2019

Development of The Adept Screening tool to Identify Health and Fitness Apps That Pose Risk For individuals Susceptible to Disordered Eating

Georgia Mcartney
georgie.mcartney@gmail.com

Follow this and additional works at: https://huskiecommons.lib.niu.edu/allgraduate-thesesdissertations

Part of the Nutrition Commons

Recommended Citation
Mcartney, Georgia, "Development of The Adept Screening tool to Identify Health and Fitness Apps That Pose Risk For individuals Susceptible to Disordered Eating" (2019). Graduate Research Theses & Dissertations. 7415.
https://huskiecommons.lib.niu.edu/allgraduate-thesesdissertations/7415

This Dissertation/Thesis is brought to you for free and open access by the Graduate Research & Artistry at Huskie Commons. It has been accepted for inclusion in Graduate Research Theses & Dissertations by an authorized administrator of Huskie Commons. For more information, please contact jschumacher@niu.edu.
ABSTRACT

DEVELOPMENT OF THE ADEPT SCREENING TOOL TO IDENTIFY HEALTH AND
FITNESS APPS THAT POSE RISK FOR INDIVIDUALS SUSCEPTIBLE
TO DISORDERED EATING

Georgia McArtney, MS
School of Health Studies
Northern Illinois University, 2019
Dr. Sheila Barrett, Director

Popular health apps could contain risk factors for disordered eating such as restriction, weight stigma, harmful exercise, or compensation. Although previous studies have examined the connection between social media and disordered eating, data is lacking regarding risk factors from health apps. The purpose of this study is to develop a rating tool to detect, examine, and categorize app-based risk factors for disordered eating.

This study utilized a comprehensive review of literature regarding risk factors for disordered eating, connections between app use and disordered eating behaviors, and existing technology rating tools. Based on this framework, the Awareness of Disordered Eating Promotion in Technology (ADEPT) was developed to contain four domains: Restriction, Weight Stigma, Exercise, and Compensatory Behaviors. The ADEPT was pilot tested on three selected apps by three trained raters in order to assess clarity and comprehension of the tool. The ADEPT was revised based on rater feedback, then used to assess the 10 most popular health and fitness apps (based on inclusion criteria).

Results from the 10-app test indicate acceptable to excellent internal consistency and large positive correlation for the following domains: Restriction, Weight Stigma, Compensatory Behaviors, and overall ADEPT total scores. The overall ADEPT results satisfied the desired level of 0.7 for Cronbach’s alpha; therefore, it is an effective tool for use in future research and indicates that disordered eating behaviors are prevalent in popular health apps.
DEVELOPMENT OF THE ADEPT SCREENING TOOL TO IDENTIFY HEALTH AND FITNESS APPS THAT POSE RISK FOR INDIVIDUALS SUSCEPTIBLE TO DISORDERED EATING

BY

GEORGIA MCARTNEY
© 2019 Georgia McArtney

A THESIS SUBMITTED TO THE GRADUATE SCHOOL IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE MASTER OF SCIENCE

SCHOOL OF HEALTH STUDIES

Thesis Director:
Dr. Sheila Barrett
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF APPENDICES</td>
<td>v</td>
</tr>
<tr>
<td>CHAPTER I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER II. REVIEW OF LITERATURE</td>
<td>8</td>
</tr>
<tr>
<td>CHAPTER III. METHODS AND DESIGN</td>
<td>21</td>
</tr>
<tr>
<td>CHAPTER IV. RESULTS</td>
<td>30</td>
</tr>
<tr>
<td>CHAPTER V. DISCUSSION</td>
<td>38</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>45</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>49</td>
</tr>
<tr>
<td>Table</td>
<td>Page</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1. Data Analysis Methods</td>
<td>27</td>
</tr>
<tr>
<td>2. Mean ± SD and p Values</td>
<td>31</td>
</tr>
<tr>
<td>3. Research Timeline</td>
<td>70</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Comparison of Apps by Raters: Restriction</td>
<td>32</td>
</tr>
<tr>
<td>2.</td>
<td>Comparison of Apps by Raters: Weight Stigma</td>
<td>33</td>
</tr>
<tr>
<td>3.</td>
<td>Comparison of Apps by Raters: Exercise</td>
<td>34</td>
</tr>
<tr>
<td>4.</td>
<td>Comparison of Apps by Raters: Compensatory Behaviors</td>
<td>36</td>
</tr>
<tr>
<td>5.</td>
<td>Overall Totals</td>
<td>37</td>
</tr>
<tr>
<td>6.</td>
<td>Most Popular Health and Fitness Apps in The United States as of May 2018</td>
<td>50</td>
</tr>
<tr>
<td>7.</td>
<td>Market Reach of the Most Popular Mobile App Categories in the United States as of July 2018</td>
<td>51</td>
</tr>
</tbody>
</table>
LIST OF APPENDICES

Appendix                              Page

A. POPULAR HEALTH APPS.............................. 49
B. SELECTED HEALTH APPS............................ 52
C. ADEPT .............................................. 54
D. FEEDBACK FORM .................................... 64
E. RATER TRAINING .................................... 66
F. RESEARCH TIMELINE ............................... 69
CHAPTER I: INTRODUCTION

The current health culture is rife with consumer-friendly tools ranging from mobile health applications (apps) to social media accounts and virtual accountability groups. As of fall 2018, there were more than 4 million apps available from Android and Apple app stores in the United States\(^1\), Health and Fitness apps have a 64.5% U.S. market reach\(^2\), and the 10 most popular health/fitness apps combined have 75.5 million monthly U.S. users\(^3\) (see Appendix A for figures). While health apps presume to make behavior change accessible to a wide audience, questions of efficacy and safety arise when we examine the quality of information and method of delivery promoted by many outlets. Several recent studies examined the theoretical framework for health apps, with disturbing conclusions that popular wellness apps have little to no evidence-based efficacy and could instead be causing harm\(^4,5,6\). With the far-reaching audience that health apps boast, it could prove dangerous if these apps are not promoting evidence-based practice and are failing to create effective behavior change.

While health apps are gaining widespread usage, public awareness of disordered eating (DE) and eating disorders (EDs) is also increasing. Eating disorders have strict diagnostic criteria defined by the American Psychiatric Association, whereas disordered eating is a descriptive term for unhealthy eating patterns not linked to a diagnosis\(^7\). Disordered eating encompasses a generalized set of symptoms which can often foreshadow development of an eating disorder.
Common disordered eating behaviors include frequent dieting, weight fluctuation, food or exercise rituals, food guilt, or compensating for “bad foods” with restriction or exercise. These disordered behaviors can closely mirror encouraged behaviors in widely used health apps. For example, the strict daily food records promoted in calorie tracking apps mirrors the obsessive eating patterns seen in disordered eating. Similarly, encouragement to exercise as an attempt to counterbalance food intake comes dangerously close to the exercise compulsion symptomatic of disordered eating. Many health apps encourage users to set a goal weight, but unrealistic goal weights can lead users to engage in disordered eating behaviors in order to achieve or maintain these goals.

Health apps often employ a “community” feature for users to hold each other accountable and share their weight loss “journeys.” It is widely accepted that two major risk factors for eating disorder development are media exposure and peer pressure, aspects that virtual health communities employ. It therefore appears that these virtual communities are in danger of creating an environment rife with disordered eating through features which could cause users to feel ashamed or inadequate for not engaging in restrictive or compulsive health behaviors. One study of an online bodybuilding community found that members considered no health cost too great in pursuit of strengthening their bodies, and while this community represents an extreme faction of health-focused people, it is not difficult to imagine justifying disordered eating in pursuit of a smaller and more toned body.

The above health app features raise the question: where is the line between a community of accountability and outright peer pressure? When health communities hold up idealized bodies
(read: thin yet somehow curvy and muscular) as the goal for their members, they ignore that this body may well be unattainable for many app users. Media exposure in the health world seems based on the myth that all people can attain a perfectly toned, thin body with the right diet and enough self-control, yet evidence shows that weight loss is unsustainable for many people\textsuperscript{10–12}. If members of health app communities are unable to maintain weight loss (and evidence suggests that dieting often leads to weight cycling, not weight loss\textsuperscript{10,13}), they could turn instead to disordered eating to achieve their weight loss goals\textsuperscript{14}. Due to the potentially detrimental habits health apps can promote, this study will examine the prevalence of disordered eating behaviors or thoughts suggested in popular health apps.

Health apps can perpetuate weight bias through visuals, language, and underlying beliefs, conveying the idea that certain body types, weight statuses and appearances are unacceptable and must be changed through use of the health app. Users can internalize this weight bias, and research suggests that weight bias internalization negatively impacts health outcomes\textsuperscript{15} and is a risk factor for disordered eating\textsuperscript{16}. One study cites weight stigmatization as a meaningful predictor of binge eating\textsuperscript{17} while a separate study suggests that internalized weight bias correlates with higher frequency of binge eating\textsuperscript{18}. The focus of many health apps on changing the users body size could pose a risk for weight bias; therefore, the proposed screening tool will explore the presence of weight bias in apps.

Evidence demonstrates that food restriction is a predominant risk factor for bingeing, as laboratory studies indicate that restrictive and stressed rats display binge eating (BE) for highly palatable foods\textsuperscript{19}. A foundational study by Dr. Linda Bacon in the field of disordered eating
examined the sustainability of weight loss. Bacon’s study established the fact that dietary restraint can actually produce weight gain over time through “weight cycling” (repeated loss and regain of weight)\(^\text{10}\). This raises questions of ethics and best practice if we provide weight loss treatment which is ineffective or dangerous, as researchers suggest that dieting is the most common predictor for Anorexia Nervosa (AN), and binge eating is the most common onset symptom for Binge Eating Disorder (BED)\(^\text{20}\). Therefore, restrictive behaviors encouraged by health apps could contribute to disordered eating behaviors. By assimilating theoretical efficacy models with DE risk behaviors such as restriction, compulsion, and weight stigmatization, this study seeks to examine the relationship between popular health apps and the risk for DE.

