

UNIVERSITY OF MIAMI

DEVELOPING AND TESTING THE EFFECTS OF AN EDUCATIONAL PODCAST
TO IMPROVE CRITICAL APPRAISAL OF HEALTHCARE CLAIMS

By

Vanesa Abigail Mora Ringle

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

August, 2020

©2020
Vanessa Abigail Mora Ringle
All Rights Reserved

UNIVERSITY OF MIAMI

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

DEVELOPING AND TESTING THE EFFECTS OF AN EDUCATIONAL PODCAST
TO IMPROVE CRITICAL APPRAISAL OF HEALTHCARE CLAIMS

Vanesa Abigail Mora Ringle

Approved:

Amanda Jensen-Doss, Ph.D.
Professor of Psychology

Brian Doss, Ph.D.
Professor of Psychology

Jill Ehrenreich-May, Ph.D.
Professor of Psychology

Amy Weisman de Mamani, Ph.D.
Professor of Psychology

Nicholas Carcioppolo, Ph.D.
Associate Professor of Communications Studies

Guillermo Prado, Ph.D.
Dean of the Graduate School

MORA RINGLE, VANESA ABIGAIL

(Ph.D., Psychology)

Developing and Testing the Effects of an
Educational Podcast to Improve Critical
Appraisal of Healthcare Claims.

(August 2020)

Abstract of a dissertation proposal at the University of Miami.

Dissertation supervised by Professor Amanda Jensen-Doss.

No. of pages in text. (72)

Most research on the use of evidence-based practices (EBPs) has focused on increasing supply through providers and organizations rather than on consumer factors that increase demand. The neglect of the consumer goes against the very definition of EBP, which includes a recognition of the patient's characteristics, values, and preferences. One such factor is a parent's ability to critically appraise healthcare claims based on EBP principles. Unfortunately, very few studies have focused on how to increase this ability in parent consumers of health information. This investigation aimed to address this research gap through two studies. In Study 1 we conducted an assessment of current levels of critical appraisal in U.S. parents to determine need for a critical appraisal intervention. Parents demonstrated poor critical appraisal abilities, thus establishing a need for educational strategies targeting those abilities. In Study 2 we developed and user-tested an educational podcast to meet that need, and tested its efficacy through an online randomized controlled trial, where 250 parents were randomly assigned to listen to an experimental critical appraisal podcast (n=128) or a control podcast (n=122). The experimental podcast was called the *Parents Making Informed*

Health Choices Podcast, covered nine EBP principles, and included physical and mental healthcare claims. We found that, relative to the control podcast, listening to the *Parents Making Informed Health Choices Podcast* improved parent critical appraisal of healthcare claims and also had a positive effect on intended behaviors, EBP attitudes, and treatment preferences. We also found several correlates of critical appraisal, including behavioral intention, attitudes, and treatment preferences. Findings provide initial support for the efficacy of a brief podcast intervention to improve U.S. parent critical appraisal of healthcare claims.

TABLE OF CONTENTS

	Page
LIST OF FIGURES AND TABLES.....	iv
Chapter	
1 INTRODUCTION	1
2 STUDY 1	15
3 STUDY 2	20
4 GENERAL DISCUSSION	37
REFERENCES	50
FIGURES	60
TABLES	61

LIST OF FIGURES AND TABLES

	Page
FIGURE 1	60
TABLE 1	61
TABLE 2	63
TABLE 3	64
TABLE 4	67
TABLE 5	69
TABLE 6	71
TABLE 7	72

CHAPTER ONE: INTRODUCTION

Effective psychological practices backed by rigorous research, or evidence-based practices (EBPs), exist, yet few people receive them, or know about how research evidence can inform decisions they make about health services (Arch, Twohig, Deacon, Landy, & Bluett, 2015; Becker, 2015; Becker, Spirito, & Vanmali, 2016; Carman et al., 2010; Mora Ringle et al., 2020). Most empirical work on the adoption and implementation of EBPs has focused on therapists, while the role of the consumer has been examined less (Becker, 2015; Friedberg & Bayar, 2017; Kirk, Broman-Fulks, & Bergquist, 2016; Santucci, McHugh, & Barlow, 2012). This neglect of the consumer goes against the very definition of *evidence-based practice*, which is the combination of the best available scientific evidence, therapist expertise, and *patient characteristics, values, and preference* (APA Presidential Task Force on Evidence-Based Practice, 2006). It is also inconsistent with “patient-centered care,” which is noted as a key component of high quality care within the Institute of Medicine’s (IOM) landmark “quality chasm” report, one of the foremost frameworks for the health care system (Alegría et al., 2014; Barry & Edgman-Levitan, 2012; Harik, Hundt, Bernardy, Norman, & Hamblen, 2016; Institute of Medicine (US) Committee on Quality of Health Care in America, 2001).

In terms of specific consumer-related barriers and needs, close to 80 million American adults have persistently low health literacy, or struggle to obtain, process, and understand health information (Berkman, Sheridan, Donahue, Halpern, & Crotty, 2011; Chen, Goodson, & Acosta, 2015; Hoffmann & Del Mar, 2015; Nutbeam, 2008). Low health literacy has been linked to worse health outcomes, higher healthcare costs, and

higher hospitalization rates (Berkman et al., 2011; Vernon, Trujillo, Rosenbaum, & DeBuono, 2007). Increasing health literacy is considered a way of empowering consumers, especially when health literacy interventions target *critical health literacy*, which is defined as advanced cognitive skills, including critical analysis of health information, or *critical appraisal skills* (Barry, Sixsmith, & D'Eath, 2012; Nutbeam, 2008). Specifically, critical appraisal of health information involves using knowledge about science and EBP principles to appraise healthcare claims and distinguish reliable claims from unreliable ones; in other words, being able to distinguish EBPs from potentially ineffective treatments (Austvoll-Dahlgren et al., 2015; Nordheim, Gundersen, Espehaug, Guttersrud, & Flottorp, 2016). Thus, critical appraisal skills enable consumers to have greater control over their health decisions by facilitating making informed decisions regarding treatments. Unfortunately, little research has been conducted on strategies to increase consumer critical appraisal skills (Barry et al., 2012). As such, low critical appraisal of healthcare claims an important and addressable consumer-related factor that may increase demand for and use of EBPs.

The low critical appraisal problem is exacerbated by the fact that more than ever before, individuals have to make health services decisions based on an overabundance of both accurate and inaccurate health information readily available via mass media and the internet (Cusack, Del Mar, Chalmers, Gibson, & Hoffmann, 2018; Eysenbach, Powell, Kuss, & Sa, 2002). Additionally, studies also reveal that the public often relies on and trusts anecdotal information more than research evidence (Eysenbach & Köhler, 2002; Glenton, Paulsen, & Oxman, 2005; Schwartz et al., 2006; Sillence, Briggs, Harris, & Fishwick, 2007). This is problematic because believing unreliable claims may lead

people to pursue treatments that are ineffective and even harmful, increasing their risks and wasting valuable resources (Cusack et al., 2018).

Other consumer-related issues include their limited understanding of evidence-based health care and their attitudes toward EBPs. Extant research on consumer knowledge consists of qualitative studies that have time and again found that U.S. adults and adolescents have a limited understanding of evidence-based approaches in medicine and psychology (Becker et al., 2016; Carman et al., 2010; Mora Ringle et al., 2020; Scheyett, McCarthy, & Rausch, 2006). In all of these studies, researchers asked consumers to define what EBP means, and few participants were able to correctly define it as treatments that are effective based on research evidence. Notably, one may wonder whether simply promoting the use of the “evidence-based” or “clinically proven” terminology when referring to treatments would be sufficient; however, such phrases could easily become misused by being applied to unreliable healthcare claims (Cusack et al., 2018). Thus, this further points to a need to teach critical appraisal of healthcare claims so people can independently distinguish unreliable claims, and increase public demand for effective healthcare practices.

In addition to a lack of knowledge of evidence-based health care, adults in the U.S. also still report inconsistent values regarding the importance of research evidence when making decisions about treatment (Mora Ringle et al., 2020). In their qualitative examinations of consumer attitudes toward evidence-based medicine, Carman et al. found that while U.S. adults value research evidence when making health decisions, there are situations in which they would trust their doctor’s experience/opinion over scientific research (Carman et al., 2016; Carman et al., 2010). Additionally, studies focusing on

psychotherapy found that both treatment-seeking adults and adults from the general public favor other treatment mechanisms (e.g., therapist characteristics; therapeutic alliance) significantly more than a treatment being evidence-based, especially when the presenting problem is less severe and impairing (Carman et al., 2010; Farrell & Deacon, 2016; Swan & Heesacker, 2013; Swift & Callahan, 2010). Conversely, one recent study found that non-treatment seeking U.S. adults actually preferred treatments that have scientific support *over* other treatment variables (e.g., clinician experience, therapeutic alliance), especially with more severe and impairing psychological conditions (e.g., post-traumatic stress disorder, schizophrenia; Kirk et al., 2016). The issue of easy access to an excess of reliable and unreliable treatments claims, along with findings regarding lack of consumer knowledge of evidence-based health care, and attitudes that are not consistently favorable toward EBPs, point to the need to develop and test interventions that improve critical appraisal skills (Nordheim et al., 2016).

Existing efforts to involve consumers and improve health literacy have mostly consisted of interventions that address functional health literacy, which includes skills such as basic literacy (reading and writing) and imparting knowledge of health conditions and services, rather than teaching the ability to critically appraise healthcare claims (Nordheim et al., 2016). Available interventions include decision-aid interventions that provide information about treatment options and associated benefits and harms (Stacey et al., 2014); public deliberation, where the public is exposed to new health-related research information through written materials and interaction with experts (Carman et al., 2016); interventions to increase active patient participation in decisions about their health care (i.e., “patient activation;” Alegría et al., 2014; Deen, Lu, Rothstein, Santana, & Gold,

2011; Eliacin, Rollins, Burgess, Salyers, & Matthias, 2016; Hibbard & Mahoney, 2010; Thomas et al., 2017); programs to increase recognition of mental health disorders, helping behaviors, and reducing stigma, (such as Mental Health First Aid Training; Kitchener & Jorm, 2002; Kitchener & Jorm, 2006); and direct-to-consumer social marketing, which includes efforts to increase demands for services and products that promote public health (e.g., commercials; Becker, 2015; Friedberg & Bayar, 2017; Kirk et al., 2016; Santucci et al., 2012; Schofield, Moore, Hall, & Coles, 2016; Schofield, Weis, Ponzini, & McHugh, 2017). Although these efforts have been successful at educating consumers about specific types of treatment decisions, and even at impacting consumer attitudes and behavior, none of them address critical appraisal of healthcare claims.

There are several advantages to developing interventions that increase the public's critical appraisal abilities: 1) Critical appraisal is a cognitive ability that uses knowledge about evidence-based/science principles applicable to all general and mental health treatments claims. 2) It does not involve educating on one treatment at a time, consequently empowering consumers to confront the nearly limitless healthcare claims they are bound to encounter. Moreover, it would be impossible for consumers to memorize all of the specific treatment options that are evidence-based, and unrealistic to think they would want to do so (Jamtvedt, Klemp, Mørland, & Nylenna, 2015). 3) Critical appraisal skills will generalize to new situations and will remain relevant as the evidence-base changes or grows (Jacque, Koch-Weser, Faux, & Meiri, 2016). 4) This will likely lead to greater patient participation in shared-decision making (Austvoll-Dahlgren, Nsangi, & Semakula, 2016). 5) Previous research has shown that critical

appraisal can be taught to individuals across education levels (Nsangi et al., 2017; Semakula, Nsangi, Oxman, Oxman, et al., 2017).

Although some resources and interventions exist that teach critical appraisal, they mostly focus on teaching this skill to health professionals (Cusack et al., 2018; Horsley et al., 2011; Milne & Oliver, 1996), and often focus on specific EBP concepts/principles at a time, such as randomization, but do not cover other science/evidence-based concepts that are important to engaging in critical appraisal of healthcare claims, such as “treatments should be compared fairly” (Austvoll-Dahlgren et al., 2015). Furthermore, very few studies have examined the effects of educational interventions on critical appraisal as a *skill in decision-making* (rather than knowledge of facts about EBP). In addition, there are effective school curricula and programs that teach critical appraisal to students from elementary to high school (Jacque et al., 2016; Nordheim et al., 2016; Nsangi et al., 2017; Nsangi, Semakula, Oxman, & Sewankambo, 2015). However, school educational programs are lengthier and require use of more resources, and as such are not easily transferable to teaching critical appraisal to the adult general public, who, as previously noted, are in great need of this skill (Cusack et al., 2018).

The expansion of research on critical appraisal interventions for the public began with the development of a list of EBP principles that people must understand to engage in critical appraisal (Austvoll-Dahlgren et al., 2015). Austvoll-Dahlgren and colleagues compiled EBP principles by conducting a review of relevant literature, and consulting with experts in areas such as research methodology, health literacy, and teaching evidence-based healthcare to medical professionals. This study yielded 32 EBP principles (see Table 1), which are now known as the Informed Health Choices (IHC) Key

Concepts (hereafter referred to as “EBP principles”), and can be divided into 6 categories: 1) recognizing the need for fair comparisons of treatments, 2) judging whether a comparison of treatment is a fair comparison, 3) understanding the role of chance, 4) considering all the relevant fair comparisons, 5) understanding the results of fair comparisons of treatments, and 6) judging whether fair comparisons of treatments are relevant. A crucial point regarding these EBP principles is that they are meant to serve as a syllabus and starting off point for developing interventions and assessment tools that are tailored for the public, and ideally, delivered via mass media (Austvoll-Dahlgren et al., 2015).

In addition, a measure that assesses people’s ability to critically appraise healthcare claims by applying EBP principles has been developed and validated (Austvoll-Dahlgren, Guttersrud, Nsangi, Semakula, & Oxman, 2017; Austvoll-Dahlgren, Semakula, et al., 2017). This instrument, known as the “Claim Evaluation Tools,” includes a total of approximately 190 multiple choice items, with several items tapping into different EBP principles. The Claim Evaluations Tools can be used by researchers to develop questionnaires alongside interventions used to improve people’s ability to critically appraise healthcare claims.

