Finish the remaining six introductory sessions on your own.
- Listen to one mindfulness session four times a week until the end of the semester. You can choose from sessions that are one to nine minutes long.
- Log the session name and length in time on an electronic document.

<u>The Last Week of the Semester:</u>

- After using the app for the rest of the semester, you will repeat the PSS-10 assessment.

Risks: Participation in this research is voluntary. Even if your parent/guardian has given consent, **you have the right to refuse to be in this study**. If you decide to be in the study and change your mind, you have the right to withdraw at any time. You may skip questions and/or choose not to complete mindfulness sessions. Your decision to participate, not participate, or withdraw will not affect your grade in Health class or result in any negative consequences.

Participation in this research may lead to a reduction in stress and poses minimal risk but may not provide any benefit to you. In the event of any discomfort during the study, you will be referred to your school counselor.

Confidentiality, Privacy and Anonymity: All student information and records will be kept private and confidential **to the extent allowed by law.** The researcher will use a subject number instead of your name on all forms, documents, and future presentations or publications. Once the study is over, all electronic data and paper documents will be destroyed. The researcher may present the outcomes of this study for academic purposes but will never reveal your name or any facts that might identify you.

If at any time you have questions or problems about this study, you can contact the researcher, Mrs. Eileen Kaskons (978-244-0791 or ekaskons@lesley.edu), or the Lesley University sponsoring faculty, Dr. Adam Meiselman (617-349-8477 or ameiselm@lesley.edu). We will give you a copy of this assent form to keep.

Signatures and Names:

_____	_____	_____
Date	Researcher's Signature	Print Name
_____	_____	_____
Date	Your Signature	Print Name

There is a standing committee for Human Subjects in Research at Lesley University to which complaints or problems concerning any research project may, and should, be reported if they arise. Contact the Committee Chairpersons at irb@lesley.edu

Parent/Guardian Informed Consent – Group 2

Title: The Effect of Mindfulness Meditation Using a Smartphone app on Stress in High School Students

Researcher: Eileen Kaskons, a doctoral student at Lesley University.

Description and Purpose: Your child is being asked to volunteer for this research study during their Health class. The purpose of this study is to investigate the effect of the Stop, Breathe & Think (SB&T) mindfulness app on student stress levels. SB&T (https://www.stopbreathethink.com/) includes lessons in nonreligious mindfulness meditation.

Procedures: This study will be conducted at Greater Lowell Technical High School from September to November 2019 with the researcher and health teacher as part of your child's Health class. Your child's class has been selected to be on a waitlist and is a very important part of the study because they will help establish pre-study stress levels. They will receive the same app used by the non-waitlist group after that group's study has been completed. Since they are waiting until the second half of the semester to participate, we will measure their stress level six times during the first ten weeks before they start using the app. By signing this form, you agree that your child can participate in the following activities:

During the First 10 Weeks of the Semester:

- Complete a short ten-question Perceived Stress Scale (PSS-10) assessment six times.

After the Non-waitlist Group has Completed the Study:
In class:

- Bring smartphone and headphones to class. (If your child does not have equipment, an iPad or headphones will be provided.)
- Download the Stop, Breathe & Think (SB&T) app to their smartphone.
- Listen to the first four of ten introductory sessions.

Outside of class:

- Finish the remaining six introductory sessions on their own.
- Listen to one mindfulness session four times a week until the end of the semester. Your child can choose from sessions that are one to nine minutes long.
- Log the session name and length in time on an electronic document.

The Last Week of the Semester:

- Your child will repeat the PSS-10 assessment.

Risks: Participation in this research is voluntary. Your child has the right to refuse to be in this study. If your child decides to be in the study, your child has the right to change their mind and stop at any time. Your child may skip questions and/or choose not to complete mindfulness sessions. All of your child's questions will be answered at any time. Your child's decision to participate, not participate, or withdraw will not affect their grade in Health class or result in any negative consequences.