Statement of the Problem

Factors such as weight bias, restriction, and exercise compulsion elevate risk for disordered eating. Health apps commonly employ techniques which utilize these same risk factors to encourage behavior change in users. This study seeks to examine the similarities in disordered behaviors and behaviors encouraged in health apps. Through development of a mobile app rating scale, screenings for DE patterns were combined with formats utilized by validated app screening tools to assess risk for DE behaviors in popular health apps. The rating tool features four behavioral categories associated with DE risk: Restriction, Weight Stigma, Exercise, and Compensatory Behaviors. The rating scale assigns a score for each category and a total score to each app assessed.
Background and Significance

Little data exists regarding DE risk from health app usage. Many studies focus on theoretical basis or efficacy for health apps, with alarming conclusions that many health apps are not supported by effective interventions or evidence-based information. As stated above, with a 64.5% market reach\(^2\) and 75.5 million monthly users\(^3\) in the U.S., health apps are a widely utilized source of health information and behavioral influence. If health apps encourage disordered eating risk factors such as restriction, weight stigma, harmful exercise, or compensation, then a growing percentage of the population is at risk for exposure to DE behaviors from app usage. By examining the behaviors and thought patterns which popular health apps encourage, this study investigated app-based risk factors for DE in order to increase awareness and prevent incidence.

Purpose

The purpose of this study is to detect, examine, and categorize app-based risk factors for disordered eating. Previous studies have explored the lack of theoretical framework and unreliable information provided by health apps\(^4,5,6\), yet little research exists determining the risk for DE from app usage. This study provides a foundation in the ongoing examination of the role of health apps in DE behaviors. Results of the study may be used in the prevention or early detection of DE from the use of health apps. Additionally, considering the 64.5% market reach of health and fitness apps\(^2\), practitioners can use this tool to promote apps that do not exacerbate DE risk. Research indicates that 62% of Registered Dietitians use health apps with their clients\(^21\);
therefore, this tool could be used to screen apps before suggesting them to clients to decrease exposure to DE behaviors.

Aims

The aim of this study is to provide a context for the current health app market from which to develop a screening tool to assess popular health apps for DE risks. This app assessment tool, called the Awareness of Disordered Eating Promotion in Technology (ADEPT), will categorize behaviors and thoughts encouraged by the apps based on their risk for or similarity to DE habits.

The study is divided into four specific aims as listed below:

Specific Aim 1: Identify popular health and fitness apps, categorize features of health and fitness apps, and identify app activities that could adversely impact individuals at risk for disordered eating behaviors.

Specific Aim 2: Develop a tool to identify features of health and fitness apps that could put individuals at risk.

Specific Aim 3: Using three trained raters, test the tool on three selected health apps to determine comprehension and clarity. Revise the tool based on feedback.

Specific Aim 4: Using the same three raters, test the revised tool on the 10 most popular health apps and perform statistical analysis to determine interrater reliability and internal consistency.

The specific aims (SA) build upon one another, as SA1 provides the foundation to guide the development of the tool in SA2, which was used in SA3 and SA4 for testing.
Research Questions

1. To what extent do popular health and fitness apps contain risk factors for disordered eating?

2. Do popular health apps contain features or content which may pose a risk for perpetuating disordered eating?

3. Can a screening tool be developed that can assist practitioners in quickly identifying apps containing features that could lead to disordered eating?

Objectives

1. To assess the extent to which popular health apps contain or encourage disordered eating behaviors.

2. To determine if popular health apps contain features or content which may pose a risk for perpetuating disordered eating.

3. To develop a screening tool to assess apps for disordered eating risk for use by consumers and practitioners to quickly identify apps containing features which could elevate risk.
CHAPTER II: REVIEW OF LITERATURE

Apps Used for Health Promotion

Pinto et al (2017) analyzed existing research regarding health promotion material to determine what types of health communication technology were available to adolescents\(^\text{22}\). The goal was to identify common themes in adolescent health education and learn what professionals were contributing to educational technology. The researchers used an integrative review framework to systematically categorize available research. Their search used keywords “information technology, adolescents, and health education” for articles published in English, Spanish, or Portuguese between January 2010 and December 2014. The inclusion criteria specified the articles must be complete, available at no cost, and responsive to the keywords specified above. Although these criteria produced 883 initial results, 23 articles were selected for analysis.

These articles were analyzed for areas such as type of study, health information and communication technology use, themes, types of technology, and professional involvement. The analysis of the 23 articles produced a categorization of types of education effectively communicated via technology. Analysis yielded 52.17\% of studies addressed adolescent sexual health, with less frequently addressed topics including asthma and diabetes; 65.21\% of the tools were implemented or developed by nursing professionals. Common media for education included text messages, websites, virtual learning games, and social media. The research
established that adolescents are more likely to gain health information through technology as opposed to conducting research or seeking a health professional. Therefore, if health professionals can develop and promote educational technology, they can ensure adolescents are accessing reliable health information.

Efficacy of Health Apps

Azar et al (2013) investigated if popular health apps were based on effective models of behavior change. Effective behavior modification techniques employed in clinical practice include self-monitoring and recording health outcomes, as compared with the goal of many apps to simply keep users engaged. The method employed was a comparative, descriptive assessment of top health apps, with the following inclusion criteria: free, top 200 health apps that track dietary and anthropometric data, and function independently. These criteria produced 23 apps, which were divided into five categories, with the top two from each category proceeding to the evaluation. Researchers used a validated behavioral health theory content survey that assessed for health belief model: the transtheoretical model, theory of reasoned action/planned behaviors, and the social cognitive/social learning theory. Apps were also assessed for persuasive ability.

The assessment included 20 intervention strategies that were categorized as knowledge, cognitive, behavior, emotion-focused, or therapeutic. These interventions are linked with the behavior theories listed above. Each app’s interaction with users was rated on five levels, ranging from simple informational exchange to individually tailored assistance. The apps were downloaded and tested for two weeks before being rated on a behavioral theory score (BTS) and persuasive technology score (PTS). The BTS score indicated the sum of the engagement in the
20 behavior intervention strategies based on the five levels of involvement, with a potential total of 100 items. The mean BTS was 8.1 (SD ± 4.2), with half of the apps receiving a BTS <10. PTS is based on persuasive technology theory, which holds that technology can change behaviors and attitudes through persuasion and social influence, so the PTS score rated the app on its ability to persuade in order to create change, with a potential score of 6. Mean PTS was 1.9 (SD ± 1.7); LoseIT scored highest, at a BTS of 14 and PTS of 5. Diet tracking apps such as LoseIT and MyFitnessPal consistently scored higher on BTS, showing that diet tracking apps are more likely to use behavioral theory. The study concluded that most apps incorporated some limited strategy, proving that using evidence-based behavioral interventions is possible in app-based education, yet all of the apps were lacking theoretical models of behavior change.

In a similar approach to the above study, Conroy et al (2014) examined behavior change techniques incorporated into fitness apps. A team of researchers characterized the behavior change techniques employed in the descriptions for popular physical activity (PA) apps. Eligible apps were selected from the top 50 paid and free lists, respectively, on iTunes and Google Play. Of these 200 potential apps, 167 involved PA and were selected for analysis. App descriptions featured between one and 13 behavior change techniques, such as instruction for how to perform a behavior, modeling a behavior, feedback, and goal setting, among others. Most app descriptions used less than four behavior change techniques. Latent class models were developed, with one to five classes. Most apps (54%) fell in the latent class of PA motivation, focused on social and self-regulated PA. The remaining 46% of apps were in the latent class of apps focused on providing instruction on behaviors and modeling. The study concluded that
behavior change techniques are not widely incorporated into health apps. Based on app descriptions and functions, many health apps are not founded on reliable methods.

A randomized controlled trial (RCT) of a smartphone application used for weight loss compared retention, adherence, and self-monitoring among app users, website users, and paper diary users\(^{24}\). The goal was to determine if app-based health interventions could provide a cost-effective, easily distributed weight management strategy. At the time the study was conducted (2013), no RCTs existed for smartphone weight loss apps focusing on eventual self-monitoring. Participants were 128 overweight volunteers randomized to one of three intervention groups: smartphone, website, or paper diary. The app used for the study was My Meal Mate (MMM), which was developed by the researchers to target behavioral factors in weight management. The main features of the app include goal setting, self-monitoring, and weekly text messages providing feedback. The website intervention was based on Weight Loss Resources, an existing website.

The trial proceeded over a six-month period at self-directed pace without intervention from the researchers. Researchers and participants met face-to-face at baseline, six weeks posttrial, and six months posttrial. The smartphone group evidenced highest retention, with 93%, followed by 55% in the website group and 53% in the diary group. Adherence was based on days diet was recorded, with a statistically significant higher rate of mean 92 days in the smartphone group, mean 35 days in the website group, and mean 29 days in the diary group. All groups exhibited a decline in self-monitoring over the course of the trial. Another outcome measure was mean weight change at 6 months: -4.6 kg in the smartphone group, -2.9 kg in the diary group,
and -1.3 kg in the website group. The BMI change at six months was -1.6 kg/m in the smartphone group, -1.0 kg/m in the diary group, and -0.5 in the website group. Interestingly the diary group displayed a higher mean weight loss and greater reduction in mean BMI despite the fact that they also evidenced the lowest adherence to diet recording. Overall the MMM app produced greater adherence, higher retention, and statistically significant results when compared with website and diary interventions. Health app usage for weight loss was more effective than conventional means for weight management. This indicates that using smartphone apps for health interventions can produce significant results.