To date, only one study has applied the EBP principles to a mass media intervention, using an educational audio podcast (Semakula, Nsangi, Oxman, Oxman, et al., 2017; Semakula, Nsangi, Oxman, & Sewankambo, 2015; Semakula, Nsangi, Oxman, Austvoll-Dahlgren, et al., 2017). The first step in the development of the podcast was to choose EBP principles that are relevant and comprehensible to the public through mass media. Semakula and colleagues (2015) did this by conducting workshops, which

included focus groups and surveys with health journalists from Uganda. Results from these groups indicated prioritizing nine of the 32 EBP principles about evidence-based health care in the development of a critical appraisal podcast (Semakula et al., 2015). The 9 chosen EBP principles are listed in Table 2. Notably, no study to date has focused on which EBP principles to prioritize in a U.S. critical appraisal intervention.

After determining the EBP principles to cover, Semakula and colleagues developed story lines that included examples of healthcare claims about health conditions relevant to individuals from Uganda such as malaria, diarrhea, and HIV/AIDS (see example health claims in Table 2). The IHC Podcast Program consisted of an introduction, 8 main episodes, 3 short recap episodes, and a conclusion. Each main episode lasted approximately 5 minutes, and included one or two concepts within a short storyline that was bookended with two examples of healthcare practice claims about medical conditions. Each story also included a question about the trustworthiness of the healthcare practice claim, which was resolved by applying the Key Concept pertinent to the claim. Via a randomized controlled trial (RCT) of 675 participants and with an effect size of .83, the educational podcast was found to significantly improve parents' critical appraisal abilities compared to those who only received a public service announcement (Semakula, Nsangi, Oxman, Oxman, et al., 2017). Moreover, improvements in critical appraisal were observed across varying levels of educational attainment, including approximately half of participants who had only completed elementary education. The IHC podcast is available to the public through here:

<https://soundcloud.com/informedhealthchoices/sets/informed-health-choices>.

Despite the many strengths of the IHC project's ground-breaking work, this podcast

intervention has yet to be adapted for and tested in Western nations, and importantly, has solely focused on medical healthcare claims, without any examples of mental healthcare claims. Although educating consumers is important in all areas of healthcare, it is especially relevant to mental healthcare as mental health treatments have long been plagued with unscientific or pseudoscientific claims and tolerance of such claims (Lilienfeld, 2007, 2012). In a seminal article on this subject, Lilienfeld (2012) posited that such tolerance for pseudoscientific claims in mental health may contribute to the public's lack of knowledge of effective psychological EBPs (Becker et al., 2016; Mora Ringle et al., 2020; Tanenbaum, 2008), and their underutilization (Freiheit, Vye, Swan, & Cady, 2004; Hipol & Deacon, 2013; Sayer et al., 2017).

Additionally, as previously mentioned, new research on public understanding and attitudes toward evidence-based mental health care indicated that people value scientific information when making mental health treatment decisions; however, few people (20%) knew the meaning of evidence-based mental health care (Mora Ringle et al., 2020). This was of particular concern given that 35% of the sample had previously sought mental health services. As such, it is imperative that mental healthcare claims are emphasized alongside medical healthcare claims in critical appraisal education through mass media.

Current Investigation

Ultimately, improving people's critical appraisal abilities should lead them to detect unreliable healthcare claims, helping them select effective, EBPs, thus decreasing the personal and public costs of ineffective health practices. Considering the importance and potential impact of increasing consumer critical appraisal of healthcare claims made anecdotally and through mass media, it is imperative to develop and test interventions

that teach this skill to the U.S. parents. Parents are a distinctly important consumer to target as they are responsible for healthcare decisions for both themselves and their children, and will at one point or another be confronted with healthcare claims. As such, improving a parent's ability to critically appraise healthcare claims will likely have double the benefits. Unfortunately, a critical appraisal intervention has yet to be developed for U.S. parents. Building off Semakula et al.'s pioneering research on the Uganda critical appraisal podcast, this investigation aimed to address this research gap by developing and testing the efficacy of a brief educational podcast series designed to improve U.S. parents' ability to critically appraise medical and mental healthcare claims. We examined podcast effects on critical appraisal ability, intended behavior, attitudes, and preference for EBPs.

Innovations of the Current Investigation

This is the first attempt to teach critical appraisal to parents through a mass media format in the U.S. This study builds off Semakula et al.'s Uganda podcast by modifying it to include mental health claims and to be relevant to American parents. Like the Uganda podcast, the content is brief and accessible to people from various backgrounds and educational levels. We ensured the podcast's accessibility by applying the same user-testing methodology used in the development of the Uganda podcast. Additionally, we utilized communications theory, including frameworks such as the extended elaboration likelihood model, as a general guide to create an educational podcast that included entertainment elements (e.g., relevant storylines, relatable characters) that could potentially increase its influence on attitudes and intended behavior and reduce resistance to the podcast's health message (Moyer-Gusé, 2008).

This is the first study to integrate examples of mental health claims when teaching critical appraisal. Since this approach to increasing health literacy targets a cognitive ability, it is generalizable to appraising the reliability and trustworthiness of claims made about both medical and mental health practices. However, given the preponderance of pseudoscientific claims related to mental health practices, the generalizability of the skills will likely be enhanced by also including examples of mental health claims. This is a unique expansion of the original podcast which only included physical health claims relevant to Uganda (e.g., Quinine can cure malaria).

Finally, the audio podcast format may be a uniquely powerful medium for communicating health information in the U.S., since it is estimated that approximately 112 million Americans over the age of 12 have listened to at least one podcast (Edison Research, 2017). Moreover, the audio format facilitates exposure to the content in that it can be listened to anytime, anywhere, and as many times as the listener prefers. In addition, this study may increase ecological validity by providing the podcast to participants via the internet which is a common way for people to seek health information (Eysenbach & Köhler, 2002; Eysenbach et al., 2002; Seçkin, Yeatts, Hughes, Hudson, & Bell, 2016). The internet also provides an efficient means for conducting research projects and collecting high-quality data that is representative of the population (Berinsky, Huber, & Lenz, 2012; Buhrmester, Kwang, & Gosling, 2011; Levay, Freese, & Druckman, 2016). If found to be efficacious, the podcast series could be made readily available on consumer-focused websites (e.g., the National Alliance of Mental Illness, <https://www.nami.org/>).

Aims

Aim 1: Assess current critical appraisal abilities in the U.S. parents to determine

critical appraisal needs and podcast intervention content. To identify adaptations and EBP principles to prioritize in the U.S., we surveyed parents on Amazon's Mechanical Turk (MTurk) using items from the Claim Evaluation Tools that represent 28 (of an ever-expanding list of 32) EBP principles (Austvoll-Dahlgren, Guttersrud, et al., 2017; Austvoll-Dahlgren, Semakula, et al., 2017). Results from this preliminary survey were used to identify U.S. parents' current critical appraisal abilities, and informed which EBP principles to prioritize in the U.S. podcast. This aim is referred to as Study 1 below. Subsequently, the podcast was user-tested, followed by an online RCT.

Aim 2: Develop a brief podcast to improve U.S. parent critical appraisal. In

collaboration and consultation with experts on evidence-based health care and health literacy, we developed the U.S. parent critical appraisal podcast by identifying mental health additions and other adaptations to make to the existing podcast that was implemented and tested with parents in Uganda. This included writing new scripts with new storylines that included characters and healthcare claims relevant to U.S. parents.

Aim 3: Test the effects of critical appraisal podcast on critical appraisal skills.

Considering that parents in the U.S. often encounter health and treatment information via the internet, we tested the efficacy of the podcast program by recruiting parent participants online via Amazon's Mechanical Turk (MTurk), an online crowdsourcing platform widely utilized for research participant recruitment and data collection by scientists across many fields of study (Buhrmester, Kwang, & Gosling, 2016; Litman, Robinson, & Abberbock, 2017). Parents were randomized to the critical appraisal podcast

condition or a control podcast of similar length.

Aim 3 Hypothesis: Parents who listen to the critical appraisal podcast will exhibit better critical appraisal skills than those who do not.

Aim 4: Test the effects of critical appraisal podcast on intended behavior. As was done in the Uganda podcast RCT, we also planned to examine the same intended behaviors they assessed after parents listened to the podcast in Uganda.

Aim 4 Hypothesis : Compared to parents who do not listen to the critical appraisal podcast, parents who listen to the critical appraisal podcast will report increased rates of a) intending to find out what a treatment claim is based on, b) intending to find out if a claim is based on a research study comparing the treatment to no treatment (a fair comparison), c) intending to participate in a research study comparing two treatments (a fair comparison) for a hypothetical illness they have.

Aim 5: Test the effects of critical appraisal podcast on attitudes toward EBPs, empiricism in mental health, and vaccines. We aimed to measure various attitudes potentially related to critical appraisal of health claims based on EBP principles. Notably, given growing concerns regarding outbreaks of diseases such as measles linked to antivaccination beliefs/attitudes, we aimed to capture vaccine information in the critical appraisal podcast, and consequently examine vaccine attitudes.

Aim 5 Hypothesis: Parents who listen to the critical appraisal podcast will have more positive attitudes toward EBPs, empiricism in mental health, and vaccines than those who do not listen to it.

Aim 6: Test the effects of critical appraisal podcast on treatment preferences.

Aim 6 Hypothesis: Parents who listen to the critical appraisal podcast will have a

higher preference for an effective treatment being evidence-based over other treatment components such as therapist experience.

Aim 7: Examine podcast satisfaction. Even after conducting user-testing, we anticipated the importance of measuring satisfaction with the critical appraisal podcast, and examining whether overall satisfaction levels differed between the two study conditions. However, in view of our user-testing efforts, we did not expect differences in satisfaction between the conditions.

Aim 8: Explore correlates and moderators of U.S. parent critical appraisal. Examine whether various demographic and other factors are associated with critical appraisal abilities, and/or moderate the effects of the podcast on critical appraisal.

CHAPTER TWO: STUDY 1

Method

Participants

Approval was obtained from the University of Miami's Institutional Review Board prior to beginning Study 1. We recruited participants on Amazon's MTurk platform, an online labor market where anonymous "workers" receive monetary payment for completing various research tasks. Mturk has been operating for 15 years and researchers from many fields use the platform to gather high quality survey, experimental, and intervention data via the internet (Bohannon, 2016; Mason & Suri, 2012). Research on MTurk samples has found them to be fairly representative of the U.S., internet-using public, and more diverse than other convenience samples such as undergraduate students and other online samples (Berinsky et al., 2012; Buhrmester et al., 2011; Chandler & Shapiro, 2016). However, it should be noted that MTurk samples tend to be younger, politically more liberal, have higher educational attainment, and include more White and Asian individuals (Chandler & Shapiro, 2016). Nevertheless, these demographic characteristics may vary based on specific MTurk populations. For example, it has been found that surveying parents on MTurk is a fast, cost-effective, and reliable way of recruiting parents from diverse backgrounds via the internet, especially when compared to other online recruitment methods such as Listservs and Facebook (Dworkin, Hessel, Gliske, & Rudi, 2016; Schleider & Weisz, 2015). Additionally, Mturk samples have been found to have elevated psychopathology, making them a particularly appropriate sample for studies about mental health services, as they are likely to seek

services (Arditte, Çek, Shaw, & Timpano, 2016; Jensen-Doss, Patel, Casline, Mora Ringle, & Timpano, 2019; Walters, Christakis, & Wright, 2018).

Notably, in light of past MTurk data integrity issues, MTurk and researchers are now increasingly employing various measures to protect data quality (see Kennedy et al., 2020). Indeed, recently (March, 2020), MTurk outlined an “abuse prevention program” consisting of technologies that detect and monitor fraudulent activity and swiftly take action in situations of fraudulence (Amazon Mechanical Turk, 2020). Accordingly, in both studies, we used recommended data quality measures such as checking study completion length, including attention check questions, and questions assessing consistency in demographic information reported (see Kennedy et al., 2020).

We recruited 179 internet-using parents through MTurk to complete a critical appraisal test. In order to participate, parents had to be U.S. residents, had to have an MTurk approval rating of 98% or higher, and had to have at least one child below age 18 (as determined by self-report). Participants received a \$2.00 compensation through MTurk for completing the critical appraisal test (see MTurk best practice recommendations; Chandler & Shapiro, 2016). Out of 179 recruited individuals, data was used for 142 parents. The data provided by 37 individuals was not utilized in analyses due to its poor quality, including completing the survey under 10 minutes (projected completion time was at least 20 minutes; $n = 23$) and failing attention check items ($n = 14$). The critical appraisal needs assessment parent sample had a mean age of 36.9 years, was predominantly Caucasian (83%), 45% female, and 55% had a Bachelor’s or higher degree. See Table 4 for more detailed demographic information.

Measures

Demographics and other characteristics. Participant's age, gender, ethnicity, educational attainment (high school, some college, bachelor's degree, etc.), and employment status (e.g., unemployed, employed part-time, etc.) among other sociodemographic variables were collected through self-report prior to completion of the critical appraisal test. See Appendix A for Study 1 questions.

Critical appraisal. We assessed critical appraisal skills using 52 multiple choice questions from the Claim Evaluation Tools (Austvoll-Dahlgren, Guttersrud, et al., 2017; Austvoll-Dahlgren, Semakula, et al., 2017). This measure was developed by the IHC team to include multiple items for each of the EBP principles that people need to understand to critically appraise healthcare claims. In total, the Claim Evaluations Tools includes an ever-growing bank of over 130 multiple choice questions that can be used with individuals age 10 and older. This measure was psychometrically validated using Rasch analysis (Austvoll-Dahlgren, Guttersrud, et al., 2017). Study 1 items were selected based on their cultural relevance to a U.S. audience, and on how well they performed in the previous validation study. Additionally, items with high performance rates in Uganda were not selected as they were considered to potentially be too easy for a U.S. audience. There were three study-created questions that were worded exactly like Claim Evaluation Tools questions but that replaced physical health conditions with behavioral health conditions. The internal consistency of this 55-item critical appraisal scale was excellent ($\alpha = .93$).

Analytic Plan

We used SPSS Version 25 to run descriptive statistics and frequencies of demographic variables and proportions of correct responses on the critical appraisal test.

Results

We found that, despite being a relatively educated sample (55% had a bachelor's degree or higher), parents struggled to engage in critical appraisal with an average proportion of correct responses of 64% (Mean = 35.15 out of 55 questions; SD = 11.8; Range = 11-53). Specifically, parents struggled to critically appraise claims based on the following EBP principles (%'s indicate proportion of correct responses): Comparisons are needed to identify effects (55%); Advantages should outweigh disadvantages (62%); Comparison groups need to be similar (63.2%); Common practice is not always evidence-based (63.9%); Single comparisons can be misleading (63.9%); Association is not causation (67%); Anecdotes are unreliable evidence (68%). Table 3 lists all findings.