Participation in this research may lead to a reduction in stress and poses minimal risk. This research may not provide any benefit to your child. Parents and guardians are encouraged to discuss this study with their child and report any concerns to the researcher. In the event of any discomfort during the study, your child will be referred to their school counselor.

Confidentiality, Privacy and Anonymity: We will use a subject number instead of your child's name on all forms and documents. Electronic data will be kept confidential and secured on a password protected computer and paper documents will be secured in a locked file cabinet. The researcher will keep one confidential hard copy and electronic back-up of a master list of students' names linked to student number secured in a locked file cabinet separate from other study information. Once the study is complete, electronic data will be destroyed and hard copies of any documents will be destroyed by a crosscut shredder.

All student information and records will be kept private and confidential **to the extent allowed by law**. The researcher may present the outcomes of this study for academic purposes such as articles, teaching, or conference presentations. The researcher may use the data and results of the study for future analysis or publications connected to the study. A subject number instead of your child's name will be used on study records and any facts that might identify your child will not appear when the results of this study are presented or published.

If at any time you have questions or problems about this study, you can contact the researcher, Eileen Kaskons (978-244-0791 or ekaskons@lesley.edu), or the Lesley University sponsoring faculty, Dr. Adam Meiselman (617-349-8477 or ameiselm@lesley.edu).
We will give you a copy of this consent form to keep.

Signatures and Names:

_____	_____	_____
Date	Researcher's Signature	Print Name
_____	_____	
Date	Print Your Child's Name	
_____	_____	_____
Date	Parent/Guardian Signature or Legally Authorized Representative	Print Name

There is a standing committee for Human Subjects in Research at Lesley University to which complaints or problems concerning any research project may, and should, be reported if they arise. Contact the Committee Chairpersons at irb@lesley.edu

Appendix D

Class Schedule

Class Period	A-Week	X-Week
1	Group A	Group A
2	Group B	Group B
3	---------	----------
4	Group A	Group A
5	---------	----------
6/7*	Group B	Group B
8/9	---------	---------
10	Group A	Group A
11	Group B	Group B

*Periods 1, 2, 3, 4, 10 and 11 are 44 minutes. Periods 5, 6, 7, 8 and 9 are 22 minutes each, therefore Period 6/7 is 44 minutes.

Appendix E

Study Calendar

	A/X Week	M	Tu	W	Th	F
September	A Week	2 Off Labor Day	3	4	5	6
	X Week	**9 Consent**	**10**	**11**	**12 Pretest T1A**	**13 Pretest T1B**
	A Week	16 Consent	17	18	19 Pretest T1A	20 Pretest T1B
	X Week	**23 Session 1**	**24 Session 2**	**25 Independent**	**26 Independent**	**27 Reminder**
	A Week	30 Session 1	1 Session 2	2 Independent	3 Independent	4 Reminder
October	**X Week**	**7**	**8**	**9**	**10**	**11**
	A Week	14 Off Columbus Day	15	16 PSAT	17	18
	X Week	**21**	**22**	**23**	**24 Midtest T2A**	**25 Midtest T2B**
	A Week	28	29	30	31 Midtest T2A	1 Midtest T2B
November	**X Week**	**4**	**5**	**6**	**7**	**8**
	A Week	11 Veteran's Day	12	13	14	15
	X Week Posttest	**18 Posttest T3A**	**19 Posttest T3B**	**20**	**21**	**22**
	A Week Posttest	25 Posttest T3A	26 Posttest T3B	27 Half-day	28 Thanksgiving	29 Off

Appendix F

Stop, Breathe & Think Approval Email

Stop, Breathe & Think Research

Stop Breathe & Think app <info@stopbreathethink.com>
Tue 8/13/2019 5:22 PM
To: Kaskons, Eileen <ekaskons@lesley.edu>

Hello Eileen Kaskons,

Your Research Group Application for The Effect of Mindfulness Meditation Using a Smartphone App on Stress in High School Students has been approved!

You will receive another email containing your login information. Please change your password after you've logged in for the first time using the "profile" link in the top right.