A systematic review examined existing health apps for adolescents to determine evidence base for efficacy. The review analyzed available research for efficacy and acceptability of mobile apps focused on adolescent mental health. Reputable article sources such as Web of Science and Journal of Medical Internet Research were searched for relevant articles published between 2008 and 2016. Inclusion criteria required that articles feature mental health apps for adolescents focusing on such issues as depression or eating disorders. Based on predetermined criteria, 24 articles were included, and 15 apps were referenced in the included studies. Several studies did not demonstrate a significant effect of the apps. No apps had undergone research evaluation. The review concluded that of the available mental health apps for adolescents, there is insufficient evidence to support efficacy. Because of the increasing use of health apps, researchers recommended future studies to establish efficacy and safety of mobile apps. The current thesis contributes to this research need by examining efficacy of existing health apps with relation to disordered eating risk.
Noting a gap in app quality assessment, a team of Australian researchers developed a Mobile App Rating Scale (MARS) to assess the quality of mobile health apps. The quality assessment was based on content, appearance and multimedia, navigation, structure and design, and uniqueness. Notably, ratings of information quality were largely lacking in previous studies. The researchers randomly selected 60 mental health apps based on the following criteria: English, free, availability in iTunes, in categories: health and fitness, lifestyle, medical, productivity, music, education, and utilities. The first 10 apps were used for pilot testing by two expert raters who trialed the apps for 10 minutes, then rated them independently using the MARS. After rating, the two met to compare and remedy any disparities to ensure that the scale was reliable. Revisions were made to the MARS based on the pilot test, then the 50 apps were trialed and rated.

MARS consistency and total quality scores were calculated with Cronbach’s alpha to indicate that the correlation of items measuring the same general construct produced similar scores. Interrater reliability was determined with intraclass correlation coefficient. The MARS total mean score supplied a rating for the overall app quality, while subscale scores addressed specific strengths or weaknesses. The app quality total score and four subscales all had high internal consistency, showing that the MARS is a reliable indicator of app quality. The interrater reliability calculated by two-way mixed ICC was 0.79 (95% CI 0.75-0.83). The MARS total score internal consistency was Cronbach’s alpha 0.90. Internal consistency of the subscales was 0.80-0.89, median 0.85 Cronbach alpha with interrater reliabilities fair to excellent (ICC 0.50-0.80). The authors recommended that future users of the MARS undergo a training of the correct
use, understand who the app target users are, and clarify any ambiguities in the items and subscales. MARS was developed based on qualitative search of existing criteria and development of categories/subcategories to assess quality of mental health apps. Similar to the current study, the development and validation of the MARS enabled raters to assess app quality; by incorporating disordered eating risk factors into a similar rating scale, this study assessed the risk of health apps for eating disorders.

Technology and Disordered Eating

A recent study examined the effects of social media and apps used to monitor diet and exercise on disordered eating and exercise\textsuperscript{14}. Based on the social cognitive theory (SCT) framework that media consumption produces socially mediated pathways, researchers set out to examine the effect of SCT in media promoting dieting and fitness. Study participants (262) were recruited through social media posts and flyers. Participants reported their BMI and rated their media usage at baseline. Outcome measures included compulsive exercise, disordered eating, social media use, mobile app use, and traditional and microblog use. An eating attitude test was used to measure disordered eating, with a score \textgreater{} 20 indicating high risk and low scores still indicative of eating issues, with mean scores of 1.10-5.0 (M=2.56, SD=0.77).

App usage was assessed, with weekly use averaging at M=1.73 (SD=1.18). A hierarchical regression analysis was conducted to find the relationship between types of media and disordered eating, finding the beta for mobile apps as a positive predictor for stronger disordered eating symptomatology (β=.34, p=.00). Likewise, the beta for mobile app relationship with compulsive exercise was β=.31, p=.000, indicating a positive predictor for compulsive
behaviors. Researchers concluded that mobile app usage was significantly and positively associated with disordered eating and compulsive exercise. The more time an individual spent using a health app, the more likely they were to experience disordered eating. The authors hypothesized that when individuals spend more time on health apps using features such as calorie counters, they are more likely to become obsessed with calorie counting, restricting intake, or compulsive exercise. The findings of this study support the objective of this thesis to examine disordered eating risks embedded within popular health apps in order to develop a screening tool for DE risk.

In a study on “apps and eating disorders” a pair of researchers examined existing apps developed for eating disorder patients and clinicians to identify, characterize, and evaluate their impact and validity. Apps were identified by first entering eating disorder criteria into the search engines of every major smartphone app store to find 805 potential subjects, which were then narrowed down to determine which apps actually applied to patients (39) or clinicians (5). Download data for each app was determined, in addition to finding average rating by users. The function of each app was characterized as follows: provision of information, self-assessment, self-monitoring, and provision of advice or treatment. Information quality was rated as Good, Variable, or Poor based on evidence from reliable sources such as the DSM-V.

The above apps’ reliability with self-assessment was tested by one author assuming the role of a patient with Anorexia Nervosa, Bulimia Nervosa, Other Specified Feeding and Eating Disorder (OSFED), or no existing eating disorder. The app assessments were again classed as Good, Variable, or Poor based on questions and outputs. Self-monitoring was tested by the ease
of which information could be entered and the ability to reflect on previous entries. One author input case information for various types of patients to assess the ease of entering information and type of output. Advice and treatment were assessed both from the perspective of a clinician and patient and rated as Good, Variable, or Poor.

Information was provided by 13 of the 39 apps for ED patients, with two apps providing good information, eight providing variable information, and three providing poor information. Self-assessment was provided by five apps, of which two provided good assessment and three were variable or poor. Self-monitoring was provided by four apps, while three provided ease of input for information and none provided a readily usable way to view past entries and monitor differences. Advice and treatment were provided by 24 apps, with seven offering good advice, five offering variable advice, and 12 giving poor advice. Of the clinician-centered apps, services were outdated, limited, or not user-friendly. Based on the lack of reliable app support for eating disorder patients and clinicians, the researchers concluded that existing apps are not reliable and encourage the development of apps to complement the treatment process and focus on personalized interventions. The current thesis study contributes to the screening of apps to determine if they are reliable and could thus prove instrumental in development of ED-conscious patient and clinician apps.

Australian-based researchers Griffiths et al (2018) assessed influence of social media on eating disorder symptoms and body image on sexual minority male users of a dating app. The researchers hypothesized that more frequent social media use would be associated with body dissatisfaction, eating disorder symptoms (EDS), and consideration of anabolic steroid usage and
that this association would be stronger for image-based social media technology such as Instagram. Participants were recruited through advertisements on the dating app. Participants rated usage frequency for Facebook, YouTube, Instagram, Snapchat and similar social media platforms. The Male Body Attitudes Scale-Revised (MBAS-R) was used to assess body dissatisfaction, while the Eating Disorders Examination Questionnaire-Short (EDE-QS) measured eating disorder symptoms; 2,733 survey responses were included in the analysis. Spearman’s correlational analyses were used to assess bivariate associations alongside linear regression and visualization of correlation or frequency. Overall there was a positive relationship between social media usage and body dissatisfaction, EDS, and steroid consideration, especially with image-based outlets. The study concluded that more frequent usage of social media contributed to male body image dissatisfaction and EDS. This study reinforces the growing evidence base for negative body image due to social media exposure.

Cho et al (2015) studied how body image factors motivate students to use apps and likewise alter their perception of apps. The premise of the study was founded on the technology acceptance model (TAM), which is comprised of perceived usefulness (PU) and perceived ease of use (PEOU) and examines the role of perceptual motivators in adoption of new technology. The focus of the study was on the PU of health apps influencing intention for use. Researchers operated on the belief that body image is composed of three aspects – attitudinal, cognitive, and behavioral – and is explained through appearance, fitness, and health; the study focused on appearance and fitness. The hypotheses were based on the theory of reasoned action (TRA),
which holds that behaviors are triggered when subjects have a reasonable desire to achieve specific goals, so body image evaluation would trigger use of health apps.

Data were collected via online survey distributed at three Korean universities, garnering 294 responses. The survey contained screenshots of wellness apps to ensure participants understood the survey’s purpose. A 5-point Likert scale measured variables regarding appearance evaluation, fitness evaluation, and appearance orientation. A path analysis with AMOS 21 was used to evaluate the comparative fit index, infinite fit index, and root mean square error of approximation. Results suggested appearance evaluation significantly negatively predicted the PU of diet and fitness apps (beta = -0.28, p<0.001). Fitness evaluation negatively influenced PU of apps (beta -0.10, p=0.06). Appearance orientation did not significantly predict app PU, but fitness orientation positively significantly predicted it (beta=0.39, p<0.001). The PU of apps significantly predicted behavioral intention for using the health apps (beta=0.75, p<0.001). Therefore, fitness orientation (positive), fitness evaluation (negative), and appearance evaluation (negative) all significantly predicted the PU of health apps. The study concluded that app marketers should use strategies to attract individuals with low body image or physical confidence. These findings suggest employing negative body image as a motivator to attract users to health apps also exposes them to disordered eating risks. The combination of negative body image and disordered behaviors could create a perfect storm of precipitating factors for disordered eating.

Another study focusing on social media, “Do You ‘Like’ My Photo? Facebook Use Maintains Eating Disorder Risk” examines the relationship between Facebook use and
development of eating disorder symptoms. The basis for the study built upon the fact that two major eating disorder risk factors are media exposure and peer pressure. Facebook combines these factors in a uniquely potent manner. Researchers employed an experimental design to survey female subjects about their Facebook use and eating disorder symptoms. The first part of the study collected survey data from 1,960 women regarding Facebook usage and disordered eating symptoms. The second part of the study randomly assigned 84 women to use Facebook or another site for 20 minutes.

The cross-sectional results indicated that more frequent Facebook use was associated with eating disorder symptoms. Weight and shape concerns and anxiety were connected to Facebook usage as opposed to alternative site usage. The authors concluded that Facebook was dangerous for contributing to eating disorder risk factors and should be targeted for intervention. This study operated on the premise that exposure to media promotes restrictive eating, weight stigma, or social pressure to eat or look a certain way and therefore contributes to eating disorder risk. These same factors are at play in popular diet apps; therefore, the current research study extrapolated this relationship between disordered eating and risk factors embedded in popular health apps.