Of the nine EBP principles covered in the Uganda podcast, U.S. parents had the most trouble with the following (%'s indicate proportion of correct responses): Treatments should be compared fairly (55.0%), Decisions about treatments should not be based on considering only their benefits (62.0%); Comparison groups need to be similar (63.2%), Common practice is not always evidence-based (63.9%), and Single comparisons can be misleading (63.9%).

Discussion

Study 1 results revealed that, although over half of the parents reported having a bachelor's or more, their average proportion of correct responses was only 64%

(answered 35 out of 55 questions correctly). Some of the specific EBP principles parents struggled with included: “Treatments should be compared fairly,” and “Common practice doesn’t mean it’s beneficial/safe.”

In view of the present findings, we decided that the U.S. podcast in Study 2 should include all nine EBP principles covered in the Uganda podcast. There were a several reasons for this decision. First, the nine EBP principles covered in the Uganda podcast are some of the most practical and common EBP principles that parents will encounter. Second, these EBP principles were carefully selected by the Uganda podcast development team after interviewing journalists and using a modified Delphi technique to determine EBP principles to prioritize in critical appraisal resources created for the lay public based on their understandability, accessibility, and relevance (Semakula et al., 2015). Third, when we examined participant performance on the critical appraisal questions that cover these nine principles, we found that participants earned a score of 65%. If they had performed better than this, we may have considered omitting some EBP principles that were easier to create a shorter podcast. Even though present critical appraisal abilities around EBP principles “Anecdotes are unreliable evidence” and “Association is not the same as causation,” were slightly above the average of 64%, we still planned to cover these EBP principles in the U.S. podcast in Study 2 given their real-world importance, and research that has found that consumers highly regard anecdotal information, sometimes even over research evidence.

CHAPTER THREE: STUDY 2

Method

Study Design and Procedures

Study 2 consisted of three phases that occurred chronologically: 1) podcast development/adaptation, 2) podcast user-testing, and 3) online RCT. The online RCT phase compared two conditions: a) the podcast intervention and b) an audio podcast control condition (see Online RCT section below). All Study 2 procedures were approved by the University of Miami's Institutional Review Board prior to commencing.

Podcast Development/Adaptation. After completing Study 1, we reviewed the scripts and storyboards of the Uganda podcast, and determined needed adaptations (e.g., changing treatment claim examples about malaria). The PI and 3 undergraduate research assistants independently listened to the existing podcast while reading the script, and made notes about needed adaptations. They then compared notes, discussed as a group and with PI advisor (Dr. Jensen-Doss), and came to a consensus about the needed alterations. The PI and undergraduate research assistants then scanned U.S. mass media websites such as Health News Review and Behind the Headlines for examples of behavioral healthcare claims that are popular and relevant to U.S. parents. Once we accumulated sufficient examples of healthcare claims relevant to U.S. parents, the PI wrote new scripts that were reviewed by Dr. Jensen-Doss and revised through an iterative process. Once the U.S. podcast script was finalized, the podcast was produced by the *Orange Umbrella*, a production company housed within the University of Miami's School of Communication. Student voice actors performed podcast characters. The final

product consisted of an introduction, 7 main episodes, and a conclusion, and was formatted to be easily played on a phone or computer (see “Podcast Intervention and Control Condition” section below for more information on podcast content). After the Orange Umbrella finalized a prototype of the podcast that included all episodes and music, parents provided feedback on it through user-testing.

User-testing. In line with best practices, we carried out user-testing where parents listened to the podcast prototype. We recruited parent end-users through the National Alliance on Mental Illness’ (NAMI) Miami Dade County chapter, advertisements in public community spaces (community clinic), and Facebook posts. Community recruitment allowed us access to the perspective of various consumers of mental health information, including individuals with mental health conditions, families and friends of individuals with these conditions, and parents/individuals from the general public seeking mental health information. In user-testing, developers obtain feedback from users during and/or shortly after they interact with a product. User-testing enables designers to know how their product might be received by the intended user, and can be used to get feedback early in the design process, which can be incorporated into the final product (Rosenbaum, 2010). The user-testing process lasted approximately 60-90 minutes, including 32 minutes of estimated individual podcast listening, a 20-30-minute semi-structured interview (See Appendix A), and a 5-minute online satisfaction survey. Participants were compensated \$15.00 through a giftcard.

The PI then compiled parent feedback, and worked with the Orange Umbrella production company to apply suggested changes. After changes were applied and reviewed by the PI and Dr. Jensen-Doss, the critical appraisal podcast was determined to

be in its final form, and was ready to be tested via an online RCT. The U.S. podcast, named *Parents Making Informed Health Choices Podcast*, is described further in the “Podcast Intervention and Control Condition” section below.

Online RCT. The study was advertised to MTurk workers in the U.S. with MTurk approval ratings of 98% or higher who did not participate in Study 1 because they would have had prior exposure to the critical appraisal questions. We restricted their ability to view this online RCT as a study opportunity by using functions built into TurkPrime, a sourcing platform for research studies recruiting on MTurk (Litman, Robinson, & Abberbock, 2017). Eligible participants were age 18 or older, had at least one child below age 18, and were fluent in English. After consent, participants completed baseline measures. Parent demographic and other characteristics data were the only data collected at baseline. All other outcome measures (e.g., critical appraisal, EBP attitudes) were completed after random assignment and listening to the podcasts. Exposing participants to questionnaires prior to listening to the podcasts would have affected podcast listening, and it was expected that randomization would ensure that participants were equivalent at baseline (see Randomization Success section in Results section below). Moreover, this was consistent with methods used in the original trial of the IHC podcast (Semakula et al., 2017).

Following baseline measures, participants were randomly assigned on a 1:1 ratio to either the critical appraisal podcast or a control podcast, and received instructions for listening to the podcast. Participants in both conditions could not move on from the podcast listening page until 32 minutes (length of podcasts) had lapsed and they also could not fast forward. After they listened to the entire podcast, participants completed

post-intervention measures. Participants received a fair wage payment of 7.25/hour (current federal hourly minimum wage) for their participation, which is in compliance with best practices guidelines for MTurk (Buhrmester, Talairar, & Gosling, 2018). Of note, online RCT data was collected in two waves in the span of seven days: the first wave aimed to recruit 20 participants, while the second wave aimed to recruit 230 participants. The purpose of this was to ensure that all planned baseline and post-intervention procedures ran smoothly with a smaller sample, so as to reduce the number of unanticipated problems when recruiting the 230 parent sample. Importantly, all procedures ran smoothly with the first wave, therefore we did not change anything with the second wave of participants.

Podcast Intervention and Control Condition

The U.S. critical appraisal podcast (*Parents Making Informed Health Choices Podcast*) consisted of nine, 2-4-minute-long episodes, including an introduction episode and a conclusion/recap episode. The entire podcast was 32 minutes long. The seven main episodes covered nine different EBP principles using medical and/or behavioral conditions (e.g., attention deficit hyperactivity disorder). Following storylines and messages developed for the Uganda IHC podcast, the U.S. podcast included two main characters who engaged in back-and-forth conversations while explaining and applying EBP principles to different medical and behavioral conditions. The *Parents Making Informed Health Choices Podcast* can be found here: <https://soundcloud.com/user-542959241-406231118/sets/parents-making-informed-health-choices>. Table 2 lists EBP principles covered and provides examples of healthcare claims.

The control podcast served as an inert condition that was similar in length to the

critical appraisal podcast, but included no content related to critical appraisal. This control condition was chosen because the only major podcast component that needed to be controlled for was the length of time spent listening or paying attention to content presented audibly. The chosen control podcast was a free, publicly available 33-minute-long mindfulness-sitting meditation podcast and can be found here:

<https://soundcloud.com/mindfulness-dublin/sitting-meditation-with-helen-byrne>.

User-Testing Participants

After completing a critical appraisal podcast, we recruited five parents from the South Florida community. Parents were recruited through the National Alliance on Mental Illness' (NAMI) Miami Dade County chapter, community mental health clinics, and community organizations by advertising on their social media pages, and by posting flyers in the community clinics. Parents were 18 years or older, fluent in English, and had children below age 18. User-testing parents were all female, 60% were White, and on average were 36 years old.

Online RCT Participants

Similar to Study 1, we recruited online RCT participants through MTurk, however, this study was only available to parents (of children 18 and younger) who had not participated in Study 1, given that they would have had prior exposure to the critical appraisal questions. (We restricted viewing this online RCT as a study opportunity by using functions built into TurkPrime, a sourcing platform for research studies recruiting on MTurk). We recruited a total of 250 online RCT participants. First, we recruited 20 participants, followed by 230 participants over the span of seven days. Similar to Study 1, participants in online RCT were 18 or older, were U.S. residents, had MTurk approval

ratings of 98% or higher, and had at least one child below age 18. Participant flow is presented in the consort diagram in Figure 1.

On average, the 250 parents that participated in the online RCT were 35 years old, and 49% were female. Parents were predominantly White (75%); 17% were African American, 5% were Asian American, and 14% were Hispanic/Latino. Fifty-five percent of parents had a Bachelor's degree or higher. Table 4 provides additional information about demographic characteristics.

Measures

Demographics and other characteristics. Participant's age, gender, ethnicity, educational attainment (high school, some college, bachelor's degree, etc.), employment status (e.g., unemployed, employed part-time, etc.), podcast listening habits/consumption (listening time in minutes), and other characteristics data (e.g., treatment seeking history) were collected at the end of the user-testing phase, and at baseline during the online RCT. See Appendix B for all Study 2 questionnaires.

Critical appraisal. We assessed the primary outcome variable using 18 items from the Claim Evaluation Tools (Austvoll-Dahlgren, Guttarsrud, et al., 2017; Austvoll-Dahlgren, Semakula, et al., 2017), the same measure used in Study 1. In the "Online RCT" we included 18 items that tapped into the 9 EBP principles covered in the critical appraisal podcast. Items were selected based on their cultural relevance to a U.S. audience, and on how well they performed in the previous validation study. Additionally, items with high performance rates in Uganda were not selected as they were considered to potentially be too easy for a U.S. audience. There were 3 study-created questions that were worded exactly like Claim Evaluation Tools questions but replaced physical health

conditions with behavioral health conditions. The internal consistency of the final 21-item critical appraisal scale was good ($\alpha = .86$).

Intended behavior. Study 2 participants also completed the intended behaviors measure used by Semakula et al (2017). The measure consists of three items asking about the likelihood that someone will 1) find out what a treatment claim is based on, 2) find out if a claim is based on a research study comparing the treatment to no treatment (a fair comparison), and 3) saying “yes” if asked to participate in research study comparing two treatments for an illness they have. Response options were on a 4-point Likert scale from very unlikely to very likely, and also included an “I don’t know” option.

Attitudes toward evidence-based practices. We used the Consumer Attitudes Towards Evidence Based Services Scale (CAEBS) to assess participant attitudes (Teh, Hayashi, Latner, & Mueller, 2016). This measure consists of 29 items which, according to results from an exploratory factor analysis conducted by the developers, load onto four factors: Factor 1: Beliefs Regarding Therapists’ Practices; Factor 2: Attitudes about Mental Health Policy; Factor 3: Negative Personal-Level Attitudes toward EBPs; and Factor 4: Negative Societal-Level Attitudes towards EBPs. Items are rated on a 5-point likert scale from strongly disagree to strongly agree. Internal consistencies in the current sample were as following: Factor 1: $\alpha = .77$; Factor 2: $\alpha = .59$; Factor 3: $\alpha = .88$; Factor 4: $\alpha = .73$. To increase reliability, we removed item 13 from the Factor 2 scale (new $\alpha = .73$), and item 24 from the Factor 4 scale (new $\alpha = .80$)

Attitudes regarding empiricism in mental health treatment. We also used EBP attitude items created by Kirk et al (2016) that assessed “Attitudes regarding empiricism in mental health treatment” on a 5-point likert scale, with higher scores indicating more

agreement. The internal consistency of this 5-item attitudes scale in this sample was adequate ($\alpha = .77$).

Attitudes toward vaccines. We examined vaccine attitudes (namely, vaccine safety concerns) through the following 3 items 1) vaccines are unsafe, 2) vaccines have long-term negative side effects, and 3) If I had another infant today, I do not want him/her to get all the recommended vaccinations, which have been used in previous studies (Gust et al., 2004; Moran, Frank, Chatterjee, Murphy, & Baezconde-Garbanati, 2016). Participants rated these on a 5-point likert scale with higher overall scores indicating more negative vaccine attitudes. The internal consistency of this 3-item vaccine attitudes scale in the current sample was $\alpha = .93$.

Treatment Preferences. We also assessed preference for various treatment components, including: 1) treatment being evidence-based, 2) therapeutic alliance, 3) therapist experience, 4) empathic qualities of therapist, and 5) client speaking for majority of sessions (Kirk et al., 2016). As executed by Kirk et al (2016), participants could allocate 99 points across the different treatment components, and were instructed that more points signify higher preference for that treatment component.

Podcast Listening Fidelity. We assessed whether participants actually listened to the podcast by asking questions about the podcast content that anyone who listened to the entire podcast should be able to answer (i.e., basic details about content). This measure included 7 questions specific to study condition. Questions 1-6 were true or false, and question 7 required an open-ended response to a question about content at the end of the podcast. Responses to question 7 were assigned 0 – 2 points by the PI, thus, participants could earn up to 8 points total for paying attention and remembering basic podcast

content.

Podcast Satisfaction. Participants in both conditions also rated their satisfaction with the podcast through 5 items that asked about overall satisfaction, likelihood of recommending to others, and relevance.

Analytic Plan

All statistical analyses were run using SPSS Version 25. Descriptive statistics were used to examine participant demographics, listening fidelity, and podcast satisfaction. We also examined the distribution of study variables, including analyzing for outliers, homoscedasticity, and kurtosis. We also conducted preliminary analyses to examine whether there were differences on demographic variables between the two study conditions.

We examined the hypotheses for Aims 3 – 7 using multiple linear regression for continuous variables and binomial logistic regression for categorical variables.