If you have any questions or need assistance, please email research@stopbreathethink.com.

Happy researching!

Stop, Breathe & Think.

Appendix G

Stop, Breathe & Think "Get Started" Lessons

Lesson #	Title	Time
1	What to Expect	3 minutes
2	Make it Routine	2 minutes
3	One Minute to Mindfulness	1 minute
4	Mindful Breathing	3 minutes
5	Lion Mind	7 minutes
6	Counting Breaths	7 minutes
7	Body Scan	8 minutes
8	Engaging Your Senses	5 minutes
9	Mindful Walk	4 minutes
10	Stop, Breathe & Think	7 minutes

Appendix H

Research Study Exit Survey

Please do NOT put your name on this. Please take a few minutes to answer this anonymous survey about your experience with the app and our research study. When you are finished, please fold this form in ½ and put your answers in the sealed container provided. Your answers **CAN NOT** be traced back to you, so we would greatly appreciate your honest input on this form.

How often did you use the app outside of Health class? Please circle one.

A. Never B. Once or twice C. A few times D. Many times

If you circled **A or B,** please **ANSWER** questions 1 and 2, skip 3 and 4, then answer questions 5-9.

If you circled **C or D**, please **SKIP** questions 1 and 2, go to question 3, and answer questions 3-9.

1. If you **DID NOT** use the app outside of Health class, please list a few reasons why.

_____
_____
_____
_____
_____
_____

2. If you **DID NOT** use the app outside of Health class, what would have helped you use it?

_____
_____
_____
_____
_____
_____**(AFTER Q. #2 go to Q. #5)**

3. If you **DID** use the app outside of Health class, please list a few reasons why.

_____
_____
_____
_____
_____
_____

4. If you **DID** use the app outside of Health class, how did you feel after a meditation session?

_____
_____
_____
_____

All participants please answer the following questions.

5. What did you like and/or dislike about practicing meditation with the app during Health class?

6. If you were allowed to use your cell phone during school, what other times during school would you use the app?

7. What changes have you noticed in yourself since the beginning of this study?

8. What did you like and/or dislike about the Stop, Breathe & Think app in general?

9. What recommendations do you have to improve this study experience for future classes?

Appendix I

Test-Retest Reliability Coefficients

Interval/Sample	Intraclass Correlation	95% Confidence Interval		Pearson
		Lower Bound	Upper Bound	
24 Hour *N* = 190	ICC = .92*	0.89	0.94	*r* = .92*
6-week *N* = 62	ICC = .69*	0.54	0.80	*r* = .69*
3-week *N* = 55	ICC = .63*	0.40	0.79	*r* = .63*

*$p < .001$, two-tailed

Appendix J

Student Meditation Frequency

Total sessions completed	*f*	%	Cum %
0	14	13.5	13.5
1	10	9.6	23.1
2	10	9.6	32.7
3	13	12.5	45.2
4	11	10.6	55.8
5	12	11.5	67.3
6	6	5.8	73.1
7	6	5.8	78.8
8	5	4.8	83.7
9	3	2.9	86.5
10	3	2.9	89.4
11	3	2.9	92.3
12	3	2.9	95.2
15	2	1.9	97.1
16	1	1	98.1
19	1	1	99
24	1	1	100
Total	104	100	

Appendix K

Qualitative Themes

Theme	Sub-theme	Codes	Data Extracts (Respondent Number)
Resistance	1. Did not value or like independent use	a. Didn't need/want/like	*I didn't need to or want to use it.* (1) *What I didn't like was the meditation.* (18) *I didn't feel like I needed it. I was fine without it.* (34) *Nothing would really help me out because I didn't see a need to use it.* (35) *I had other better things that helps me relax.* (60) *I get bored using the app.* (65)
		b. Didn't help/work	*I didn't like it because it was like homework. I felt like we had to do it. We couldn't just go on our free time and do it.* (86) *Founded it uneffective* [sic]. *It was repetitive. Didn't feel like I needed it.* (87) *I didn't use the app because it wasn't useful for me outside of school. Wasn't helpful.* (13) *I never felt a big impact in my life while using it and it didn't really relax me as much as it was meant to.* (23) *Didn't really help me in class. Don't know why it would help me outside of class. Never had time. Wasn't necessary for me.* (7) *It was pretty neat to try it out, but it didn't help me.* (30) *It didn't make me feel different.* (*heart symbol)* (69)