Research conducted by Hefner et al (2016) examined the use of mobile health apps and blogs in relationship to disordered eating and compulsive exercise. The basis for the study was the recent increase of fitspiration and thin idealization in social media and the use of mobile apps to record diet and exercise. The framework for the study employed social cognitive theory as a driving force behind the socially mediated pathways of media consumption and disordered
eating. A sample of 262 participants were recruited through social media and completed an online questionnaire regarding exercise, eating habits, health app usage, and blog or microblog use; 76% of participants were female, aged 18-27 years old. Measures included the compulsive exercise test (CET), the eating attitude test, overall social media use, blog and microblog use, and phone app use. Hierarchical regression analysis of control variables (age, sex, BMI, social media use) and disordered eating produced significant results, with the beta for microblog use $\beta = 0.24$, $p = 0.00$, and for mobile app use $\beta = 0.34$, $p = 0.00$. Therefore, the more time individuals spent using mobile applications and microblogs with fitspiration content, the more likely they were to report disordered eating behavior and compulsive exercise. The authors concluded by noting that because microblogs and mobile apps are implicated in disordered eating and compulsive exercise behaviors, parents, educators, and users should be made aware of these risks in order to reduce exposure to harmful technology.

Summary of Literature Review

In reviewing current literature regarding mobile apps usage, social media, and the risk for disordered eating, we see a lack of research connecting popular health apps with disordered eating risk. This risk is clearly established from social media, peer pressure, weight stigma, and healthy eating messages $^{8,16,17,18,20,26,29}$, and we see these same risk factors at play in popular health apps. This study seeks to link the above risk factors with health apps by creating an app rating scale to determine disordered eating risk based on themes within the apps.
CHAPTER III: METHODS & DESIGN

Research Design

The research design is based on the following specific aims (SAs):

Specific Aim 1: Identify popular health and fitness apps, categorize features of health and fitness apps, and identify app activities that could adversely impact individuals at risk for disordered eating behaviors.

Specific Aim 2: Develop a tool to identify features of health and fitness apps that could put individuals at risk.

Specific Aim 3: Using three trained raters, test the tool on three selected health apps to determine comprehension and clarity. Revise the tool based on feedback.

Specific Aim 4: Using the same three raters, test the revised tool on the 10 most popular health apps and perform statistical analysis to determine interrater reliability.

While SA1 was addressed in the literature review, the second stage in this mixed-methods analysis of health technology and disordered eating behaviors involved the development of a rating scale based on eating disorder screening criteria. The development of this tool was guided by grounded theory, which holds that gathering, studying, and categorizing data leads researchers to discover social processes and formulate theories\textsuperscript{30}. In this study, the extensive review of literature guided the development of a screening tool, the Awareness of Disordered
Eating Promotion in Technology (ADEPT), to determine and categorize the social process of disordered eating behaviors in apps. Inductive reasoning was utilized to search for patterns of DE behaviors in health technology and develop a theory regarding the presence of DE in popular apps. The ADEPT questions were categorized based on themes, with four themes emerging: Restriction, Weight Stigma, Exercise, and Compensatory Behaviors. Each question of the ADEPT rated the specified risk factor on a Likert-type scale of 1-5. Answers were totaled within each section to yield a domain score for Restriction, Weight Stigma, Exercise, and Compensatory Behaviors, in addition to an overall Total score of all the categories together.

In SA3, the screening tool underwent pilot testing for clarity and comprehension when three trained raters used it on three apps (selected based on downloads and inclusion criteria). No data collection occurred in SA3, as this pilot test was exclusively for feedback and development. The raters were three first-year graduate students in Northern Illinois University’s Master of Nutrition/Dietetic Internship program who were gaining a specialized certification in eating disorder treatment. They all held bachelor’s degrees in nutrition and had completed 1.5 semesters of graduate work prior to participation. The raters were recruited by email, based on their participation in the Graduate Certificate in Eating Disorders and Obesity.

The ADEPT was developed and described in such a way that raters needed a foundational understanding of eating disorder behaviors but did not need prior experience with app rating in order to use it effectively and accurately. After completion of the pilot test, it was found that one of the trained raters was not able to proceed with the study due to time constraints. Therefore,
SA4 was revised to include the remaining two trained raters who assessed the 10 most popular health apps.

After initial testing and revision, the ADEPT proceeded to SA4 for testing of the 10 most popular health apps meeting inclusion criteria (based on downloads) by the same two trained raters who performed the initial test. Raters used the app for at least 10 minutes, exploring the features while analyzing images, messages, and tools according to the screening tool. After the 10-minute trial period, raters completed the ADEPT based on their experience. The entire process of viewing the app and completing the tool took less than 30 minutes per app, with a goal completion time of 20-30 minutes per app.

Analysis of results yielded interrater reliability and internal consistency to determine if the new screening tool was valid and reliable. The goal was to achieve a value of 0.7 agreement among the raters. The data collected were cross-sectional, collected at one point in time, and not followed for any period. There was no intervention in this study; the purpose was to assess existing research regarding health apps and DE in order to develop and validate a rating scale of health app risk for DE. Refer to Appendix G for the research timeline.

Apps of Interest

This study focused on health apps selected from the iTunes store’s most popular health and fitness apps. These apps typically address weight loss and dieting; therefore, they are likely to contain DE risk factors. Common features include calorie tracking, compensatory exercise, or restriction of food groups. The apps were selected based on number of downloads from the 200
most popular Health and Fitness app list and the Healthy Eating list in April 2019, with the following inclusion criteria: available in English, free, and including facets of dieting, weight loss, body size/shape, or exercise to lose weight. Fitness apps were included if they incorporated some aspect of nutrition, calorie counting, or dietary suggestions. Exclusion criteria included apps strictly focused on mental health, sleep, hydration, or meditation. Apps that relied exclusively on supporting technology such as smartwatches or specific program memberships were excluded. The iTunes app store Health and Fitness category was searched within the parameter of free health and fitness apps and the results were screened based on the inclusion criteria. Apps were selected or excluded after the author read through the description and downloaded each potential app to determine if it met the inclusion criteria. A complete list of the selected apps is available in Appendix B.

Procedures

The ADEPT (Appendix C) was used to determine the extent to which popular health apps contained risk factors for DE. The researcher developed the tool so as not to require existing knowledge of app rating, enabling this study to be replicated easily by eating disorder practitioners and the scale adapted to wide usage. The raters had a foundational knowledge of DE risk factors, signs, and symptoms from at least one graduate-level course on eating disorders; therefore, they did not require additional training on eating disorder awareness. The raters received training on how to accurately complete the ADEPT. After training, a pilot test of the ADEPT was conducted by the three raters using three apps from the list in Appendix B.
Pilot Testing

The three raters individually interfaced with the same three apps for a 10-minute trial period, after which time they completed the ADEPT in order to produce a score for each app. After the pilot test, raters convened with the researcher to compare scores and discuss ambiguities in their results. The focus of this initial test was to discover inconsistencies in the language used or misinterpretations of the screening tool. The discussion after the pilot test was guided by the feedback questionnaire in Appendix D, and the ADEPT was adapted according to the feedback. At this step, one rater left the study and the procedure was adjusted to reflect the remaining two raters.

Testing

After being revised based on rater feedback, ADEPT version 2 was sent to the raters along with a training refresher based on feedback from the pilot test. This training clarified the procedure and can be found in Appendix E. The scale was used to test the 10 most popular health apps (based on number of downloads for qualifying apps in April 2019, as determined by inclusion criteria). The two trained raters conducted the test by using each app for at least 10 minutes. After the trial period, the raters used the updated ADEPT to rate the apps with the goal of achieving high levels of interrater reliability and internal consistency to validate the ADEPT.

Data Collection

Results of the screening were entered into IBM Statistical Package for the Social Sciences (SPSS) and statistical analysis was run to determine the Cronbach’s alpha, Pearson’s
correlation coefficient, and $p$ values for the ADEPT, in addition to basic descriptive statistics such as mean, frequency, and standard deviation. These statistical measures determine if the screening tool is reliable between raters and yields consistent results.

Outcome Measures

The ADEPT is based on eating disorder screening tools\textsuperscript{31} and formatted according to app rating scale procedures\textsuperscript{6}. Outcome measures include weight or shape concerns, weight stigma, compulsive eating or exercise, and restriction. These measures are dependent variables, contingent upon app content both implicit and explicit. These and similar behaviors are noted as risks for development of DE/EDs and as such are included on the NEDA eating disorder screening tool. The desired Cronbach’s alpha is 0.70 or higher based on established reliability standards\textsuperscript{32}.

Data Analysis

Table 1 outlines the outcome measures dependent on health app content alongside the statistical tests used to analyze results.
Table 1: Data Analysis Methods

<table>
<thead>
<tr>
<th>Outcome Measures</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight/shape concern</td>
<td>Internal Consistency by Cronbach’s alpha with desired level of 0.70 or greater</td>
</tr>
<tr>
<td>Discouragement of weight gain</td>
<td></td>
</tr>
<tr>
<td>Dieting</td>
<td></td>
</tr>
<tr>
<td>Value based on weight/image</td>
<td>Strength of correlation by Pearson’s correlation coefficient</td>
</tr>
<tr>
<td>Weight stigma</td>
<td></td>
</tr>
<tr>
<td>Compulsive eating or exercise</td>
<td></td>
</tr>
<tr>
<td>Purging</td>
<td></td>
</tr>
<tr>
<td>Caloric restriction</td>
<td></td>
</tr>
<tr>
<td>Disconnection from eating/appetite</td>
<td></td>
</tr>
<tr>
<td>Avoidance of food groups</td>
<td></td>
</tr>
<tr>
<td>Significant weight loss</td>
<td></td>
</tr>
<tr>
<td>Body-based worth</td>
<td></td>
</tr>
</tbody>
</table>

Monitoring of Data/Safety

The study does not contain human subjects, therefore there was no need to keep data confidential or anonymous. Similarly, participating in the scale cannot cause harm to the apps. The raters, however, could be exposed to disordered eating risk factors through app usage and analysis. Raters were therefore required to complete the NEDA Eating Disorder Screening prior to participating in the study, with the stipulation that anyone flagged at risk would be referred to
treatment and excluded from the rating panel. All raters were deemed safe to proceed with the study and informed that if at any time throughout the process they struggled with disordered eating thoughts or behaviors they should discontinue participation.