For exploratory Aim 8, we examined correlates and moderators of critical appraisal abilities, using t-tests, multiple linear regression analyses, and by adding interaction terms to multiple regressions analyses. Cohen's *d* values of .20, .50, and .80 indicated small, medium, and large effects; R^2 values of .01, .09, and .25 indicated small, medium, and large effects (Cohen, 1988).

Results

User-Testing

Out of the five user-testing parents who listened to the critical appraisal podcast prototype, four reported overall being either “very satisfied” or “satisfied” with the podcast prototype ($n = 4$), and one parent reported being “unsure”. From the semi-

structured interviews, feedback from parents was mostly positive; they noted that the critical appraisal podcast was relatable, interesting, and helpful. Their recommendations for changes were that length should be shortened (prototype length was 39 minutes), conclusion episode should be shorter/include less recapping details, and should “be more to the point”. Additionally, parents noted that while the stories and the information discussed by the fictional characters were realistic, the back-and-forth dialogue between the characters sometimes sounded contrived. After obtaining user-testing feedback, the critical appraisal podcast was edited to incorporate feedback from end-users. We made the following changes: 1) Shortened the length by editing the script; 2) Edited character dialogue as much as possible, where appropriate, to reduce “contrivedness.”

Unfortunately, rerecording of the voices and readings was not possible, therefore, the production company made audio edits throughout to cut words or character expression/laughter that sounded contrived. 3) Rewrote conclusion to be shorter and include straightforward takeaways rather than go through the content of each episode in detail.

Online RCT - Descriptive Statistics

Preliminary analyses indicated that none of the assumptions of normality were violated except for moderately negatively skewed distributions (close to 1) for the main outcome variable, critical appraisal (skewness statistic = $-.64$), and for Factor 1 of the CAEB (Beliefs Regarding Therapists' Practices skewness statistic = $-.77$). Thus, we performed log and square root transformations of these variables which successfully reduced the skewness statistic to be closer to zero (log transformed critical appraisal skewness statistic = $.42$; square root transformed CAEB Factor 1 skewness statistic = -

.08). Table 5 includes descriptive statistics for study measures, including means and standard deviations for this sample.

As previously explained on pages 14 and 15, we utilized multiple validity measures, including length of study participation (we anticipated that it would take at least 60 minutes to complete), attention check questions, and two questions measuring consistency in report of demographic information (age and youngest child age) strategically placed at the beginning and at the end of the study. Participants who were inconsistent in their report of their age and their youngest child's age by more than a year were considered to potentially provide poor quality data. Out of 250 recruited individuals, six participants completed the study under 42 minutes (which is 70% of the projected completion time of 60 minutes); 18 participants failed attention check/listening fidelity questions (i.e., obtained a score of 4 out of 8 on the listening fidelity measure); and 25 participants provided inconsistent demographic data (e.g., providing a birth year that does not match age reported earlier in the study by more than a year). However, consistent with an intent to treat approach, because these participants were randomized to conditions and completed all study procedures, they were included in all study analyses.

Listening Fidelity. Parents in both conditions could obtain 8 possible points on the tests of attention to the respective podcasts. Parents in the critical appraisal podcast condition obtained an average score of 6.73 on the listening fidelity test (SD = 1.68, range = 1-8); parents in the control podcast condition earned an average score of 6.76 (SD = .84, range = 1-8).

Randomization Success

There were no significant differences between the critical appraisal podcast

condition and control conditions at pre-intervention with respect to demographic variables of interest, including gender ($\chi^2[1] = .679, p = .413$), education level ($\chi^2[1] = .88, p = .349$) or annual household income ($t[248] = .720, p = .472$). However, the control group was significantly older by 2 years ($t[248] = 2.43, p = .016$); therefore we controlled for age in all analyses. We also did not find significant between group differences in podcast listening habits ($t[248] = -.744, p = .457$), and podcast listening fidelity scores ($t[248] = .155, p = .877$). However, given some variability and to reduce the possibility that significant relationships are due to latent third variables, we also included education and listening fidelity as control variables in outcome analyses.

Aim 3: Critical Appraisal Outcomes

As hypothesized, parents who listened to the critical appraisal podcast performed significantly better on the critical appraisal measure than those who listened to the control podcast ($B = .450; p < .0001$). The set of predictors including control variables (age, educational attainment, and listening fidelity) and intervention condition explained 35% of the variance in critical appraisal, with podcast condition alone accounting for 7% of the variance (ΔR^2 for intervention condition = .07; a small effect). In terms of “high vs. low score” categories, a natural split occurred in the critical appraisal outcome such that 53.2% participants earned a score of 16 or lower, and 46.8% of participants earned scores of 17 or higher. Thus, we dichotomized the critical appraisal variable based on this natural split to examine high vs. low score differences between the podcast conditions. When examined in this way, results confirmed continuous outcome findings in that participants in the critical appraisal podcast condition were more than four times more likely to obtain scores in the “high score” category than those in the control podcast

condition ($OR = 4.50, p < .001$). Table 6 presents all results from linear regression analyses.

Aim 4: Intended Behavior Outcomes

Compared to parents who listened to the control podcast, parents in the critical appraisal podcast condition reported a significantly higher likelihood of “*finding out if a claim was based on a fair comparison study*” ($B = .252; p = <.01; \Delta R^2$ for intervention condition = .04). There were no significant differences between the conditions regarding the likelihood of engaging in the other two intended behaviors assessed: “*finding out what a claim was based on,*” and “*participating in fair comparison study*” (see Table 6)

Aim 5: Attitude Outcomes

Between group differences in EBP attitudes were examined across the four attitudes factors of the CAEBS (see Table 6). We found that parents in the critical appraisal condition had lower scores on the Negative Personal-Level Attitudes towards EBPs scale (more positive EBP attitudes) ($B = -1.96; p < .05; \Delta R^2$ for intervention condition = .015). There were no significant differences between podcast groups in the three other CAEBS EBP attitudes scales (Beliefs Regarding Therapists’ Practices, Attitudes about Mental Health Policy, Negative Societal-Level Attitudes towards EBPs), as well as on attitudes regarding empiricism in mental health treatment, and attitudes toward vaccines.

Aim 6: Treatment Preferences Outcomes

In terms of participant preferences for various treatment components, parents who listened to the critical appraisal podcast reported significantly higher preference for receiving effective therapies backed by scientific studies compared to parents who

listened to the control podcast ($B = 4.89$; $p < .05$; ΔR^2 for intervention condition = .02). Parents did not differ significantly in other treatment preferences assessed, including, therapeutic alliance, therapist experience, empathic qualities of therapist, and client speaking for the majority of therapy sessions (see Table 6).

Aim 7: Podcast Satisfaction

The majority of parents reported being very satisfied (40%) or satisfied (47%) with the critical appraisal podcast. Four percent of parents reported being unsure about their level of satisfaction, 8% reported being unsatisfied, and 2% reported being very unsatisfied with the critical appraisal podcast. There were no between group differences in podcast satisfaction. See Tables 5 and 6 for all podcast satisfaction data.

Aim 8: Correlates and Moderators of Critical Appraisal (Exploratory Analyses)

Parent Demographics. None of the examined parent demographic characteristics (age, income, educational attainment, and receiving medication for a psychological condition, and psychotherapy) were related to critical appraisal.

Intended Behavior. When controlling for podcast listening, one of the 3 intended behaviors assessed was associated with parent critical appraisal abilities. Specifically, higher intention to “*find out if a claim was based on a fair comparison study*” was significantly associated with higher critical appraisal scores ($B = .27$; $p < .01$; ΔR^2 for intended behavior = .04).

EBP Attitudes. Despite the lack of significant EBP attitude differences between the podcast groups, we found that, when controlling for condition, all EBP attitudes scales were significantly related to critical appraisal. Namely, higher scores on the Beliefs Regarding Therapists’ Practices scale (more positive EBP attitudes) ($B = .344$; $p < .001$;

ΔR^2 for Beliefs Regarding Therapists' Practices scale = .32) were significantly related to better critical appraisal abilities; while higher scores on the Attitudes about Mental Health Policy scale ($B = -.04, p < .05, \Delta R^2$ for Attitudes about Mental Health Policy scale = .02), and Negative Personal-Level Attitudes towards EBPs scale (more negative EBP attitudes) were associated with lower critical appraisal abilities ($B = -.042, p < .001, \Delta R^2$ for Negative Personal-Level Attitudes towards EBPs scale = .15). Interestingly, higher scores on the Negative Societal-Level Attitudes towards EBPs scale (more negative EBP attitudes) were associated with better critical appraisal ($B = .033, p < .05, \Delta R^2$ for Negative Societal-Level Attitudes towards EBPs scale = .02).

Attitudes Regarding Empiricism in Mental Health Treatment. When controlling for condition, more positive attitudes regarding empiricism in mental health treatment were associated with better critical appraisal ($B = .46, p < .001, \Delta R^2 = .11$).

Vaccine Attitudes. Higher vaccine attitudes scores (i.e., less favorable attitudes toward vaccines or more vaccine hesitance) were significantly associated with lower critical appraisal skills when controlling for study condition ($B = -.082, p < .001, \Delta R^2 = .12$).

Treatment Preferences. When controlling for condition, we found that greater preference for receiving effective therapies backed by scientific studies ($B = .014, p < .001, \Delta R^2 = .09$), less preference for having a therapist with many years of professional experience ($B = -.022, p < .001, \Delta R^2 = .08$), and less preference for a therapist allowing them to speak the majority of the therapy session ($B = -.02, p < .001, \Delta R^2 = .06$) were significantly related to better critical appraisal skills.

Moderators. A final aim of this study was to explore whether parent

sociodemographic characteristics moderated the critical appraisal podcast effect on parent critical appraisal. None of the moderators examined approached significance, including parent age, educational attainment, income, and receiving medical or psychological treatment for a psychological problem. See Table 7 for moderation analyses results.

Post Hoc Analyses Without Intent to Treat Sample

We ran all study analyses excluding the 49 participants who failed data validity measures and, in terms of critical appraisal outcomes, found that the magnitude of the intervention effect increased from 7% (a small effect) to 11% (a medium effect) when we did not include these parents ($B = .525$; $p < .0001$). In regards to other outcomes, when we removed the intent to treat sample, we no longer found a significant difference between podcast groups in the Negative Personal-Level Attitudes towards EBPs scale, but did find that parents who listened to the critical appraisal podcast reported significantly more positive attitudes regarding empiricism in mental health treatment ($B = .210$; $p < .05$; $\Delta R^2 = .03$).

In terms of correlates, all findings remained the same with the exception of one parent characteristic—taking medication for a psychological condition, which became significantly related to higher critical appraisal scores ($M = 16.83$, $SD = 4.15$) compared to parents who reported never taking medication for a psychological condition ($M = 15.25$, $SD = 4.65$); ($t[199] = 2.14$, $p < .05$, $d = .42$). Additionally, upon removing parents with questionable validity data, the CAEBS Attitudes about Mental Health Policy scale was no longer significantly related to lower critical appraisal.

Discussion

In Study 2, we developed and tested the efficacy of a story-based, educational

podcast—*The Parents Making Informed Health Choices Podcast*—an intervention designed to teach the application of EBP principles through stories, and increase U.S. parent critical appraisal of healthcare claims. As part of an online randomized controlled trial (RCT), we examined the effects of the podcast intervention on the following outcomes: critical appraisal, intended behaviors, EBP attitudes, attitudes toward empiricism in mental health treatment, vaccine attitudes, and treatment preferences. We found that listening to the *Parents Making Informed Health Choices Podcast* improved parent critical appraisal of healthcare claims and had an effect on intended behaviors, attitudes, and treatment preferences. To our knowledge, this is the first online RCT of a mass media critical appraisal intervention for U.S. parents.

CHAPTER FOUR: GENERAL DISCUSSION

With the aim of empowering the public to pursue effective healthcare practices, researchers in Uganda and Norway (the Informed Health Choices project; www.informedhealthchoices.org) are conducting ground-breaking work to improve consumer critical appraisal of healthcare claims through awareness of EBP principles (Semakula et al., 2017). Unfortunately, no such work to date has focused on U.S. parents, who, as gatekeepers of healthcare services for themselves and their families, are doubly responsible for making healthcare choices based on an overabundance of unreliable healthcare claims. Additionally, despite a longstanding history of acceptance of pseudoscience in mental healthcare, no study or intervention to date has prioritized the inclusion of critical appraisal of *mental* healthcare claims. Thus, this investigation focused on bringing consumer-focused, critical appraisal research to the U.S. through two distinct but connected studies focused on parents.

Study 1

The purpose of Study 1 was twofold: first, we wanted to characterize current U.S. parent critical appraisal levels, and secondly, we aimed to use this information to determine podcast intervention content. To the best of our knowledge, this was the first time that parent critical appraisal of healthcare claims has ever been evaluated in the U.S. at all, and precisely through the Claim Evaluation Tools, a reliable and valid battery of multiple-choice questions designed to assess a person's (ages 10 and older) ability to critically appraise medical/physical healthcare claims (Austvoll-Dahlgren, Guttersrud, et al., 2017; Austvoll-Dahlgren, Semakula, et al., 2017). We found that, even though over half of the parents reported having a bachelor's or more, their average proportion of

correct responses was only 64% (answered 35 out of 55 questions correctly). Some of the specific EBP principles parents struggled with included: “treatments should be compared fairly,” and “common practice doesn’t mean it’s beneficial/safe.” These findings are similar to results from a study assessing the critical appraisal abilities of adults in Norway (Dahlgren, Furuseth-Olsen, Rose, & Oxman, under review).

Study 1 findings suggest that college education alone may not be sufficient to adequately prepare parents to critically appraise healthcare claims. Notably, given the higher levels of education in the present sample (which is not uncommon for MTurk parent samples; Jensen-Doss et al., 2019) relative to the general population of parents in the U.S., it is possible that critical appraisal of healthcare claims may be even poorer among parents from the U.S. public (U.S. Census Bureau, 2019). Admittedly, consumer decision-making behavior in healthcare is nuanced, comprising of attitudes, beliefs, and psychosocial factors beyond sociodemographic characteristics such as socioeconomic status and educational attainment. Indeed, studies on parent vaccine attitudes are starting to find that parent education and other sociodemographic factors such as age, are not commonly significantly related to vaccine attitudes (in either direction) (Hornsey, Harris, & Fielding, 2018; Rozbroj, Lyons, & Lucke, 2019). Still, our findings indicate that U.S. parents do not understand and apply many EBP principles that are crucial for engaging in critical appraisal of healthcare claims and empowering parents to make informed health choices for themselves and their families.