		c. Not my thing	*It's not the thing for me I just don't do that type of stuff.* (5) *It's not my go to thing to destress myself. I have my own ways to calm down. Whenever I'm stressed about something like music which helps me a lot to relax.* (9)
	2. Would not attempt independent use	a. Nothing would motivate use	*Nothing that you could do to the app would make me use it. It's a good app and it would help people who want to be helped.* (5) *I did not like meditation, I found it boring and a waste of time for me. I did like that the app ran in the background so I could use Reddit.* (65)
		b. Would not use in school	*I wouldn't use the app I would use my phone for other things.* (3) *Never because most likely forget.* (18) *I honestly don't think I would use the app if phones were allowed during school.* (36)
	3. Did not enjoy classroom use	a. Didn't like	*What I didn't like was the meditation.* (18)
		b. Took too long	*I dislike how long some of the sessions were.* (92)
		c. Was boring/repetitive	*It was boring. Wasn't really that helpful. I didn't feel any difference before and after.* (31)
Theme	Sub-theme	Codes	Data Extracts (Respondent Number)
Hesitance	1. Difficult to schedule	a. Too busy/no time	*Throughout my day, I usually am doing chores or working around the house, or on my car. To be honest, I always was busy and when I did, I didn't think to go on my phone.* (38) *Because I was too busy with work and sports. Had no time to stop and do this.* (6) *I think I did not use the app because most of the time I couldn't do the sessions when I set my reminder and/or being distracted by more interesting things.* (81)

			After school I have sports till 5 and when I get home I eat, shower, and do homework and never really think about going on the app and doing a stress thing. (36) *I was busy with football. Practice every day watching film every day made me really tired. Had to take care of my father.* (86)
	2. Not a primary concern	a. Forgot	*Most of the time it was just me forgetting. The couple times I used it was when I felt like my mental health was falling apart.* (33) *I mostly forgot about it or did not want to do it.* (3) *I honestly completely forgot to most times or did not feel the need to.* (27)
		b. Not a priority	*I forgot to do it a lot of the time. I also do not constantly use my phone. As a result, the reminder tended to get buried in other notifications.* (28) *I didn't have time and it wasn't really a priority for me. I have a job and I leave at 7 PM so I'm usually exhausted when I get home.* (75)
	3. Would only use if needed for mental health	a. If stressed	*If I was more stressed in my life I might of tried it and if it was made mandatory* [sic] *for students.* (36) *Stress, anxiety.* (83)
		b. If depressed	*Maybe if I was depressed or going through things.* (8) *If I thought I needed it. Or I was desperate enough to use it.* (34) *I probably would have used it if I felt sad or depressed.* (76)

	4. Experienced fluctuating results	a. Unsure about classroom experience	*Liked that it was quick and easy. Disliked that somethings seemed repetative* [sic]. (22) *I like how it made me feel calm. I disliked it because it's not really my thing.* (91) *I didn't like that I was with people. I'd feel more relaxed if I was by myself.* (42) *Liked having some quiet time to myself. Dislike having to make a log for every session.* (29)
		b. Inconsistent independent experience	*Sometimes I felt the same but sometimes more calmer* [sic]. (48) *After each session, I would feel even slightly calmer or I would feel indifferent. It really depended on specific sessions.* (90) *Sometimes I would be in a good mood after doing a session, and sometimes I would be still in a bad mood after a session.* (98)
Theme	Sub-theme	Codes	Data Extracts (Respondent Number)
Acceptance	1. Used as a strategy to induce serenity	a. Used to relax/calm down/pause and think	*Helped me relax. It became a part of my daily routine. I enjoyed using it.* (45) *It was helpful to relax and help take my mind off of a stressful situation. Although it was usually not the first thing I thought of when stressed or anxious, it was still a helpful tool.* (49) *Because it was relaxing. It calms me down.* (54) *It calmed me down when I was paranoid. It grounded me when I didn't think I could make it through something. It helped me improve my mental health a little bit and it changed my perspective of myself, so I hate myself less.* (55) *Sometimes I felt like I needed to have something to escape with and make me feel better.* (97)