Summary of Methods

In reviewing current literature regarding DE/ED risk and health technology usage, it appears that common DE behaviors such as restriction and compensation are encouraged in health apps. Research suggests that many health apps are not evidence based\textsuperscript{5}, and usage of social media is correlated with eating disorder symptoms and body dissatisfaction\textsuperscript{18}. Evidence also suggests that media exposure and peer pressure are major risks for eating disorder development\textsuperscript{20}, and health apps often utilize community features to create accountability, which closely mirrors peer pressure by creating an environment rife with weight stigma and negative media exposure.

After examining the evidence and lack of research regarding health apps and DE, this study created a screening tool to assess popular health apps for DE risk factors. The ADEPT screening tool underwent initial testing on three apps by three trained raters. Based on rater feedback from the initial test, the ADEPT version 2 was developed to ensure clarity and consistency among raters. ADEPT version 2 was used to rate the 10 most popular health and fitness apps to determine their risk for DE. The results of this test were entered in SPSS and analyzed for internal consistency, interrater reliability, and basic descriptive statistics. These results indicated the reliability and consistency of the proposed screening tool and determine its use in future research.
Similar research has been conducted with rating app quality\textsuperscript{6}, and the present study builds upon that foundation to explore the risk for DE from health app usage. The ADEPT is adapted from validated eating disorder screening tools used for human subjects and influenced by existing mobile app rating studies. This study indicates the risk level of health apps for DE, providing direction for mobile health technology interventions. If popular health apps perpetuate DE behaviors, it is a vital public health issue to identify high-risk apps, develop DE-sensitive apps, and educate app users about DE risk.
CHAPTER IV: RESULTS

After the two raters completed testing of the top 10 health apps, their results were submitted to the researcher for statistical analysis using IBM SPSS to determine interrater reliability and internal consistency. The purpose of this statistical analysis was to validate the ADEPT tool by achieving internal consistency by Cronbach’s alpha of 0.70 or greater. This chapter presents each of the four domains within the Awareness of Disordered Eating Promotion in Technology (ADEPT) screening tool individually for discussion of data in addition to providing an overall analysis of the ADEPT. The domains include Restriction, Weight Stigma, Exercise, and Compensatory Behaviors.

Of note, each question in the ADEPT offered five possible answer choices based on a Likert-type scale divided into the following levels: none, mild, moderate, strong, or extremely strong. A lower score indicates less risk for disordered behavior; a higher score indicates greater risk for disordered behavior. Table 2 shows the mean (M) ± standard deviation (SD) and p value for each domain as well as for the overall total, which represents a summation of scores from the four domains. Of note, the p value indicates significant difference between raters for each of the categories except Compensatory Behaviors. The scores in Table 2 represent the comparison between the domains for all apps, divided by each rater. The M indicates that Rater 1 consistently assigned lower values than Rater 2 in every category except for Compensatory
Behaviors. This difference contributes to the variability in $p$ values. Results show a large variation in SD for many categories.

Table 2: Mean ± SD and $p$ Values

<table>
<thead>
<tr>
<th>Raters</th>
<th>Restriction</th>
<th>Weight Stigma</th>
<th>Exercise</th>
<th>Compensatory</th>
<th>Total Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td>Mean ±</td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
</tr>
<tr>
<td></td>
<td>$p=0.03^*$</td>
<td>$p=0.001^{**}$</td>
<td>$p=0.872$</td>
<td>$p=0.000^{***}$</td>
<td>$p=0.016^{**}$</td>
</tr>
<tr>
<td>Rater 1</td>
<td>26.90 ± 8.62</td>
<td>11.90±4.23</td>
<td>12.20±4.44</td>
<td>7.60±2.01</td>
<td>58.60±11.49</td>
</tr>
<tr>
<td>Rater 2</td>
<td>29.90 ±6.40</td>
<td>19.30±5.40</td>
<td>16.20±1.93</td>
<td>7.50±2.50</td>
<td>72.90±7.33</td>
</tr>
</tbody>
</table>

P = 0.05 * $p= 0.01^{**}$ $p=0.001^{***}$

Restriction

ADEPT Section I focused on restrictive behaviors and included 10 statements addressing calorie counting, food tracking, fasting/severe restriction, and the degree to which food obsession was encouraged. An example statement follows: “App encourages users to deliberately limit food intake to influence weight or shape.” The overall Cronbach’s alpha for restriction across all 10 apps was 0.938, indicating excellent internal consistency among raters in the Restriction domain. This high internal consistency confirms that the ADEPT is an accurate measure of the
variable of interest: restrictive behaviors. Pearson’s correlation coefficient was $r=0.922$, indicating large positive association\textsuperscript{34} between the two raters. The M $\pm$ SD for the two raters were $26.90 \pm 8.62$ and $29.90 \pm 6.40$ for Raters 1 and 2 respectively. Figure 1 compares the responses of the two raters on each of the 10 apps within the Restriction domain. Restriction totals were within one point for six of the ten apps: Lifesum, MyPlate, Carb Manager, Verv, MyFitnessPal, and LoseIT, which accounts for the high correlation ($r=0.922$).

Figure 1: Comparison of apps by raters: Restriction.

Weight Stigma

ADEPT Section II included six questions and addressed weight-biased images, language, and self-perception. Questions focused on how the app treated varying body sizes and the degree to which weight was associated with worth. An example statement follows: “App encourages user to think about weight or shape to what degree?” Overall Cronbach’s alpha for weight stigma
in all 10 apps was 0.714, indicating acceptable internal consistency among raters in the Weight Stigma domain. Pearson’s correlation was \( r = 0.571 \), indicating large positive correlation between the two raters. The M ± SD for the two raters were 11.90 ± 4.23 and 19.30 ± 5.40 for Raters 1 and 2 respectively. Figure 2 compares the responses of the two raters on each of the 10 apps within the Weight Stigma domain. Rater 2 consistently scored higher on all ten apps, but this was especially true in the Weight Stigma domain. Weight Stigma totals were within one point for only two of the ten apps: MyPlate and Carb Manager, which contribute to the large positive correlation (\( r = 0.571 \)).

![Comparison of Apps by Raters: Weight Stigma](image)

**Figure 2:** Comparison of apps by raters: Weight Stigma.

**Exercise**

ADEPT Section III addressed rigid or compensatory exercise behaviors and contained five questions. The questions assessed behaviors for signs of compulsive movement, over-
exercise, and guilt or shame as motivation for exercise. An example statement follows: “App suggests exercise in a driven or compulsive manner as a means of controlling weight, altering shape, or burning off excess calories.” Cronbach’s alpha for exercise in all 10 apps was 0.353, indicating unacceptable internal consistency among raters in the Exercise domain. Pearson’s correlation was $r=0.293$, indicating small positive correlation between the two raters. The $M \pm SD$ for the two raters were $12.20 \pm 4.44$ and $16.20 \pm 1.93$ for Raters 1 and 2 respectively. As evidenced by the graph in Figure 3, it is clear that the two raters were inconsistent in their ratings, hence the unacceptable Cronbach’s alpha and small positive association. Exercise totals were within two points for only two of the ten apps: Carb Manager and Verv, which contributed to the small positive correlation ($r=0.293$).

![Comparison of Apps by Raters: Exercise](image)

Figure 3: The above graph illustrates exercise totals for each of the 10 apps by raters.
Compensatory Behaviors

ADEPT Section IV focused on behaviors that endeavored to “make up for” eating habits or to relieve food guilt and included four questions. Statements assessed whether compensation was driven by in-app suggestions and if these behaviors could contribute to obsessive health behaviors. An example statement follows: “App suggests controlling weight or shape with compensatory measures.” Cronbach’s alpha for compensation in all 10 apps was 0.785, indicating acceptable internal consistency among raters in the Compensatory Behaviors domain. Pearson’s correlation of $r=0.662$ indicated large positive correlation between the two raters. The $M \pm SD$ for the two raters were $7.60 \pm 2.01$ and $7.50 \pm 2.51$ for raters 1 and 2 respectively. As evidenced by the graph in Figure 4, both raters remained close in their scoring of compensatory behaviors, hence the large positive association ($r=0.662$). Compensatory Behaviors totals were within three points for all ten apps and within one point for six apps: Lifesum, MyPlate, Noom, BetterMen, BetterMe, and LoseIT.
Figure 4: Comparison of apps by raters: Compensatory Behaviors

Adept Overall Totals

After rating the four sections individually, a total score was calculated when each rater added up the scores from each domain. This score is referred to as the overall Total in data analysis and accounts for sum disordered eating risk behaviors across all categories. The Total Cronbach’s alpha was 0.829, indicating overall good internal consistency among raters. Pearson’s correlation was $r=0.781$, showing overall large positive correlation between the two raters. The overall $M \pm SD$ for the two raters were $58.60 \pm 11.49$ and $72.90 \pm 7.33$ for Raters 1 and 2 respectively. As evidenced by the graph in Figure 5, both raters assigned overall totals within 10 points for four of the ten apps: MyPlate, Carb Manager, Verv, and MyFitnessPal, hence the large positive association ($r=0.781$).
Figure 5: Overall Totals
CHAPTER V: DISCUSSION

While Specific Aims 1-3 were fulfilled through the review of literature, study design, and pilot testing, Specific Aim 4 was fulfilled through the above data analysis. After the ADEPT was tested using the revised tool and a reduced number of two raters, data analysis yielded good internal consistency and large positive correlation between raters. Considering the data by category, the Restriction results showed excellent internal consistency and large positive correlation ($\alpha=0.938$, $r=0.922$). Weight Stigma results showed acceptable internal consistency and large positive correlation ($\alpha=0.714$, $r=0.571$). Exercise results showed unacceptable internal consistency and small positive correlation ($\alpha=0.353$, $r=0.293$).

A possible reason for the poor results in the Exercise domain could be due to inclusion criteria being too inclusive, allowing exercise-focused apps which lacked core diet-centered DE behaviors. Alternatively, the Exercise category itself could prove too broad and distant from the central, diet-focused behavior of DE which the ADEPT set out to assess. Compensatory Behaviors results showed acceptable internal consistency and large positive correlation ($\alpha=0.785$, $r=0.662$). The Compensatory Behaviors category was the only domain to achieve a $p$ value indicative of no significant difference between raters, while $p$ values for the other sections indicated significant difference between raters. The overall Total Cronbach’s alpha and
Pearson’s $r$ indicate ADEPT has good internal consistency and large positive correlation between raters ($\alpha=0.829$, $r=0.781$).