Study 2

Our findings of low levels of critical appraisal amongst U.S. parents confirmed a need for a critical appraisal intervention. Thus, in Study 2 we developed, user-tested,

and through an online RCT, tested the efficacy of the *Parents Making Informed Health Choices Podcast*. The podcast included seven story-based episodes where different characters modeled how to critically appraise physical and mental health claims through the application of nine EBP principles (e.g., treatments should be compared fairly), which were also covered in the original Uganda podcast (Semakula et al., 2017). We included these nine EBP principles after establishing a need to increase critical appraisal around these principles in Study 1, and because they were also relevant to healthcare choices U.S. parents have to make. In addition to the seven main episodes, the *Parents Making Informed Health Choices Podcast* also included an introduction and a conclusion, with a total running time of 32 minutes. Communications theory on learning through entertainment partially informed the format of this study's critical appraisal intervention to ensure that messaging was not merely didactic, and there was a sense of relatability in the storylines and characters (Moyer-Gusé, 2008).

Parents who participated in user-testing provided generally very positive podcast feedback. Nevertheless, parents also suggested shortening the overall podcast length, especially the length of the conclusion episode, as well as improving the dialogue/acting of the voice actors to sound more natural, and less contrived or "like reading a script." Although in our short timeframe we were only able to recruit 5 of the 10 user-testing parents we had planned for, their feedback was an invaluable part of the podcast development process. As such, future consumer-directed intervention efforts should prioritize user-testing and start recruitment for this early in the process, perhaps even before the prototype is finished. Guiding frameworks for conducting user-testing of critical appraisal learning resources are starting to be developed, expanded, and

refined (Nsangi et al., 2020; Rosenbaum et al., 2019; Semakula et al., 2019).

After updating and finalizing the podcast, we tested its efficacy via an online RCT where parents (18 years or older) of children below age 18, who had not already participated in Study 1, were randomly assigned to listen to the *Parents Making Informed Health Choices Podcast* or an inert, control podcast of the same length (to control for time listening/paying attention).

The primary aim of Study 2 was to examine the efficacy of the critical appraisal podcast. Current findings support our hypothesis regarding its potential effect on U.S. parents' critical appraisal of healthcare claims. After listening to the podcasts, parents in the critical appraisal podcast condition performed significantly better on the critical appraisal test than parents in the control condition. When we examined critical appraisal as a dichotomous variable, with scores falling in either "low" or "high" categories, we also found that parents who listened to the critical appraisal podcast were more likely to obtain "high scores." Our findings with U.S. parents add to the growing body of empirical evidence demonstrating the efficacy of critical appraisal learning resources designed for the lay public (Castle et al., 2017; Chalmers et al., 2019; Cusack et al., 2018). The critical appraisal effects observed here are similar to those observed in the original Uganda podcast where they also found a significant difference in post-intervention critical appraisal performance favoring parents in the critical appraisal condition.

It should be noted that the *Parents Making Informed Health Choices Podcast* achieved this effect through an online, one-time, brief, audio intervention whereas the original Uganda podcast RCT utilized a much more controlled and lengthier podcast

listening procedure. Parents in the Uganda study were supervised when listening to two new episodes of the podcast and a recap episode per week across a 7 – 10 ten-week span. The Uganda podcast developers chose this more controlled and lengthier listening procedure given past findings that indicated single-session interventions for medical professionals were not efficacious (Ilic, Tepper, & Misso, 2012). In our study, we chose to prioritize ecological validity based on the fact that U.S. parents are ubiquitously exposed to and seek out health information on the internet. Additionally, after study completion, we planned to disseminate the podcast solely through the internet. The current study provides preliminary evidence that a single-session critical appraisal intervention geared toward internet-using parents may work. We must note, however, that parents in this study were tested on their critical appraisal abilities a few minutes after listening to the podcasts. As such, a follow-up study is necessary to examine long-term effects on critical appraisal as well as behavioral intention, attitudes, and treatment preference outcomes.

We also acknowledge that although parents in the control condition earned a significantly lower critical appraisal score (63% compared to 74%), this performance is not severely poor, and it is similar to the performance of Study 1 participants on this subset of critical appraisal questions (65%). The *moderately* low critical appraisal performance of participants in the control condition and Study 1 could be due to the critical appraisal questions being too easy or not sufficiently capturing the nuance of encountering health claims in everyday life; it could also have something to do with the higher levels of educational attainment in the present samples.

An additional point to highlight is that the present trial included three mental

health-related questions in the critical appraisal test in addition to 18 physical health-related questions, and this measure had overall good internal consistency. This suggests that capturing both psychological and medical aspects of healthcare through both an intervention and a measure is not only important, but feasible. Thus, given the particularly problematic acceptance of pseudoscientific practices in behavioral health services, future research should focus on further developing critical appraisal questions that capture a wide range of psychological conditions.

Next, we examined and found an intended behavior difference by condition after parents listened to the podcasts. Out of the three intended behaviors assessed, parents who listened to the critical appraisal podcast were more likely to report intending to find out if a treatment claim is based on a fair comparison study. Following Semakula and colleagues' approach, we also examined their intention to find out what a claim is based on, and intention to participate in a fair comparison study for a hypothetical illness they have. Contrary to hypotheses, these two other intended behaviors were not significantly different between the two study conditions. A possible explanation for this may be that the podcast directly and repeatedly mentions "fair comparisons" whereas the other intended behaviors are less frequently mentioned. These findings also differ from the Uganda podcast RCT, where they did not find any significant behavioral intention differences between study conditions. This may be due to a difference in our measurement of behavioral intention in that we examined it as a continuous variable as opposed to creating a dichotomous category (unlikely versus likely) as done by Semakula et al (2017). However, creating such a dichotomous grouping does not properly capture the range of higher intentionality represented by the difference in choosing the "very

likely” versus “likely” category. In the present sample 72% of parents in the critical appraisal podcast condition chose “very likely,” and 20% chose “likely;” whereas in the control condition, 56% of parents chose “very likely,” and 32% chose “likely.”

Admittedly, the real-world, actual behavior implication for choosing a “likely” versus “very likely” response on a single question regarding *intended* behavior remains elusive.

We also examined between group differences in EBP attitudes, attitudes regarding empiricism in mental health treatment, vaccine attitudes, and treatment preferences after parents listened to the podcasts. We found that, compared to parents who listened to the control podcast, parents who listened to the critical appraisal podcast had less negative personal-level attitudes toward EBPs, and reported higher preference for receiving effective therapies backed by scientific studies. The CAEBS personal-level attitudes toward EBPs scale included items such as, “I don’t feel comfortable making treatment decisions,” and “Regardless of what the evidence says, I know what works best for me,” which are in line with the critical appraisal’s podcast overall theme of individual action in healthcare decision-making.

Contrary to expectations, participants did not differ by condition with respect to any other EBP attitudes and vaccine attitudes. We also found that overall satisfaction with the podcasts was similar between the two conditions. The null findings regarding most EBP attitudes as measured by 3 out of the 4 CAEBS scales were surprising given the *general* similarity between the concepts covered in the critical appraisal podcast and the factors assessed by the CAEBS (Teh et al., 2016). Nevertheless, upon closer scrutiny of the CAEBS’ scales and items, some possible explanations become more apparent. First, while the *Parents Making Informed Health Choices Podcast* centers on EBP

principles and making them accessible to lay parents, it rarely mentions the term “EBP.” This is because the podcast explains, models action (i.e., engagement in critical appraisal), and makes specific recommendations about what to ask rather than *telling* the audience they should seek out EBPs. On the other hand, the CAEBS frequently and explicitly mentions the term EBP (and defines it in the instructions). Perhaps this suggests that evidence-based healthcare proponents and researchers have some middle ground to reach in regards to balancing efforts to teach a *science and health literacy-based skill* (i.e., critical appraisal of healthcare claims) and increasing consumer knowledge of *healthcare terms*. It would be interesting and elucidating for future studies to examine whether explicitly and repeatedly mentioning EBPs in the podcast has a different effect on EBP attitudes as measured by the CAEBS. Another potential explanation of the null findings is that CAEBS items have a strong emphasis on society and policy themes, whereas the critical appraisal podcast solely focused on individual actions around EBP principles.

An exploratory aim of this study was also to examine sociodemographic and other correlates of critical appraisal and moderators of podcast effects. We found various correlates of critical appraisal, but none of the proposed variables attained significance as moderators of podcast effects. Better critical appraisal abilities were related to the following factors: 1) taking medication for a psychological condition, 2) higher intention to find out if a treatment claim is based on a fair comparison study, 3) more positive EBP attitudes on the Beliefs Regarding Therapists’ Practices CAEB scale, 4) less negative EBP attitudes on the Negative Personal-Level Attitudes towards EBPs CAEBS scale, 5) more negative attitudes on the Negative Societal-Level Attitudes towards EBPs CAEBS

scale, 6) more positive attitudes regarding empiricism in mental health treatment, 7) more favorable attitudes toward vaccines, 8) greater preference for receiving effective therapies backed by scientific studies, 9) less preference for having a therapist with many years of professional experience, and 10) less preference for a therapist allowing them to speak the majority of the therapy session. Although these were exploratory examinations given the limited extant research on critical appraisal in U.S. parents, the direction of these relationships is mostly what would be expected given outcome results previously described. The only unexpected relationship directions were that better critical appraisal was related to more negative EBP attitudes on the Negative Societal-Level Attitudes CAEBS scale, and lower critical appraisal was related positive EBP attitudes on the Attitudes about Mental Health Policy scale. This CAEBS scale included five items (e.g., “all cultures might not believe in EBPs”) which again appear to tap into themes not covered by the podcast. Given that this is the first time these relationships with critical appraisal have been examined, they merit replication.

Notably, we did not find that the podcast had an effect on several of these critical appraisal correlates. Thus, future studies should consider exploring potential alterations to the critical appraisal podcast content to increase its impact on critical appraisal correlates. Interestingly, all of these critical appraisal correlates are specifically about mental health care topics, which was a unique expansion of this study as no previous study has examined critical appraisal alongside mental health-related constructs.

We also conducted post hoc analyses examining podcast effects and correlates of critical appraisal without the intent to treat sample of 49 parents who failed one or more validity measures. We included validity measures in accordance with best practices

guidelines for conducting experimental research online, especially on MTurk. However, we chose to keep these individuals in the primary analyses in light of best practices guidelines for RCTs which recommend including all participants who were randomized into a condition in analyses to decrease the bias their exclusion may introduce. In the analyses excluding the intent to treat sample, we found the following: a) the magnitude of podcast intervention effect increased from 7% to 11%, b) no intervention effect on the CAEBS Negative Personal-Level Attitudes Toward EBPs scale, c) intervention effect on attitudes regarding empiricism in mental health treatment, d) correlation between taking medication for a psychological condition and higher critical appraisal scores, and f) no correlation between the CAEBS Attitudes about Mental Health Policy scale and lower critical appraisal. The increase in the effect of the *Parents Making Informed Health Choices Podcast* after excluding the intent to treat sample questionable validity data suggests that more controlled podcast listening procedures (such as the one implemented in the Uganda RCT by Semakula and colleagues) where participants are less likely to become distracted when listening, may be important to establish the efficacy of new critical appraisal interventions such as a mass media podcast. (Of note, in the current trial participants could not fast forward or move on from the podcast listening page until 32 minutes had lapsed from first entering the page). However, as previously noted, examining the effectiveness of these interventions under real-world conditions where parents will not be provided significant assistance in paying attention is also as important—if not—more important. Notably, it is also a possibility that parents could have “tuned out” during podcast listening but still completed the post intervention measures carefully and/or correctly.

Future studies should also determine who is most likely to benefit from these critical appraisal interventions; for example, including a sample with a broader range of levels of educational attainment to determine if education is a moderator. Another important future direction for critical appraisal research is how to incorporate pragmatic tips on how parents can carry out critical appraisal in conversations with healthcare providers, especially in situations where the evidence-base may not be strong, but where a treatment is still recommended either because it is the only option or for other valid reasons. There are EBP principles that address these realities (e.g., EBP principle “how certain is the evidence?” Chalmers et al., 2018), and future podcast development should focus on creating short episodes that capture these EBP principles in decision-making. Furthermore, critical appraisal interventions could complement a shared-decision making framework (Cheng et al., 2017; Elwyn et al., 2012) well in that critical appraisal interventions encourage patient activation with a provider. Although notably, consumer-focused critical appraisal interventions such as the *Parents Making Informed Health Choices Podcast* do emphasize question asking and independent action more than collaborating with providers on a decision. Thus, future studies should focus on combining or embedding critical appraisal interventions within shared decision-making frameworks. Indeed, although this is not directly communicated in the *Parents Making Informed Health Choices Podcast*, characters in the stories often do model appropriate interactions with healthcare providers, however, the storylines do not conclude in a shared decision about healthcare services.

Limitations and Future Directions

As with any investigation, this study has several limitations that suggest additional directions for future research. First, the current online-convenience sample is not representative of all U.S. parents; especially of parents in low-resource, community settings. Thus, future studies should focus on testing the efficacy of the podcast with more representative samples of parents from community settings. Along those lines, the *Parents Making Informed Health Choices Podcast* is currently only available to English-speaking parents, creating a disparity for parents who speak other languages, especially predominantly Spanish-speaking parents, who make up the fastest growing linguistic population in the U.S. Future efforts should focus on translating these materials and testing their efficacy in other languages, especially with Spanish-speaking parents. Some of this work is already starting with children in Spain (García et al., 2019). Second, currently we have found evidence of only the very short-term efficacy of the *Parents Making Informed Health Choices Podcast*. Therefore, an important future direction is to conduct a long-term follow up assessment of critical appraisal with parents who participated in the current online RCT, especially given that a one-year follow up study on the effects of the Uganda critical appraisal podcast found considerable decline in critical appraisal abilities (Semakula et al., 2020). Third, we only measure self-reported, not actual, critical appraisal and application of EBP principles in parents' daily lives. Finally, all outcome variables were about parent attitudes and preferences regarding mental health services for themselves, but we did not specifically ask about their attitudes and preferences in regards to mental health services for their children, which may be an important addition to future studies.