		b. Used it to aid sleep	*It helps me to fall asleep and helped me stop and think for a bit.* (48) *I mostly use it when I was trying to relax or sleep just to quiet my mind a bit.* (51) *It helped me fall asleep a few times and helped me relax before going to bed.* (95)
		c. To cope/reduce stress	*Helping coping w/ stress/anxiety. Calming. Sleep.* (56) *I had a death in my life and needed some mindfulness to get my head straight. I also was just stressed.* (57) *To calm down any excess anxiety. To take me time to chill out. To take extra stress off my mind.* (90) *Helping coping w/ stress/anxiety. Calming. Sleep.* (94)
	2. Positive independent experience	a. Calm/relaxed	*I felt a little calm. Made me feel I was in control of my emotions. I really think it can help.* (42) *The sessions were very calming. It was easier to process new information and sort through problems after a session.* (49) *I felt like a new person I was so relaxed.* (54) *I felt calmer, less scared because of how fast my heart was beating, relaxed.* (55) *After the sessions, I did feel calmer and more peaceful than I had been just before the session.* (51) *I felt calmer, relaxed, and I would always get in a better mood.* (96)

	3. Tranquil classroom experience	a. Calming/peaceful/relaxed/less stressed	*I liked how it calms me down and relaxes me. I like how it kind of makes me escape reality and focus on myself. I did not really dislike anything.* (74) *I liked it because it was peaceful. Dislike. None.* (8) *I like practicing in class because even if I forgot to do it we still had an opportunity to relax.* (55) *It's relaxed me when I was stressed.* (72)
		b. Good start to the day or class/nice break	*I liked it because it gave me a good start to the day and I was relaxed all day.* (96) *It was a good break from a difficult school day and presented an alternative path rather than listening to negative conversations.* (49)
		c. Enjoyed quiet/ environment	*I liked how quiet everyone was and how concentrated people were.* (15) *Liked having some quiet time to myself. Dislike having to make a log for every session.* (29)
		d. Increased focus	*I liked that I could concentrate for the rest of the day, and that I had a period of peace where I could just relax.* (27) *I like how he gave me a chance to recalibrate myself and actually focus in shop.* (62)
	4. Would use in school	a. Lunch/free time	*During lunch because that's when I get to hang out with friends to relax from the work in shop or class.* (9) *I would use the app if I'm done with all of the classwork early.* (19)

			A class when I'm done with the work or when I need a break because I just can't seem to do the work right. (70) *In the halls, before test, before a game, while studying.* (93)
		b. Before a test/quiz/presentation	*I would 100% use it before I took a test. It would help so much.* (27) *Before a presentation, during times of stress.* (55) *Probably right before a major test or quiz.* (62) *I would use the app before test/exams when my anxiety is higher than normal or before heading to a stressful class.* (90)
		c. In times of stress	*Math, the class is very stressful.* (15) *I would maybe use it at the times that I was anxious or nervous or when I just needed to calm down.* (21) *I probably would not use the app unless I was having a stressful moment in school.* (61) *I would use it after a stressful test or if I'm really tired/angry/sad.* (73)
	5. Experienced an increase in overall well-being	a. Calmer/relaxed/less stressed	*I was stressed before but meditation kind of helps me calm down.* (25) *I feel more relaxed and not stressed anymore.* (58) *I'm more calm with situations that I used to get mad at.* (65) *I have noticed I'm almost more capable of handling my stress over minor things.* (90) *A bit calmer and I would hear people out.* (4)
		b. Increased emotional awareness/regulation	*I've been more observant* [sic] *with myself and how I feel during certain situations.* (42)