While the Exercise subsection did not provide internal consistency and correlation to the significance levels set in this study (0.7), the remaining three categories delivered acceptable consistency and correlation. Moreover, the ADEPT tool as a whole delivered both good internal consistency and large positive correlation between raters. These overall scores met the desired Cronbach’s alpha of 0.70 established by the author; therefore, the ADEPT satisfied the initial testing requirements.

The results indicate the ADEPT as a promising tool for future use in determining disordered eating risk from health apps. As discussed in the review of literature, the more time individuals spend using mobile applications with fitspiration content, the more likely they are to report disordered eating behavior and compulsive exercise$^{14}$. Research suggests that mobile health app usage is a positive predictor for stronger disordered eating symptomatology$^{14}$. Similarly, more frequent usage of social media contributes to body image dissatisfaction and eating disorder symptoms$^{26}$. These studies support the view that exposure to media such as health apps can promote disordered eating. The ADEPT provides a valuable tool for assessing risk factors in popular health apps and could be used to decrease exposure to harmful apps.

Strengths and Limitations

The ADEPT was developed based on the NEDA Eating Disorder Screening alongside other validated screening tools, all respected and widely used standards for assessing ED risk in
participants. The ADEPT relies on the support of multiple, validated ED screening tools, and these served as guidance in the development and utilization of the ADEPT. This study did not hazard exposure of human subjects to DE risk; therefore, by rating the apps directly rather than gathering data from human subjects, the researcher reduced the risk of encouraging DE behaviors.

The raters were not required to be experts in the field of mobile application research because the ADEPT was designed for use by raters with no previous app rating experience. This enables ease of replicating results so that any interested parties can use the screening tool after completing perfunctory training. The 10-minute trial period allowed the rater a general overview of the app at a broad level and reflects practices used in other app rating studies. Data is lacking on DE risk from health app usage; therefore, this data can contribute to the foundation of app and DE risk research. The overall ADEPT scores satisfied the desired level for Cronbach’s alpha of 0.70 for good internal consistency.

The limitation of basing the study on multiple validated ED screening tools lies in the fact that these tools were developed for human subjects self-assessing their behaviors and emotions regarding food and their bodies. Adapting the NEDA screening to apps rather than humans, therefore, may have failed to detect risks embedded in the app design or caused risk to be assigned to the app where none should be. Research is lacking on eating disorder risk from app usage, so there is not a clear protocol for how to assess this risk or account for the difference between a human subject and a mobile app. This study adapted procedures previously used for assessing ED risk in humans and copied procedures from app quality rating studies rather than
ED risk studies. While app raters were not required to be ED professionals, in order to accurately apply the ADEPT, raters must have insight into the subtle ways which images or wording could encourage DE behaviors.

While the 10-minute trial period for each app enabled the tool to be widely used with results readily replicated, raters may have overlooked subtle signs of DE due to lack of time to assess the app or lack of knowledge regarding DE/EDs. The raters’ lack of experience with the various subtleties of DE/ED behaviors could have led them to overlook aspects of the app which could contribute to DE risk. The ADEPT was designed for use after a 10-minute trial period of the app. This time was not adequate to fully explore every feature on selected health apps, meaning that risk factors may have been hidden in facets of the app not explored by the rater or not accessible until further use. If the raters were personally experiencing DE/EDs, they may have failed to identify risk factors in the apps because those behaviors seemed normal to them. The training prior to the rating process, however, was designed to educate raters about these risks and screen them for DE/EDs.

The rater’s variability in answers in the Exercise domain produced unacceptable reliability and correlation. While this poor result does not detract from the overall reliability of the scale, it could indicate need for refining app inclusion criteria to focus explicitly on diet apps. This variability could also indicate a need for either removing entirely or reworking the exercise section. Additionally, the \( p \) values showed significant difference between raters for every category except Compensatory Behaviors, in large part due to the fact that Rater 2 consistently scored apps higher than Rater 1. The variability in responses between raters does, however,
provide insight into the fact that while app content may not trigger disordered eating habits in some users, other users could be triggered by the same content.

Additional consideration should be given to rater fatigue during the testing process. Raters were given a two-week period during which to complete the 10-app test. Raters could complete the 10 apps all at once or space them out in their own timing. This variability in rating conditions could produce rater fatigue and variation in circumstances between the raters. Such variability could not be controlled for in the present study, but in future research raters could complete the ADEPT in a controlled, uniform environment to ensure minimal variability.

Implications for Dietetic Practice

Widely used health apps may pose a risk for DE; therefore, it proves crucial to develop educational materials to inform health app users of the risk and direct them towards preventative resources. Future efforts could include development of DE-sensitive apps to support the wellness of users without compromising their eating competency. A gap may exist in the market for apps which promote health while screening for eating disorders and offering DE prevention tools such as local services, websites, and screenings.

The ADEPT could serve as a valuable guide for eating disorder professionals to screen popular health apps before making recommendations to clients. It could also serve as a tool in eating disorder recovery for clinicians to work through alongside clients, helping participants to recognize DE behaviors in their media. The ADEPT could be particularly useful for Registered Dietitian Nutritionists (RDNs) to refer to in practice. Due to the popularity of health apps, RDNs
will likely encounter clients utilizing apps and receive questions about which apps to use. By incorporating the ADEPT into their practice, RDNs can identify apps that pose risk for DE behaviors and suggest lower risk health apps.

Future Research

There is little known data regarding the DE risk associated with health app usage; therefore, this study could contribute to future research regarding such risk. In future research, the ADEPT could be combined with quantitative data gathered from human subjects regarding their DE behaviors before and after health app usage to provide a more detailed picture of the risk these apps pose. This study could be expanded by using a larger pool of raters to repeat the test with the 50 most popular health apps and results could be analyzed against data from the present study. Alternatively, a panel of raters could utilize the ADEPT to assign risk scores to popular apps. These scores could be distributed in educational materials to promote awareness for DE behaviors in apps and rank apps based on DE risk. By using a large rating panel to assess many of the most popular apps, a validated ADEPT index of standardized app scores could be established for distribution to health professions so that they can access app scores without having to repeat the test. The purpose of the current study was to develop and pilot test the ADEPT, but further studies could utilize data from the ADEPT to validate the tool.

Conclusion

Common risk factors for DE include dieting, compulsive exercise, and compensatory behaviors\textsuperscript{20,25}. Common behaviors suggested in health apps include dieting, forcing oneself to
participate in frequent or excessive exercise\textsuperscript{35}, and using restriction or exercise to compensate for “cheat days”. The striking resemblance between DE and health app content indicates that popular health apps could pose a risk for DE.

The majority of existing research focuses on DE behaviors from social media usage, whereas the current study extended these established risk factors to develop the ADEPT for health apps. During research, the following categories of DE behavior emerged: Restriction, Weight Stigma, Exercise, and Compensatory Behaviors. The findings of this study show that the 10 most popular health apps do contain DE behaviors, indicating a need for new health apps that promote wellness without increasing risk for DE. Further research is needed to validate the ADEPT; however, results indicate that it can be used in future studies to assess apps for DE risk. RDNs and other practitioners in the ED field may also utilize the ADEPT to screen apps before suggesting them to clients.
REFERENCES


33. Gliem JA, Gliem RR. Calculating, Interpreting, and Reporting Cronbach’s Alpha Reliability Coefficient for Likert-Type Scales.

APPENDIX A

POPULAR HEALTH APPS
Figure 6: Most popular health and fitness apps in the United States as of May 2018. This graph gives information on active reach of the most popular health and fitness apps in the United States, sorted by monthly active users. As of May 2018, Fitbit was ranked first with 27.4 million unique U.S. users. MyFitnessPal was ranked second with a 19.1 million user strong audience.
Figure 7: Market reach of the most popular mobile app categories in the United States as of July 2018. This graph gives information on the market reach of the most popular app categories in the United States as of July 2018, 64.5% mobile market reach of health apps among users in the U.S.
APPENDIX B
SELECTED HEALTH APPS
Pilot Test:

1. Keto Diet (#50 hint: use the ‘I’ tab to navigate diet info)
2. Zero – fasting tracker (#51)
3. My Diet Coach (#54)

10 App Test:

4. myfitness pal (#2, MyFitnessPal.com)
5. Better men (#5, Genesis Technology Partners)
6. better me (#6, Genesis Training & Diet Plan)
7. fitbit (#10, FitBit, Inc)
8. loseit (#12, FitNow)
9. Carb manager - keto diet app (#14, Wombatt Apps LLC)
10. Weight loss fitness by verv (#16, Verv, Inc, Hint: refer to tips section for nutrition info)
11. Lifesum: diet and macro tracker (#21, LifeSum AB)
12. Myplate calorie counter (#22, LIVESTRONG.com)
13. Noom (#26, Noom, Hint: don’t follow the prompts for ‘get customized course’, choose ‘maybe later’ for free option)
APPENDIX C
ADEPT
Awareness of Disordered Eating Promotion in Technology
ADEPT Screening Tool (Version 2)

App Name & Version: _____________________________ Platform □ iPhone □ Android
Rater Number: ________________ Time Spent Reviewing: ____________________

Brief Description of App:

_____________________________________________________________________________
_____________________________________________________________________________

App Focus: (Select all that apply)

□ Weight loss       □ Wellness
□ Diet               □ Health Education
□ Exercise           □ Other (please specify)
□ Calorie counting

Technical Specifications: (Select all that apply)

□ Connects with social media  □ Provides community feature
□ Connects with smart device □ Sends notifications

Scoring: Add up the total value for each category based on the number circled. For example, if option 3 is selected from section 1, question 1, the numeric value 3 will be assigned to that question and should be added to other values in the restriction section to calculate a score for restriction. To calculate the total score, add the scores from each section together. The number in parentheses represents the number of questions in each category/total number of questions.