Conclusion

This investigation provides empirical evidence for a need to increase critical appraisal of healthcare claims among U.S. parents, as well as initial support for the efficacy of a podcast designed to meet this need. We found that listening to a relatively brief podcast improved parents' ability to critically appraise healthcare claims, as well as increased self-reported intended behavior, positive attitudes toward evidence-based practices, and preference for evidence-based practices. Additional trials with larger, more diverse community samples, are needed. Engaging in critical appraisal of healthcare claims is essential for parents to make informed health choices for themselves and their families. Accordingly, efforts are underway to disseminate the *Parents Making Informed Health Choices Podcast* through consumer-focused websites.

References

- Alegría, M., Carson, N., Flores, M., Li, X., Shi, P., Lessios, A. S., . . . Interian, A. (2014). Activation, self-management, engagement, and retention in behavioral health care: A randomized clinical trial of the DECIDE intervention. *JAMA Psychiatry, 71*(5), 557-565.
- Amazon Mechanical Turk. (2020). Important updates on MTurk marketplace integrity, worker identity and requester tools to manage task quality. Retrieved from <https://blog.mturk.com/important-updates-on-mturk-marketplace-integrity-worker-identity-and-requester-tools-to-manage-206e4e90da0c>
- APA Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *American Psychologist, 61*(4), 271-285.
- Arch, J. J., Twohig, M. P., Deacon, B. J., Landy, L. N., & Bluett, E. J. (2015). The credibility of exposure therapy: Does the theoretical rationale matter? *Behaviour Research and Therapy, 72*, 81-92.
- Arditte, K. A., Çek, D., Shaw, A. M., & Timpano, K. R. (2016). The importance of assessing clinical phenomena in Mechanical Turk research. *Psychological Assessment, 28*(6), 684.
- Austvoll-Dahlgren, A., Guttersrud, Ø., Nsangi, A., Semakula, D., & Oxman, A. D. (2017). Measuring ability to assess claims about treatment effects: a latent trait analysis of items from the ‘Claim Evaluation Tools’ database using Rasch modelling. *BMJ Open, 7*(5), 1-10.
- Austvoll-Dahlgren, A., Nsangi, A., & Semakula, D. (2016). Interventions and assessment tools addressing key concepts people need to know to appraise claims about treatment effects: a systematic mapping review. *Systematic Reviews, 5*(1), 215.
- Austvoll-Dahlgren, A., Semakula, D., Nsangi, A., Oxman, A. D., Chalmers, I., Rosenbaum, S., & Guttersrud, Ø. (2017). Measuring ability to assess claims about treatment effects: the development of the ‘Claim Evaluation Tools’. *BMJ Open, 7*(5), e013184.
- Austvoll-Dahlgren, A., Oxman, A. D., Chalmers, I., Nsangi, A., Glenton, C., Lewin, S., . . . Sewankambo, N. (2015). Key concepts that people need to understand to assess claims about treatment effects. *Journal of Evidence-Based Medicine, 8*(3), 112-125.
- Barry, M. J., & Edgman-Levitan, S. (2012). Shared decision making—the pinnacle of patient-centered care. *New England Journal of Medicine, 366*(9), 780-781.

- Barry, M. M., Sixsmith, J., & D'Eath, M. (2012). *A rapid evidence review of interventions for improving health literacy*. Retrieved from <https://www.ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/1205-TER-Improving-Health-Literacy.pdf>
- Becker, S. J. (2015). Direct-to-consumer marketing: A complementary approach to traditional dissemination and implementation efforts for mental health and substance abuse interventions. *Clinical Psychology: Science and Practice*, 22(1), 85-100.
- Becker, S. J., Spirito, A., & Vanmali, R. (2016). Perceptions of 'Evidence-Based Practice' among the consumers of adolescent substance use treatment. *Health Education Journal*, 75(3), 358-369.
- Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon. com's Mechanical Turk. *Political Analysis*, 20(3), 351-368.
- Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J., & Crotty, K. (2011). Low health literacy and health outcomes: an updated systematic review. *Annals of Internal Medicine*, 155(2), 97-107.
- Bohannon, J. (2016, June 7). Psychologists grow increasingly dependent on online research subjects. *Science Magazine*. Retrieved from <https://www.sciencemag.org/news/2016/06/psychologists-grow-increasingly-dependent-online-researchsubjects>.
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6(1), 3-5.
- Buhrmester, M. D., Talaifar, S., & Gosling, S. D. (2018). An evaluation of amazon's Mechanical Turk, its rapid rise, and its effective use. *Perspectives on Psychological Science*, 13(2), 149-154.
- Carman, K. L., Maurer, M., Mangrum, R., Yang, M., Ginsburg, M., Sofaer, S., . . . Richmond, J. (2016). Understanding an informed public's views on the role of evidence in making health care decisions. *Health Affairs*, 35(4), 566-574.

- Carman, K. L., Maurer, M., Yegian, J. M., Dardess, P., McGee, J., Evers, M., & Marlo, K. O. (2010). Evidence that consumers are skeptical about evidence-based health care. *Health Affairs*, *29*(7), 1400-1406.
- Castle, J. C., Chalmers, I., Atkinson, P., Badenoch, D., Oxman, A. D., Austvoll-Dahlgren, A., . . . Woloshin, S. (2017). Establishing a library of resources to help people understand key concepts in assessing treatment claims—the “Critical thinking and Appraisal Resource Library”(CARL). *PLoS One*, *12*(7), 1-12.
- Chalmers, I., Atkinson, P., Badenoch, D., Glasziou, P., Austvoll-Dahlgren, A., Oxman, A., & Clarke, M. (2019). The James Lind Initiative: books, websites and databases to promote critical thinking about treatment claims, 2003 to 2018. *Research Involvement and Engagement*, *5*(1), 6.
- Chalmers, I., Oxman, A. D., Austvoll-Dahlgren, A., Ryan-Vig, S., Pannell, S., Sewankambo, N., . . . Glasziou, P. (2018). Key concepts for informed health choices: a framework for helping people learn how to assess treatment claims and make informed choices. *BMJ Evidence-Based Medicine*, *23*(1), 29-33.
- Chandler, J., & Shapiro, D. (2016). Conducting clinical research using crowdsourced convenience samples. *Annual Review of Clinical Psychology*, *12*, 53-81.
- Chen, X., Goodson, P., & Acosta, S. (2015). Blending health literacy with an English as a second language curriculum: A systematic literature review. *Journal of Health Communication*, *20*(sup2), 101-111.
- Cheng, H., Hayes, D., Edbrooke-Childs, J., Martin, K., Chapman, L., & Wolpert, M. (2017). What approaches for promoting shared decision-making are used in child mental health? A scoping review. *Clinical Psychology & Psychotherapy*, *24*(6), 1-17.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (Second ed.). Hove, Hillsdale, NJ: Academic Press.
- Cusack, L., Del Mar, C. B., Chalmers, I., Gibson, E., & Hoffmann, T. C. (2018). Educational interventions to improve people’s understanding of key concepts in assessing the effects of health interventions: A systematic review. *Systematic Reviews*, *7*(1), 68.
- Dahlgren, A., Furuseth-Olsen, K., Rose, C. J., & Oxman, A. D. J. F. (under review). The Norwegian public’s ability to assess treatment claims: results of a cross-sectional study of critical health literacy. *F1000Research*, *9*, 179.

- Deen, D., Lu, W.-H., Rothstein, D., Santana, L., & Gold, M. R. (2011). Asking questions: the effect of a brief intervention in community health centers on patient activation. *Patient Education and Counseling*, *84*(2), 257-260.
- Dworkin, J., Hessel, H., Gliske, K., & Rudi, J. H. (2016). A comparison of three online recruitment strategies for engaging parents. *Family Relations*, *65*(4), 550-561.
- Edison Research. (2017). *The Infinite Dial 2017: The premier study of trends in media usage and consumption habits*. Retrieved from <https://www.edisonresearch.com/infinite-dial-2017/>
- Eliacin, J., Rollins, A. L., Burgess, D. J., Salyers, M. P., & Matthias, M. S. (2016). Patient activation and visit preparation in African American veterans receiving mental health care. *Cultural Diversity and Ethnic Minority Psychology*, *22*(4), 580.
- Elwyn, G., Frosch, D., Thomson, R., Joseph-Williams, N., Lloyd, A., Kinnersley, P., . . . Rollnick, S. (2012). Shared decision making: A model for clinical practice. *Journal of General Internal Medicine*, *27*(10), 1361-1367.
- Eysenbach, G., & Köhler, C. (2002). How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. *BMJ*, *324*(7337), 573-577.
- Eysenbach, G., Powell, J., Kuss, O., & Sa, E.-R. (2002). Empirical studies assessing the quality of health information for consumers on the world wide web: A systematic review. *JAMA*, *287*(20), 2691-2700.
- Farrell, N. R., & Deacon, B. J. (2016). The relative importance of relational and scientific characteristics of psychotherapy: Perceptions of community members vs. therapists. *Journal of Behavior Therapy and Experimental Psychiatry*, *50*, 171-177.
- Freiheit, S. R., Vye, C., Swan, R., & Cady, M. (2004). Cognitive-behavioral therapy for anxiety: Is dissemination working? *The Behavior Therapist*, *27*(2), 25-32.
- Friedberg, R. D., & Bayar, H. (2017). If It works for pills, can it work for skills? Direct-to-consumer social marketing of evidence-based psychological treatments. *Psychiatric Services*, *68*(6), 621-623.

- García, L. M., Alonso-Coello, P., Ministral, L. A., Ballesté-Delpierre, C., Aybar, C. C., de Britos, C., . . . Llaquet, P. (2019). Learning to make informed health choices: Protocol for a pilot study in schools in Barcelona. *Manuscript submitted for publication*.
- Glenton, C., Paulsen, E. J., & Oxman, A. D. (2005). Portals to Wonderland: Health portals lead to confusing information about the effects of health care. *BMC Medical Informatics and Decision Making*, 5(1), 7.
- Gust, D. A., Strine, T. W., Maurice, E., Smith, P., Yusuf, H., Wilkinson, M., . . . Schwartz, B. (2004). Underimmunization among children: Effects of vaccine safety concerns on immunization status. *Pediatrics*, 114(1), 16-22.
- Harik, J. M., Hundt, N. E., Bernardy, N. C., Norman, S. B., & Hamblen, J. L. (2016). Desired involvement in treatment decisions among adults with PTSD symptoms. *Journal of Traumatic Stress*, 29(3), 221-228.
- Hibbard, J. H., & Mahoney, E. (2010). Toward a theory of patient and consumer activation. *Patient Education and Counseling*, 78(3), 377-381.
- Hipol, L. J., & Deacon, B. J. (2013). Dissemination of evidence-based practices for anxiety disorders in Wyoming: A survey of practicing psychotherapists. *Behavior Modification*, 37(2), 170-188.
- Hoffmann, T. C., & Del Mar, C. (2015). Patients' expectations of the benefits and harms of treatments, screening, and tests: A systematic review. *JAMA Internal Medicine*, 175(2), 274-286.
- Hornsey, M. J., Harris, E. A., & Fielding, K. S. (2018). The psychological roots of anti-vaccination attitudes: A 24-nation investigation. *Health Psychology*, 37(4), 307-315.
- Horsley, T., Hyde, C., Santesso, N., Parkes, J., Milne, R., & Stewart, R. (2011). Teaching critical appraisal skills in healthcare settings. *Cochrane Database Systematic Reviews*.
- Ilic, D., Tepper, K., & Misso, M. (2012). Teaching evidence-based medicine literature searching skills to medical students during the clinical years: A randomized controlled trial. *Journal of the Medical Library Association*, 100(3), 190-196.
- Institute of Medicine (US) Committee on Quality of Health Care in America. (2001). *Crossing the Quality Chasm: A New Health System for the 21st Century*.: National Academies Press (US),.

- Jacque, B., Koch-Weser, S., Faux, R., & Meiri, K. (2016). Addressing health literacy challenges with a cutting-edge infectious disease curriculum for the high school biology classroom. *Health Education & Behavior, 43*(1), 43-53.
- Jamtvedt, G., Klemp, M., Mørland, B., & Nylenna, M. (2015). Responsibility and accountability for well informed health-care decisions: a global challenge. *The Lancet, 386*(9995), 826-828.
- Jensen-Doss, A., Patel, Z., Casline, E., Mora Ringle, V., & Timpano, K. (2019). Using mechanical turk to study parents and children: an examination of data quality and representativeness. *Manuscript submitted for publication.*
- Kennedy, R., Clifford, S., Burleigh, T., Waggoner, P. D., Jewell, R., & Winter, N. J. (2020). The shape of and solutions to the MTurk quality crisis. *Political Science Research and Methods, 1-16.*
- Kirk, A., Broman-Fulks, J., & Bergquist, J. (2016). National attitudes towards mental health treatment: the importance of research evidence. *Cognitive Behaviour Therapy, 45*(6), 458-472.
- Kitchener, B. A., & Jorm, A. F. (2002). Mental health first aid training for the public: evaluation of effects on knowledge, attitudes and helping behavior. *BMC Psychiatry, 2*(1), 10.
- Kitchener, B. A., & Jorm, A. F. (2006). Mental health first aid training: review of evaluation studies. *Australian and New Zealand Journal of Psychiatry, 40*(1), 6-8.
- Levay, K. E., Freese, J., & Druckman, J. N. (2016). The demographic and political composition of Mechanical Turk samples. *Sage Open, 6*(1), 1-17.
- Lilienfeld, S. (2007). Psychological treatments that cause harm. *Perspectives on Psychological Science, 2*(1), 53-70.
- Lilienfeld, S. (2012). Public skepticism of psychology: why many people perceive the study of human behavior as unscientific. *American Psychologist, 67*(2), 111.
- Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods, 49*(2), 433-442.
- Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods, 44*(1), 1-23.