		c. More positive/confident	*I am more concious* [sic] *of how anxious I let myself get, so I use some of the exercises to re-evaluate.* (50) *I've noticed my emotions a lot more.* (73) *I have been more self-aware of what I am doing.* (75) *I ended up having a more positive outlook in life.* (19) *How positive I come off with somethings [sic].* (76) *I have more confidence in handling situations.* (49) *I've been happier, less worried about everything, and I'm trying to speak up for myself more.* (55) *Feeling miserable à not so miserable.* (97)
		d. Increased focus	*Helped me focus and relax more.* (8) *I notice that well* [sic] *I was studying this app it was helping me and make me focus at school.* (46) *I have noticed that I am quieter and able to listen and focus better.* (51)
Theme	Sub-theme	Codes	Data Extracts (Respondent Number)
Recommendations	1. Provide variety of class/app activities	a. Add games/fun activities	*Put meditation games on the app.* (2) *Make it less repetitive and boring, add fun things to it.* (7) *Make it fun.* (11) *The app was cool, the app should add more talks and some interaction.* (40) *More topics on the app to cover more areas and situations.* (76)
		b. Include varied audio	*A music session for those who want a type of music genre to listen to than listen to someone talk.* (60)

		c. Allow collaboration	*Make different audio for background.* (48) *In the activeitys* [sic] *have a partner.* (18) *I recommend that you try to get everyone envolved* [sic] *and after you share your thoughts and feelings with each other.* (24)
	d. Build into the curriculum/ school day	e. Make it a requirement/assignment	*I think this study would've been better if some of the meditation was mandatory, because this would increase student motivation.*(81) *I would have used it more if it was assigned and not a voluntary thing.* (23) *If it had to be done. If it were mandatory, then I would've used it a lot more.* (27) *If the teachers told me to use it that's when imma* [sic] *use it.* (58) *If it was a homework grade or classwork grade.* (69)
		f. Allow more time in class/increase length of the study	*I would focus on doing it more in class & having more features because kids (like myself) get distracted by fun things to do at home.* (61) *Designate more time to use app in class.* (64) *More time in class to do it.* (39) *Have them do the study for a longer time.* (15) *Give more time for sessions.* (83)
		g. Allow in other classes	*Have more classes use it at the beginning of class.* (28) *I enjoyed it and I thought it was fun so I cant* [sic] *really add a recommendation because I enjoyed it. I think people should do this even after you stop going to health.* (45)

			I would recommend asking teachers if they would allow/offer group meditation before test: teacher would pull up a meditation video that the students could listen to. (51) *Give out permission slips for their teachers to sign and give them permission to use the app before tests, during stressful assignments, etc.…* (89)
	h. Buy premium	a. Disliked being limited to free sessions	*I didn't like certain ones being locked.* (17) *I wish that we could access more sessions (the locked ones.)* (37) *The fact that some of the sessions had to be paid for was something that I strongly disliked.* (62)
	b. Facilitate independent use	a. Find time	*If I was not as busy as I am right now during weekdays and weekends I would probably use the app.* (19) *If I was a little more organized at home, it definitely would've helped in using the app at home.* (62)
		b. Optimize notifications	*If the reminder sat at the top of my notifications.* (28) *If I had a reminder. There was a notification from the app, but I would see it and maybe forget or was busy.* (33) *Reminder to use it, and have more responsibility with it.* (85)
		c. Emphasize results	*Explain the importance of practicing at home.* (49) *Maybe talk about how its* [sic] *successful.* (57)

	c. No recommendations for study	a. No recommendations	*I wouldn't know what to improve.* (1)
			Nothing. Since the app has what it needs to help someone relax. (30)
		b. Positive experience	*No recommendations!! It was a good experience.* (50)
			Nothing yall [sic] *doing great.* (88)
			Nothing, everything was run perfectly. (96)