Section A Restriction (10): _______

Section B Weight Stigma (6): _______

Section C Exercise (5) _______

Section D Compensatory Behaviors (4): _______

Total Score (25): _______

Directions: This tool assess promotion of disordered eating behaviors in four categories: Restriction, Weight Stigma, Exercise, and Compensatory Behaviors. Select the numbered option which best reflects the strength of the disordered behavior promoted through app content:
1: (None) indicates app does not contain the specified behavior

2: (Mild) indicates there may be a reference to the behavior or a side option to explore it

3: (Moderate) indicates the behavior is available in the app or is discussed with the option to try it out

4: (Strong) indicates the behavior is recommended

5: (Extremely Strong) indicates the behavior is either mandated, penalized if done/not done, or the behavior is recommended as in level 4 but with additional conditions

Section A (10 Questions)

Restriction: dieting, calorie counts, control of food intake, calculation of calorie needs, advising users to consume <1,200 kcal, food group removal (i.e. avoid carbs).

1. Calorie calculation: app calculates calorie needs for user and restricts their intake to a strict calorie value or users select a calorie value as their maximum intake.
   1. None: No calculation of calories
   2. Mild: Mentions calculating calories (i.e. gives an example that some people calculate calories or refers to an article about calorie counting)
   3. Moderate: Users can choose to calculate calories, but it is not required or central to app functioning
   4. Strong: Calculates calories for user
   5. Extremely Strong: Calculates calories for user then subtracts an amount for weight loss (i.e. calculates calorie needs as 1,600 kcal, subtracts 500 kcal/day for wt. loss)

2. Calorie counts: app utilizes calorie counts and users count their calories each day.
   1. None: No calorie counts
   2. Mild: Mentions calorie counting – offers food tracking as an elective/additional feature/function
   3. Moderate: Suggests users count their calories (i.e. as an optional technique or feature)
   4. Strong: Encourages users to count calories
   5. Extremely Strong: Users must count calories and are penalized if they exceed calorie count (i.e. loss of points or rewards)

3. Food tracking: food intake is recorded in the app and users are prompted to record what/when/how much they eat.
   1. None: Food intake is not tracked
   2. Mild: Mentions food tracking – offers food tracking as an elective/additional feature/function
   3. Moderate: Provides option for tracking food intake as a main feature
   4. Strong: Recommends users track their food intake (i.e. with reminders, notifications, incentives)
5. Extremely Strong: Users must track food intake and are penalized (i.e. loss of points/failure to gain points) if they don’t track their intake

4. Restricted intake: app encourages users to deliberately limit food intake to influence weight or shape (ex. consuming <1,200 kcal, or subtracting 500 kcal from estimated needs for weight loss)
   1. None: Restricted intake not addressed
   2. Mild: Refers to or mentions restricting intake
   3. Moderate: Provides option for users to restrict intake or suggests restriction as an option
   4. Strong: Recommends users restrict their intake (i.e. calorie restriction is a main feature/focus of the app)
   5. Extremely Strong: Users are penalized if they do not restrict their intake OR must restrict for the purpose of influencing weight/shape

5. Fasting: app encourages users to go without food for long periods of time (ex. 6 or more waking hours, intermittent fasting) in order to influence weight or shape.
   1. None: Fasting is not addressed
   2. Mild: Refers to or mentions fasting (i.e. an article about intermittent fasting is available)
   3. Moderate: Provides an option for users to fast
   4. Strong: Encourages or recommends that users try fasting or fasting as a main app function
   5. Extremely Strong: Recommends users try fasting in order to influence weight/shape

6. Food obsession: app encourages users to think about food, eating, or calories throughout the day or could credibly create repetitive thoughts of food, eating, or calories (ex. ‘think about what you are going to eat in advance’ or ‘keep track of calories throughout the day’).
   1. None: There is no chance of food obsession with this app
   2. Mild: Refers to repetitive thoughts about food/eating/calories (eg. casual mention)
   3. Moderate: Suggests users think of food/eating/calories throughout the day (eg. offering as an option)
   4. Strong: Recommends users should think of food/eating/calories throughout the day OR use of this app could lead to such repetitive thoughts
   5. Extremely Strong: Users should think about food/eating/calories throughout the day; repetitive thoughts about food are central to the app

7. Disconnection from appetite: app encourages disinterest in eating and disconnection from appetite (i.e. ignore hunger, don’t snack, don’t exceed daily calorie allotment, push through the urge to eat).
   1. None: Connection with appetite not implied or addressed
   2. Mild: Refers to or mentions ignoring appetite
3. Moderate: Provides option or suggestion to practice ignoring appetite (i.e. user could try ignoring stomach growling)
4. Strong: Recommends users should ignore appetite (i.e. “don’t give in to hunger”)
5. Extremely Strong: Ignoring appetite is central to app function and app provides techniques for users to disconnect from or push through hunger

8. Elimination: app suggests users avoid certain foods (i.e. eliminate bread) or entire food groups (i.e. carbohydrates) for health reasons, weight loss, or shape alteration. Note: this excludes elimination for those with medically diagnosed food allergies.
   1. None: Elimination is not addressed
   2. Mild: Mentions elimination of foods/groups
   3. Moderate: Option for users to try eliminating foods/groups
   4. Strong: Recommends users eliminate foods/groups (i.e. don’t eat bread)
   5. Extremely Strong: Users should eliminate foods/groups with the intent of influencing health/weight/shape (i.e. eliminate bread to help you lose x pounds)

9. Food Morals: app promotes or implies a simplistic view of food as good or bad, wherein good foods promote health while bad foods are detrimental to health. This could be explicitly calling foods bad or recommending to ‘avoid refined grains’ or ‘avoid junk food’. (for example: eating a ‘bad food’ such as pizza will directly negatively impact health, OR consistently eating ‘bad foods’ equals being unhealthy).
   1. None: Food morals are not implied or addressed
   2. Mild: Refers to good foods/bad foods (i.e. articles about ‘bad foods)
   3. Moderate: Suggests the concept of good/bad foods
   4. Strong: Recommends consumption of good foods, avoidance of bad foods
   5. Extremely Strong: Recommends consumption of good foods, penalizes or requires compensation for bad food consumption

10. Food as nourishment: app promotes intuitive eating (IE) and responding to hunger/fullness cues (H/F) and consuming a variety of foods/groups (not as a weight loss tool/incentive).
    1. Extremely Strong: Recommends viewing food as nourishment, IE, or H/F as the alternative to dieting/weight loss
    2. Strong: Recommends viewing food as nourishment, IE, or H/F
    3. Moderate: Suggests the concept of food as nourishment, IE, or H/F
    4. Mild: Refers to food as nourishment, IE, or H/F
    5. None: No information/discussion of food as nourishment, IE, or H/F

Section B (6 Questions)

Weight Stigma: weightism, weight bias, anti-fat bias, weight-based discrimination; promotes negative view of larger bodies or higher weights; targeting individuals because of their weight.

1. Weight centric: app encourages user to think about weight or shape to what degree:
   1. None
2. Mild
3. Moderately/Regularly
4. Strong
5. Extremely Strong (makes it difficult to focus on life)

2. Fat phobia: app content could make users scared of gaining weight, either through explicit content or implied badness associated with being fat such as: dramatic titles/articles which awfulize fat; before/after content which promotes fear of returning to the before stage; or displaying/speaking of fat bodies as flawed.
   1. None: Fat phobia not implied or addressed
   2. Mild: Some additional/elective features show content which could be perceived as fat phobic
   3. Moderate: Content or attitudes are suggestive of fat phobia
   4. Strong: Encourages fat phobia implicitly
   5. Extremely Strong: Recommends or promotes fat phobia explicitly (e.g. uses fear of fatness as a motivator for weight loss behaviors)

3. Negative self-perception: app usage could reasonably create or fuel a desire to lose weight or alter body shape based on negative view of large bodies as promoted in app or through messages about achieving a 'normal weight'.
   1. None: No chance that negative self-perception could be promoted by app usage
   2. Mild: Mentions content which could promote negative self-perception
   3. Moderate: Suggests content which could promote negative self-perception (i.e. negative self-perception is available/possible, but not central)
   4. Strong: Recommends content or behaviors which promote negative self-perception (i.e. negative self-perception is a central function of app)
   5. Extremely Strong: Encourages negative self-perception as a motivator or tool for weight loss (i.e. messaging such as “try to achieve a normal weight of ‘x’ pounds)

4. Weight and worth: app usage could suggest that weight or shape influence how the person view themselves or how others view them; weight is connected to worth through content/behaviors/images (e.g. images which portray before and after pictures); implied lower value to large bodies/higher value to small bodies; places undue/exaggerated amount of value on weight.
   1. None: Weight is not associated with worth
   2. Mild: Mentions an association between weight and worth
   3. Moderate: Suggests an association between weight and worth
   4. Strong: Recommends or creates views or behaviors which associate weight with worth
   5. Extremely Strong: Central messaging that weight is a measure of worth by promoting weight loss to increase value to self or others
5. Weight-biased imagery: app displays pictures which could produce fear of fatness or perpetuate weight stigma (ex. large people sitting on the couch/thin people being active); before/after pictures to motivate weight loss; large bodies associated with health complications/thin bodies associated with health.

1. None: No weight-biased images present
2. Mild: Images could refer to/imply weight bias
3. Moderate: Images suggest weight bias
4. Strong: Images encourage or create weight bias
5. Extremely Strong: Images encourage weight bias and use weight bias to promote behavior change

6. Thin bias: app content implies positivity related to thinness; thin bodies are portrayed in a positive/healthy light; thin bodies are used to sell a behavior (ex. Thin people drinking water or running); thin bodies are shown achieving goals.

1. None: Content does not portray thin bodies as superior to fat bodies
2. Mild: Refers to/implies thin bias in few content items
3. Moderate: Suggests thin bias in some content
4. Strong: Encourages thin bias in most content
5. Extremely Strong: Encourages thin bias as a motivator for weight loss/behavior change

Section C (5 Questions)

Exercise: exercise includes planned activity, activities of daily living, and app suggested or created exercise plans. Definition: “Exercise is any bodily activity that enhances or maintains physical fitness and overall health and wellness.”