- Milne, R., & Oliver, S. (1996). Evidence-based consumer health information: developing teaching in critical appraisal skills. *International Journal for Quality in Health Care*, 8(5), 439-445.
- Mora Ringle, V., Walsh, L. M., Maxwell, C. A., Smith, A. M., Grossman, R. A., Becker, S. J., & Jensen-Doss, A. (2020). Understanding of evidence-based mental health care and the perceived importance of scientific information in a sample of US adults. *Journal of Clinical Psychology*, 76(1), 161-175.
- Moran, M. B., Frank, L. B., Chatterjee, J. S., Murphy, S. T., & Baezconde-Garbanati, L. (2016). Information scanning and vaccine safety concerns among African American, Mexican American, and non-Hispanic White women. *Patient Education and Counseling*, 99(1), 147-153.
- Moyer-Gusé, E. (2008). Toward a theory of entertainment persuasion: Explaining the persuasive effects of entertainment-education messages. *Communication Theory*, 18(3), 407-425.
- Nordheim, L. V., Gundersen, M. W., Espehaug, B., Gutterud, Ø., & Flottorp, S. (2016). Effects of school-based educational interventions for enhancing adolescents abilities in critical appraisal of health claims: A systematic review. *PLoS One*, 11(8), 1-21.
- Nsangi, A., Semakula, D., Oxman, A. D., Austvoll-Dahlgren, A., Oxman, M., Rosenbaum, S., . . . Kaseje, M. (2017). Effects of the Informed Health Choices primary school intervention on the ability of children in Uganda to assess the reliability of claims about treatment effects: A cluster-randomised controlled trial. *The Lancet*, 390(10092), 374-388.
- Nsangi, A., Semakula, D., Oxman, A. D., & Sewankambo, N. K. (2015). Teaching children in low-income countries to assess claims about treatment effects: Prioritization of key concepts. *Journal of Evidence-Based Medicine*, 8(4), 173-180.
- Nsangi, A., Semakula, D., Rosenbaum, S. E., Oxman, A. D., Oxman, M., Morelli, A., . . . Uwitonze, A.-M. (2020). Development of the informed health choices resources in four countries to teach primary school children to assess claims about treatment effects: A qualitative study employing a user-centred approach. *Pilot and Feasibility Studies*, 6(18), 1-15.
- Nutbeam, D. (2008). The evolving concept of health literacy. *Social Science & Medicine*, 67(12), 2072-2078.

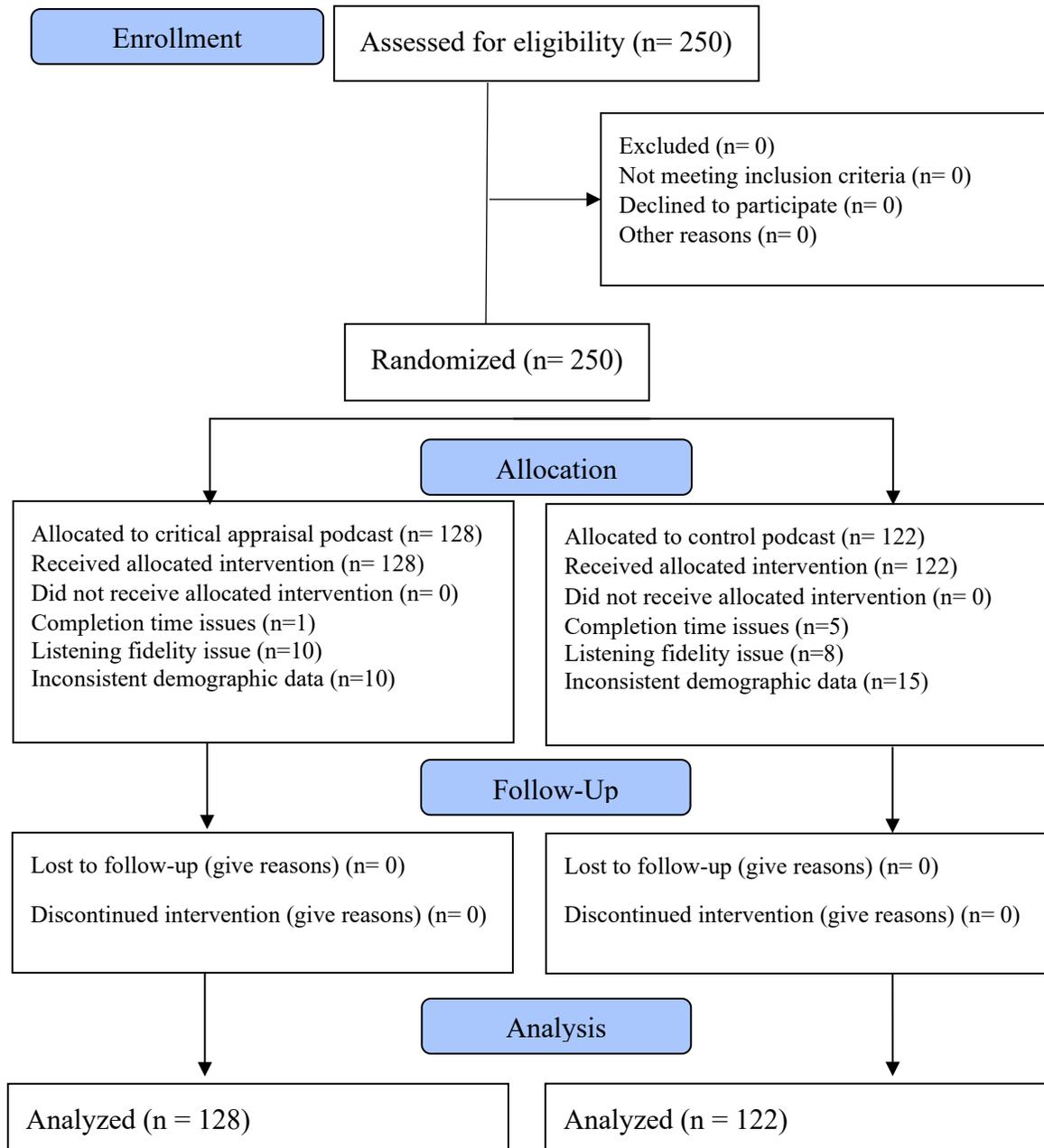
- Rosenbaum, S. (2010). *Improving the user experience of evidence: A design approach to evidence-informed health care*. (Unpublished doctoral dissertation), The Oslo School of Architecture and Design, Oslo, Norway.
- Rosenbaum, S., Oxman, M., Oxman, A., Chelagat, F., Mugisha, M., Ssenyonga, R., . . . Semakula, D. (2019). Human-centred design development of Informed Health Choices (IHC) learning resources for secondary school students. *Manuscript submitted for publication*.
- Rozbroj, T., Lyons, A., & Lucke, J. (2019). Psychosocial and demographic characteristics relating to vaccine attitudes in Australia. *Patient Education and Counseling*, *102*(1), 172-179.
- Santucci, L. C., McHugh, R. K., & Barlow, D. H. (2012). Direct-to-consumer marketing of evidence-based psychological interventions: Introduction. *Behavior Therapy*, *43*(2), 231-235.
- Sayer, N. A., Rosen, C. S., Bernardy, N. C., Cook, J. M., Orazem, R. J., Chard, K. M., . . . Crowley, J. (2017). Context matters: team and organizational factors associated with reach of evidence-based psychotherapies for PTSD in the veterans health administration. *Administration and Policy in Mental Health and Mental Health Services Research*, 1-15.
- Scheyett, A., McCarthy, E., & Rausch, C. (2006). Consumer and family views on evidence-based practices and adult mental health services. *Community Mental Health Journal*, *42*(3), 243-257.
- Schleider, J. L., & Weisz, J. R. (2015). Using Mechanical Turk to study family processes and youth mental health: A test of feasibility. *Journal of Child and Family Studies*, *24*(11), 3235-3246.
- Schofield, C. A., Moore, C. D., Hall, A., & Coles, M. E. (2016). Understanding perceptions of anxiety disorders and their treatment. *The Journal of Nervous and Mental Disease*, *204*(2), 116-122.
- Schofield, C. A., Weis, H., Ponzini, G., & McHugh, R. K. (2017). Improving perceptions of cognitive-behavioral therapy with brief website exposure: Does neuroscientific allure impact attitudes? *Journal of Behavior Therapy and Experimental Psychiatry*, *57*, 53-61.
- Schwartz, K. L., Roe, T., Northrup, J., Meza, J., Seifeldin, R., & Neale, A. V. (2006). Family medicine patients' use of the Internet for health information: a MetroNet study. *The Journal of the American Board of Family Medicine*, *19*(1), 39-45.

- Seçkin, G., Yeatts, D., Hughes, S., Hudson, C., & Bell, V. (2016). Being an informed consumer of health information and assessment of electronic health literacy in a national sample of internet users: validity and reliability of the e-HLS instrument. *Journal of Medical Internet Research, 18*(7).
- Semakula, D., Nsangi, A., Oxman, A. D., Oxman, M., Austvoll-Dahlgren, A., Rosenbaum, S., . . . Kaseje, M. (2017). Effects of the Informed Health Choices podcast on the ability of parents of primary school children in Uganda to assess claims about treatment effects: a randomised controlled trial. *The Lancet, 390*(10092), 389-398.
- Semakula, D., Nsangi, A., Oxman, A. D., Oxman, M., Austvoll-Dahlgren, A., Rosenbaum, S., . . . Nyirazinyoye, L. (2020). Effects of the Informed Health Choices podcast on the ability of parents of primary school children in Uganda to assess the trustworthiness of claims about treatment effects: one-year follow up of a randomised trial. *Trials, 21*(1), 1-18.
- Semakula, D., Nsangi, A., Oxman, A. D., & Sewankambo, N. K. (2015). Priority setting for resources to improve the understanding of information about claims of treatment effects in the mass media. *Journal of Evidence-Based Medicine, 8*(2), 84-90.
- Semakula, D., Nsangi, A., Oxman, M., Austvoll-Dahlgren, A., Rosenbaum, S., Kaseje, M., . . . Oxman, A. D. (2017). Can an educational podcast improve the ability of parents of primary school children to assess the reliability of claims made about the benefits and harms of treatments: study protocol for a randomised controlled trial. *Trials, 18*(1), 31.
- Semakula, D., Nsangi, A., Oxman, M., Rosenbaum, S. E., Oxman, A. D., Austvoll-Dahlgren, A., . . . Morelli, A. (2019). Development of mass media resources to improve the ability of parents of primary school children in Uganda to assess the trustworthiness of claims about the effects of treatments: a human-centred design approach. *Pilot and Feasibility Studies, 5*(1), 1-17.
- Sillence, E., Briggs, P., Harris, P. R., & Fishwick, L. (2007). How do patients evaluate and make use of online health information? *Social Science & Medicine, 64*(9), 1853-1862.
- Stacey, D., Légaré, F., Col, N. F., Bennett, C. L., Barry, M. J., Eden, K. B., . . . Thomson, R. (2014). Decision aids for people facing health treatment or screening decisions. *Cochrane Database Systematic Reviews*(4).
- Swan, L. K., & Heesacker, M. (2013). Evidence of a pronounced preference for therapy guided by common factors. *Journal of Clinical Psychology, 69*(9), 869-879.

- Swift, J. K., & Callahan, J. L. (2010). A comparison of client preferences for intervention empirical support versus common therapy variables. *Journal of Clinical Psychology, 66*(12), 1217-1231.
- Tanenbaum, S. (2008). Perspectives on evidence-based practice from consumers in the US public mental health system. *Journal of Evaluation in Clinical Practice, 14*(5), 699-706.
- Teh, L. B., Hayashi, K., Latner, J., & Mueller, C. W. (2016). Consumer attitudes towards evidence based mental health services among American mental health consumers. *International Journal of Mental Health Nursing, 25*(5), 462-471.
- Thomas, K. C., Stein, G. L., Williams, C. S., Jolles, M. P., Sleath, B. L., Martinez, M., . . . Morrissey, J. P. (2017). Fostering activation among Latino parents of children with mental health needs: an RCT. *Psychiatric Services, 68*(10), 1068-1075.
- U.S. Census Bureau. (2019). *Parents With Coresident Children Under 18, by Living Arrangement, Sex, and Selected Characteristics*. Retrieved from <https://www.census.gov/data/tables/2019/demo/families/cps-2019.html>
- Vernon, J. A., Trujillo, A., Rosenbaum, S. J., & DeBuono, B. (2007). Low health literacy: Implications for national health policy. Washington, DC: Department of Health Policy, School of Public Health and Health Services, The George Washington University.
- Walters, K., Christakis, D. A., & Wright, D. R. (2018). Are Mechanical Turk worker samples representative of health status and health behaviors in the US? *PLoS One, 13*(6), 1-10.