1. Exercise compulsion: app suggests exercise in a driven or compulsive manner as a means of controlling weight, altering shape, or burning off excess calories. Compensatory exercise is defined herein as exercise to reduce or offset the effects of eating or prior lack of exercise, performed rigidly without regard to injury, need for rest, or desire to exercise; exercise can be used to mask/numb anxiety or stress; exercise as a primary coping mechanism.
1. None: Exercise compulsion/as a compensatory mechanism is not addressed or implied
2. Mild: Mentions behaviors which are compulsive or compensatory exercise
3. Moderate: Suggests using exercise as a compulsive or compensatory behavior, offers compulsive/compensatory exercise as an option to try out
4. Strong: Recommends using exercise as a compulsive or compensatory behavior
5. Extremely Strong: Exercise is primarily viewed as a compensation for eating, is promoted in a compulsory manner, or is used to deduct calories from daily totals

2. Excessive exercise: app encourages excessive exercise, which is exercise that is accompanied by guilt, undertaken solely to manipulate weight or alter shape, with the purpose of burning off calories, consistently pushing to the limits of endurance, interfering with other commitments, or undertaken despite injury.
   1. None: Excessive exercise is not addressed or implied
   2. Mild: Mentions behaviors which could create or promote excessive exercise
   3. Moderate: Option for users to exercise excessively
   4. Strong: Recommends users should exercise excessively
   5. Extremely Strong: Displays excessive exercise as the norm for exercise

3. Rigid exercise: app creates a rigid structure of exercise which could create reasonable distress in the user if they cannot follow it exactly; exercise routines are inflexible and must be followed exactly; exercise is habitual not mindful; exercise routine could socially isolate user; workout doesn’t ‘count’ unless it’s a certain time or intensity.
   1. None: No rigid rules or structure in exercise
   2. Mild: Mentions the possibility of exercise as rigid
   3. Moderate: Suggests rigid exercise routines as an option
   4. Strong: Recommends users create/follow rigid exercise plans
   5. Extremely Strong: Mandatory adherence to a strict exercise plan with missed workouts penalized

4. Exercise Guilt: app creates an adverse response to inability to exercise; promotes guilt when days are missed or when exercise is not to a certain degree of exertion or length of time; exercise is done out of obligation; motivated by guilt/dislike for body size/shape; utilizes guilt as a motivator to exercise; notifications about needing to exercise contribute to guilt; (i.e. exercise will make you feel better about your slip-ups).
   1. None: No chance that users could feel guilt for missing exercise
   2. Mild: Mentions guilt might result from failure to exercise
   3. Moderate: Suggests users could feel guilt for not exercising
   4. Strong: Addresses or implies users should feel guilt for not exercising
   5. Extremely Strong: Implies users should feel guilt for not exercising and uses guilt as a motivator for exercising
5. **Mindful Movement**: app promotes joyful movement motivated by general health improvement; users are encouraged or taught to listen to how their body feels (not focused on weight loss/body alteration).
   1. **Extremely Strong**: Recommends users listen to their bodies and move mindfully and teaches intuitive movement as the alternative to rigid exercise routines
   2. **Strong**: Recommends users move mindfully
   3. **Moderate**: Suggests mindful movement as an option
   4. **Mild**: Mentions mindful movement
   5. **None**: Does not promote or address mindful movement

---

Section D (4 Questions)

Compensatory Behaviors: behaviors to make up for or undo eating; behaviors to relieve guilt or anxiety associated with things such as eating, not exercising, or emotional distress. Behaviors could include but are not limited to: vomiting, exercising, restriction, laxative abuse, fasting, misuse of enemas or diet pills, or chewing and spitting.

1. **User-driven compensation**: app suggests controlling weight or shape with compensatory measures.
   1. **None**: Compensatory behaviors are not addressed or implied
   2. **Mild**: Mentions compensatory behaviors
   3. **Moderate**: Suggests compensatory behaviors as an option for users to engage in
   4. **Strong**: Recommends compensatory behaviors
   5. **Extremely Strong**: Recommends compensatory behaviors as the major method for controlling weight or shape

2. **App-driven compensation**: app addresses binge eating, cheat days, or ‘slip-ups’ of diet or exercise pattern with the suggestion to engage in a compensatory behavior; compensation is the primary coping mechanism.
   1. **None**: No compensation is suggested for eating or exercise behaviors
   2. **Mild**: Mentions compensatory behaviors
   3. **Moderate**: Suggests compensatory behaviors as an option for users to engage in
   4. **Strong**: Recommends compensatory behaviors
   5. **Extremely Strong**: Recommends compensatory behaviors as the major method for making up for ‘cheats’, binges, or ‘slip-ups’

3. **Compensatory guilt**: app response or education regarding binges/slip-ups/cheat days could realistically create distress or guilt in users; information regarding binges/slip-ups/cheat days could create guilt which could produce compensatory behaviors; app subscribes to a specific pattern of behavior which could produce compensatory behaviors if the pattern is deviated from (i.e. users must eat only 2x daily; if they don’t a compensatory behavior could be performed to make up for eating an extra meal/snack).
1. None: No guilt is implied by information regarding ‘slip-up’ behaviors
2. Mild: Mentions users could feel guilt regarding ‘slip-up’ behaviors
3. Moderate: Suggests users feel guilt regarding ‘slip-up’ behaviors
4. Strong: Recommends users feel guilt regarding ‘slip-up’ behaviors
5. Extremely Strong: Recommends users feel guilt regarding ‘slip-up’ behaviors and use this guilt to drive compensatory behaviors

4. Health Obsession: App encourages users to obsessively engage in health behaviors; app could create or affirm a fear of ‘unhealthy’ foods or behaviors; behaviors could include: compulsive label checking, cutting out food groups, distress when ‘healthy’ foods aren’t available, or obsessive following of ‘healthy lifestyle’ apps or media.
   1. None: There is no chance of health obsession with use of this app
   2. Mild: Refers to repetitive health-focused thoughts (casual mention)
   3. Moderate: Suggests users think about health throughout the day or offers an option to think repetitively about health
   4. Strong: Recommends users think about health behaviors throughout the day OR app use could create repetitive health-obsessive thoughts
   5. Extremely Strong: Users should think about health throughout the day in an obsessive manner; health obsession is central to the app (e.g. used as motivation)
APPENDIX D

FEEDBACK FORM
ADEPT Feedback Form

1. Describe your overall experience with ADEPT:

2. Which specific questions did not make sense or were consistently difficult to answer?

3. Which questions were redundant?

4. Did you notice something when viewing apps that you couldn’t categorize in the tool?

5. What additional questions or areas do you think should be included?

6. How could the instructions and rating calculation directions be made clearer?

7. Are the categories clear? What do you need to clarify them?

8. Which questions were difficult to differentiate between the potential answers? Please provide question numbers.

9. What worked well and what didn’t?
APPENDIX E
RATER TRAINING
1. Please record all responses online and email them to me; this will make it easier for me to copy data over since I am performing data analysis on these results. You don’t have to email all 10 at once – please email as you complete and I can begin data entry that way.

2. **Don’t discuss** the apps or ratings with each other until after they are completed and submitted.

3. In response to difficulty staying consistent or choosing between options: keep in mind that you are using this app as if you are a clinician screening for a client to use. Imagine your client asked you if the app is appropriate for them to use or could trigger disordered behavior. You are looking out for things which might trigger your client. There is no right or wrong answer, rather it is based on your professional judgment from your knowledge of disordered eating.

4. Notice the addition of the ‘time spent reviewing’ line on the first page. Please record how long you interact with the app. Include the initial ~10-minute period and any time you refer to the app while you are filling out the screening. The goal is for all of you to spend similar amounts of time as this will produce a more standard experience. It seems that 20-30 minutes is the average time it takes to complete the ADEPT; this is great, but please try not to exceed 30 minutes because we don’t want to compare results from 20 minutes spent in the app with results from 45 minutes in the app. Similarly, if you are rushing through the screening then the app will likely receive a lower score because risk factors will not be adequately explored.

5. Please pay close attention to content; if you aren’t sure if an app addresses a particular question, double check. I noticed several instances of questions which were marked with 1 (none) even though the behavior was clearly present in the app after a brief interaction.

6. Notifications: I understand you will not actually receive any notifications during your brief trial. To evaluate notifications, check the settings to see how often notifications will be sent, or if it offers a way to tailor notifications. If there is a preview of notifications, notice the language and how that could make a user feel. Recognize that the notification content will align with app content; if the app is promoting disordered behaviors, the notifications will likely encourage similar behaviors.

7. Difficulty distinguishing between 4th and 5th levels: it seemed everyone was hesitant to assign an ‘extreme’ rating to an app, which is absolutely fine. To clarify, if the app has what seems even a touch of extreme, I would prefer you err on the side of extreme. Extreme behavior could be explicit or implicit – anything that makes you concerned the app could create/promote extreme disordered behavior. I understand there will likely not be a lot of extreme ratings given out – this is good. But don’t be scared to assign it! If you judge app content to have or imply extreme disordered behaviors, then give it a 5.

8. Per your suggestions, I added a question in the compensation section on orthorexia nervosa behaviors titled ‘health obsession’. Note that there is now a total of 25 questions.

9. If there are any apps you feel are not a good fit for the test, let me know and I will re-evaluate them. I am open to changing them out, but I need to hear from you on this.
10. Previously the 5 levels referred to frequency; the change in wording reflects that the levels are now based on strength. If a behavior is only mentioned/implied infrequently but is very dangerous or strong, it should receive a higher rating. Ask yourself if there is a chance that language or behaviors in the app could create the stated behavior for your imaginary client; if so, make sure to indicate it. The new levels are none, mild, moderate, strong, and extremely strong.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Assemble Research Committee</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proposal Defense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Revise Proposal &amp; Develop Screening Tool</td>
<td>✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prepare and Submit IRB Application</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Train Raters</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pilot Test (three apps)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Revise Screening Tool</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pilot Test (top 10 apps)</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Entry and Analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manuscript Preparation and Submission</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Thesis Defense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>