Figures

Figure 1.
Consort Diagram



Tables

Table 1

Informed Health Choices Key Concepts (Evidence-Based Practice Principles)
(Austvoll-Dahlgren et al., 2015)

1. Recognizing the need for fair comparisons of treatments

- 1.1 Treatments may be harmful
- 1.2 Personal experiences or anecdotes (stories) are an unreliable basis for assessing the effects of most treatments
- 1.3 An ‘outcome’ may be associated with a treatment, but not caused by the treatment
- 1.4 Widely used treatments or treatments that have been used for a long time are not necessarily beneficial or safe
- 1.5 New, brand-named, or more expensive treatments may not be better than available alternatives
- 1.6 Opinions of experts or authorities do not alone provide a reliable basis for deciding on the benefits and harms of treatments
- 1.7 Conflicting interests may result in misleading claims about the effects of treatments
- 1.8 Increasing the amount of a treatment does not necessarily increase the benefits of a treatment and may cause harm
- 1.9 Earlier detection of disease is not necessarily better
- 1.10 Hope or fear can lead to unrealistic expectations about the effects of treatments
- 1.11 Beliefs about how treatments work are not reliable predictors of the actual effects of treatments
- 1.12 Large, dramatic effects of treatments are rare

2. Judging whether a comparison of treatment is a fair comparison

- 2.1 Evaluating the effects of treatments requires appropriate comparisons
- 2.2 Apart from the treatments being compared, the comparison groups need to be similar (i.e. 'like needs to be compared with like')
- 2.3 People’s outcomes should be counted in the group to which they were allocated

2.4 People in the groups being compared need to be cared for similarly (apart from the treatments being compared)

2.5 If possible, people should not know which of the treatments being compared they are receiving

2.6 Outcomes should be measured in the same way (fairly) in the treatment groups being compared

2.7 It is important to measure outcomes in everyone who was included in the treatment comparison groups

3. Understanding the role of chance

3.1 Small studies in which few outcome events occur are usually not informative and the results may be misleading

3.2 The use of p values to indicate the probability of something having occurred by chance may be misleading; CIs are more informative

3.3 Saying that a difference is statistically significant or that is not statistically significant can be misleading

4. Considering all of the relevant fair comparisons

4.1 The results of single tests of treatments can be misleading

4.2 Reviews of treatment tests that do not use systematic methods can be misleading

4.3 Well-performed systematic reviews often reveal a lack of relevant evidence, but they provide the best basis for making judgements about the certainty of the evidence

5. Understanding the results of fair comparisons of treatments

5.1 Treatments may have beneficial and harmful effects

5.2 Relative effects of treatments alone can be misleading

5.3 Average differences between treatments can be misleading

6. Judging whether fair comparisons of treatments are relevant

6.1 Fair comparisons of treatments should measure outcomes that are important

6.2 Fair comparisons of treatments in animals or highly selected groups of people may not be relevant

6.3 The treatments evaluated in fair comparisons may not be relevant or applicable

6.4 Results for a selected group of people within fair comparisons can be misleading

Table 2

EBP Principles and Examples of Claims Covered in the Parents Making Informed Health Choices Podcast and the Informed Health Choices Podcast (Semakula et al., 2017)

EBP Principle	U.S. Claim Examples	Uganda Claim Examples
(1) Treatments should be compared	Elderberry is an effective treatment for child flu.	Quail eggs make you very strong.
(2) Treatments should be compared fairly	Cognitive-Behavioral Therapy for youth anxiety and depression works.	Group support is helpful for someone who is depressed.
(3) Findings from small studies can be misleading	We can know that vaccines cause autism based on information from one small study.	Washing hands with soap does not stop children from getting diarrhea.
(4) Association is not the same as causation*	Contraceptive pills cause women to gain weight.*	A lot of women gain weight when they take contraceptive pills.
(5) Expert opinion is not always right*	Contraceptive pills cause women to gain weight.* Butter can heal burns.	Eating some hot pepper will heal ulcers. Putting cooking oil on a burn will heal it.
(6) Anecdotes are unreliable evidence	An herbal treatment for ADHD with no side effects exists.*	Quinine can cure malaria. It can also give you nausea and make you vomit.
(7) Treatments might be harmful*	Herbal treatment for ADHD with no side effects exists.*	Herbal medicines exist for malaria treatment that cure malaria and do not have any bad effects.
(8) Treatments have benefits and harms*	Physical discipline is the best strategy for managing child behavior problems.	An herbal treatment called kyogero stops babies from getting infections.

*Covered in the same episode of the *Parents Making Informed Health Choices Podcast*

Table 3
Proportion of correct responses per EBP principle

EBP Principle	Proportion of correct responses on two EBP principle items
Recognizing an unreliable basis for a claim	
1.1 Treatments may be harmful	0.715*+
1.2 Personal experiences or anecdotes (stories) are an unreliable basis for assessing the effects of most treatments	0.677*+
1.3 An ‘outcome’ may be associated with a treatment, but not caused by the treatment	0.670*+
1.4 Widely used treatments or treatments that have been used for a long time are not necessarily beneficial or safe	0.639*+
1.5 New, brand-named, or more expensive treatments may not be better than available alternatives	0.701
1.6 Opinions of experts or authorities do not alone provide a reliable basis for deciding on the benefits and harms of treatments	0.757*
1.7 Conflicting interests may result in misleading claims about the effects of treatments	0.805
1.8 Increasing the amount of a treatment does not necessarily increase the benefits of a treatment and may cause harm	0.729
1.9 Earlier detection of disease is not necessarily better	NA
1.10 Hope or fear can lead to unrealistic expectations about the effects of treatments	0.653
1.11 Beliefs about how treatments work are not reliable predictors of the actual effects of treatments	0.701
1.12 Large, dramatic effects of treatments are rare	0.529

Understanding whether comparisons are fair and reliable

2.1 Evaluating the effects of treatments requires appropriate comparisons	0.552*+
2.2 Apart from the treatments being compared, the comparison groups need to be similar (i.e. 'like needs to be compared with like')	0.632*+
2.3 People's outcomes should be counted in the group to which they were allocated	NA
2.4 People in the groups being compared need to be cared for similarly (apart from the treatments being compared)	0.687
2.5 If possible, people should not know which of the treatments being compared they are receiving	0.635
2.6 Outcomes should be measured in the same way (fairly) in the treatment groups being compared	0.718
2.7 It is important to measure outcomes in everyone who was included in the treatment comparison groups	NA
2.8 The results of single comparisons of treatments can be misleading	0.639*+
2.9 Reviews of treatment comparisons that do not use systematic methods can be misleading	0.562
2.10 Unpublished results of fair comparisons may result in biased estimates of treatment effects	0.531
2.11 Results for a selected group of people within a systematic review of fair comparisons of treatments can be misleading	NA
2.12 Relative effects of treatments alone can be misleading	0.483
2.13 Average differences between treatments can be misleading	NA
2.14 Small studies in which few outcome events occur are usually not informative and the results may be misleading	0.687

2.15 The use of p-values to indicate the probability of something having occurred by chance may be misleading; confidence intervals are more informative	NA
2.16 Saying that a difference is statistically significant or that it is not statistically significant can be misleading	0.514
2.17 A lack of evidence is not the same as evidence of “no difference”	NA
Making informed choices	
3.1 A systematic review of fair comparisons of treatments should measure outcomes that are important	0.567
3.2 A systematic review of fair comparisons of treatments in animals or highly selected groups of people may not be relevant	0.611
3.3 The treatments evaluated in fair comparisons may not be relevant or applicable	0.601
3.4 Well done systematic reviews often reveal a lack of relevant evidence, but they provide the best basis for making judgements about the certainty of the evidence	NA
3.5 Decisions about treatments should not be based on considering only their benefits	0.615*+

*Covered in the IHC podcast (Semakula et al, 2017).

+Covered in U.S. podcast

Table 4
Participant Demographics

Characteristic	Study 1: Needs Assessment (n=142)	Study 2: User-Testing (n=5)	Study 2: Online RCT (n=250)
	Mean (SD, range) or % (n)		
Age	36.86 (7.93, 23-64)	36.40 (6.31, 30-43)	34.98 (7.8, 20-77)
Gender			
Female	45.1% (n=64)	100% (n=5)	49.0% (n=121)
Race/Ethnicity*			
African American	7.7% (n=11)	0	16.5% (n=41)
American Indian or Alaska Native	0.7% (n=1)	0	0.8% (n=2)
Asian	5.6% (n=8)	20% (n=1)	5.2% (n=13)
Hispanic	4.2% (n=6)	40% (n=2)	14.1% (n=35)
White	83.1% (n=118)	60% (n=3)	74.6% (n=185)
Highest Level of Education Achieved+			
Some high school, no diploma	0	0	0.4% (n=1)
High school	10.6% (n=15)	0	11.7% (n=29)
Some college, no degree	19.7% (n=28)	0	17.7% (n=44)
Associate's or technical degree	14.8% (n=21)	20% (n=1)	10.5% (n=26)
Bachelor's degree	43.0% (n=61)	40% (n=2)	49.6% (n=123)
Master's degree	12.0% (n=17)	40% (n=2)	9.3% (n=23)
Doctoral or other graduate degree	0	0	0.8% (n=2)
Employment Status+			
Currently working	82.4% (n=117)	100% (n=5)	85.9% (n=213)
Unemployed	4.2% (n=6)	0	3.6% (n=9)
Retired	0	0	0.8% (n=2)

	12.7% (n=18)	0	8.9% (n=22)
Homemaker			
Student or other	0.7% (n=1)	0	0.8% (n=2)
	Study 1: Needs Assessment (n=142)	Study 2: User-Testing (n=5)	Study 2: Online RCT (n=250)
Characteristic	Mean (SD, range) or % (n)		
Annual Household Income+			
Less than \$19,999	3.5% (n=5)	--	8.0% (n=20)
\$20,000 to \$39,999	27.5% (n=39)	--	22.4% (n=56)
\$40,000 to \$59,999	25.4% (n=36)	--	25.6% (n=64)
\$60,000 to \$79,999	18.4% (n=26)	--	22.0% (n=55)
\$80,000 to \$99,999	9.9% (n=14)	--	9.6% (n=24)
More than \$100,000	15.5% (n=22)	--	11.6% (n=29)
Insured	--	--	88.0% (n=220)
Children #	2.03 (1.31, 1-10)	2.20 (.84, 1-3)	1.73 (.94, 1-6)
Mental Health History and Service Use+			
Mental health diagnosis	--	--	43.6% (n=109)
Ever received psychotherapy for a psychological problem	--	--	45.6% (n=114)
Ever taken medication for a psychological problem	--	--	36.4% (n=91)
Podcast Listening Habits+			
< 30 minutes	--	--	28.4% (71)
30-60 minutes	--	--	33.6% (84)
61-90 minutes	--	--	16.8% (42)
> 90 minutes	--	--	20.8% (52)

*Indicates percentages do not sum to 100 due to overlap across categories

+Data missing (n=1-2)

Table 5*Descriptive Data for Study Variables by Condition and Overall*

Variable	Overall (n=250)	Critical Appraisal Podcast (n=128)	Control Podcast (n=122)
	Mean (SD, range) or % (n)		
Critical Appraisal – Continuous	14.36 (5.41, 2-21)	15.48 (5.19, 2-21)	13.19 (5.42, 2-21)
Critical Appraisal – Dichotomous			
High Score (1) – (18 or more correct out of 21)	39.6% (99)	52.3% (67)	26.2% (32)
Low Score (0) – (17 or less correct out of 21)	60.4% (151)	47.7% (61)	73.8% (90)
Intended Behavior (1-4)			
Find out what a claim was based on	3.53 (.71, 1-4)	3.54 (.73, 1-4)	3.52 (.70, 1-4)
Find out if a claim was based on a fair comparison study	3.61 (.63, 1-4)	3.72 (.55, 1-4)	3.49 (.69, 1-4)
How likely are you to participate in a fair comparison study	3.08 (.87, 1-4)	3.13 (.85, 1-4)	3.03 (.88, 1-4)
EBP Attitudes (1-5)			
Beliefs Regarding Therapists' Practices	24.74 (3.54, 10-30)	24.85 (3.57, 13-30)	24.61 (3.51, 10-30)
Attitudes about Mental Health Policy	13.40 (3.34, 5-20)	13.65 (3.19, 6-20)	13.14 (3.48, 5-20)
Negative Personal-Level Attitudes toward EBPs	23.97 (7.82, 9-45)	23.01 (7.43, 9-41)	24.98 (8.11, 9-45)
Negative Societal-Level Attitudes toward EBPs	14.48 (3.47, 5-20)	14.36 (3.51, 5-20)	14.60 (3.44, 4-20)
Attitudes Regarding Empiricism (1-5)	4.08 (.63, 2.2-5)	4.15 (.60, 2.4-5)	4.01 (.66, 2.2-5)
Vaccine Attitudes (1-5)	6.1 (3.58, 3-15)	6.03 (3.57, 3-15)	6.12 (3.60, 3-15)
Treatment Preferences			
Therapeutic alliance	24.66 (16.14, 0-99)	23.57 (16.51, 0-99)	25.80 (15.73, 0-99)
Scientific studies show therapy is highly effective	28.20 (18.82, 0-99)	30.02 (20.55, 0-99)	26.29 (16.70, 0-89)
Therapist experience	16.33 (10.63, 0-49)	16.02 (11.04, 0-49)	16.66 (10.21, 0-48)
Empathic therapist	18.39 (12.74, 0-71)	17.77 (12.53, 0-70)	19.05 (12.97, 0-71)
Client speaking majority of session	11.42 (10.04, 0-52)	11.63 (10.90, 0-52)	11.20 (9.09, 0-42)

Variable	Overall (n=250)	Critical Appraisal Podcast (n=128)	Control Podcast (n=122)
	Mean (SD, range) or % (n)		
Podcast Satisfaction			
Overall satisfaction (1-5)	4.18 (.89, 1-5)	4.16 (.93, 1-5)	4.20 (.85, 1-5)
Continue listening (1-5)	3.61 (1.16, 1-5)	3.56 (1.19, 1-5)	3.65 (1.14, 1-5)
Recommend to others (0-10)	6.71 (2.89, 0-10)	6.76 (2.79, 0-10)	6.66 (3.01, 0-10)
Relevance to mental health questions (1-4)	2.97 (.94, 1-4)	3.05 (.95, 1-4)	2.88 (.92, 1-4)
Relevance to physical health questions (1-4)	2.80 (1.03, 1-4)	3.14 (.89, 1-4)	2.43 (1.04, 1-4)

*Indicates percentages do not sum to 100 due to overlap across categories

Table 6

Linear Regression Analyses Examining Critical Appraisal Podcast Effect on Critical Appraisal, Intended Behavior, Attitudes, Treatment Preferences, and Overall Podcast Satisfaction (Control variables: age, education, and listening fidelity)

	Critical Appraisal		
	<i>B</i>	SE	β
Critical Appraisal	.450	.089	.266***
Intended Behavior			
Find out what a claim was based on	.042	.092	.029
Find out if a claim was based on a fair comparison	.252	.078	.201**
How likely are you to participate in a fair comparison study	.109	.116	.063
EBP Attitudes			
Beliefs Regarding Therapists' Practice	.073	.093	.050
Attitudes about Mental Health Policy	.495	.416	.074
Negative Personal-Level Attitudes toward EBPs	-1.96	.893	-.127*
Negative Societal-Level Attitudes toward EBPs	-.160	.446	-.023
Attitudes Regarding Empiricism	.156	.080	.124
Vaccine Attitudes	-.185	.400	-.026
Treatment Preferences			
Therapeutic alliance	-1.82	2.08	-.056
Scientific studies show therapy is highly effective	4.89	2.34	.131*
Therapist experience	-1.50	1.33	-.071
Empathic therapist	-159	1.64	-.063
Client speaking majority of session	.024	1.26	.001
Overall Podcast Satisfaction	-.050	.113	-.028

* $p < .05$; ** $p < .01$; *** $p < .001$

Table 7
Exploratory Analyses of the Interaction between Study Condition and Parent Demographic Variables

Demographic Variables	Critical Appraisal		
	<i>B</i>	SE	<i>p</i>
Age x Condition	.015	.013	.259
Education x Condition	.019	.084	.822
Income x Condition	.020	.038	.594
Medication for psychological problem x Condition	.227	.228	.298
Psychotherapy for a psychological problem x	.111	.211	.598

* $p < .05$; ** $p < .01$; *** $p < .001$