# JMIR Mental Health

Journal Impact Factor (JIF) (2023): 4.8
Volume 9 (2022), Issue 7    ISSN 2368-7959    Editor in Chief:  John Torous, MD, MBI

## Contents

### Review

### Original Papers

Review

# The Impact of Mobile Technology-Delivered Interventions on Youth Well-being: Systematic Review and 3-Level Meta-analysis

Colleen S Conley[1], PhD; Elizabeth B Raposa[2], PhD; Kate Bartolotta[2*], BA; Sarah E Broner[1*], BA; Maya Hareli[1], MA; Nicola Forbes[2], MA; Kirsten M Christensen[3], MA; Mark Assink[4], PhD

[1]Department of Psychology, Loyola University Chicago, Chicago, IL, United States

[2]Department of Psychology, Fordham University, Bronx, NY, United States

[3]Department of Psychology, University of Massachusetts Boston, Boston, MA, United States

[4]Faculty of Social and Behavioural Sciences, University of Amsterdam, Amsterdam, Netherlands

*these authors contributed equally

**Corresponding Author:**
Colleen S Conley, PhD
Department of Psychology
Loyola University Chicago
1032 W. Sheridan Road
Chicago, IL, 60660
United States
Phone: 1 7735083603
Email: cconley@luc.edu

## Abstract

**Background:** Rates of mental health problems among youth are high and rising, whereas treatment seeking in this population remains low. Technology-delivered interventions (TDIs) appear to be promising avenues for broadening the reach of evidence-based interventions for youth well-being. However, to date, meta-analytic reviews on youth samples have primarily been limited to computer and internet interventions, whereas meta-analytic evidence on mobile TDIs (mTDIs), largely comprising mobile apps for smartphones and tablets, have primarily focused on adult samples.

**Objective:** This study aimed to evaluate the effectiveness of mTDIs for a broad range of well-being outcomes in unselected, at-risk, and clinical samples of youth.

**Methods:** The systematic review used 5 major search strategies to identify 80 studies evaluating 83 wellness- and mental health-focused mTDIs for 19,748 youth (mean age 2.93-26.25 years). We conducted a 3-level meta-analysis on the full sample and a subsample of the 38 highest-quality studies.

**Results:** Analyses demonstrated significant benefits of mTDIs for youth both at posttest ($g$=0.27) and follow-up (range 1.21-43.14 weeks; $g$=0.26) for a variety of psychosocial outcomes, including general well-being and distress, symptoms of diverse psychological disorders, psychosocial strategies and skills, and health-related symptoms and behaviors. Effects were significantly moderated by the type of comparison group (strongest for no intervention, followed by inert placebo or information-only, and only marginal for clinical comparison) but only among the higher-quality studies. With respect to youth characteristics, neither gender nor pre-existing mental health risk level (not selected for risk, at-risk, or clinical) moderated effect sizes; however, effects increased with the age of youth in the higher-quality studies. In terms of intervention features, mTDIs in these research studies were effective regardless of whether they included various technological features (eg, tailoring, social elements, or gamification) or support features (eg, orientation, reminders, or coaching), although the use of mTDIs in a research context likely differs in important ways from their use when taken up through self-motivation, parent direction, peer suggestion, or clinician referral. Only mTDIs with a clear prescription for frequent use (ie, at least once per week) showed significant effects, although this effect was evident only in the higher-quality subsample. Moderation analyses did not detect statistically significant differences in effect sizes based on the prescribed duration of mTDI use (weeks or sessions), and reporting issues in primary studies limited the analysis of completed duration, thereby calling for improved methodology, assessment, and reporting to clarify true effects.

**Conclusions:** Overall, this study's findings demonstrate that youth can experience broad and durable benefits of mTDIs, delivered in a variety of ways, and suggest directions for future research and development of mTDIs for youth, particularly in more naturalistic and ecologically valid settings.

*(JMIR Ment Health 2022;9(7):e34254)* doi:10.2196/34254

**KEYWORDS**

meta-analysis; mental health; well-being; intervention; treatment; youth; technology; smartphone; mobile phone; app; mobile health

## Introduction

### Youth Mental Health Needs

Rates of mental health problems among youth, including children, adolescents, and young adults, are alarmingly high and appear to have risen in recent decades [1,2]. Rates of impulse control disorders (eg, attention deficit hyperactivity disorder and conduct disorder) and some anxiety disorders begin rising as early as the age of 4 years, with sharp increases in the prevalence of anxiety, mood, and substance use disorders across adolescence and young adulthood [3]. Indeed, nationally representative samples of adolescents and young adults show that >40% of youth in this age range experience psychological disorders in a given year and lifetime prevalence rates are estimated to approach 60% [4,5]. Beyond diagnosable mental disorders, many youths struggle with a diverse array of subclinical emotional, behavioral, interpersonal, and academic challenges [6-10].

Despite the high prevalence of youth facing mental health problems, only one-third to one-half of those in need receive mental health treatment [11-14], and treatment rates are even lower among low-income youth [15] and those with marginalized racial and ethnic identities [13,16,17]. This treatment gap is also evident among college students [11], which is notable given that these youth often have convenient access on campus to no- or low-cost mental health services [18,19].

Youth and families face several barriers to receiving mental health services [20-24]. Many lack knowledge and awareness of common mental health problems and may assume that certain behavioral or emotional issues are simply temporary phases or difficulties they can address on their own. These types of assumptions may be compounded by the stigma about mental illness and psychological services within the youth's family, culture, or broader community [25,26]. Many youths and caregivers also lack knowledge of available evidence-based treatments for mental health problems. Furthermore, several structural barriers severely limit the access of many youths and families to culturally sensitive, effective mental health care. Low-cost, evidence-based treatments for youth mental health problems are not available in many underserved communities across the world, including low-income rural and urban areas and countries with limited health care infrastructure. Even when such treatments are available, families may lack the time or resources needed to travel and take advantage of these treatments [27-30].

### Technology's Role in Youth Mental Health

Although efforts should continue to address the barriers that prevent formal services by qualified mental health professionals, it is also important to consider alternative ways of fulfilling the unmet mental health needs among youth. Mobile technology–delivered interventions (mTDIs), including mental health content delivered via mobile phones, tablets, and wearable smart devices (eg, watches, glasses, and virtual reality [VR] headsets), are potential ways of meeting this need. As of 2018, youth smartphone ownership and use were remarkably high, with 95% of teenagers having access to smartphones [31]. In 2016, the average age of first owning a smartphone was 10 years in the United States [32], and younger children commonly have access to smart devices through parents, siblings, or schools that provide tablets and other mobile devices to students. These mobile devices may be overlooked conduits for mental health information. Indeed, a recent survey of teenagers and young adults [33] reported that among those with moderate to severe depressive symptoms, 90% had searched the internet for information about mental health, and 38% used a mental health app. Parents also often use the internet for resources on health-related issues among their young children [34].

Technology can offer easy ways of connecting with mental health resources, such as mood-enhancing and skill-building apps purported to improve mental health. The ubiquitous, self-guided nature of such technology-delivered tools makes them appealing alternatives for those who are limited by access to, or trust in, formal mental health services [35]. Key themes in a recent review of research on internet-based help seeking for mental health difficulties among young people (aged up to 25 years) [36] showed that youth frequently engaged in technology-based (eg, internet-based) help seeking late at night (when traditional in-person mental health services are typically not available) and that youth endorsed several specific benefits of seeking help this way, including anonymity and privacy; lower perceived stigma and judgment; accessibility, including in times of crisis; and connection to others with similar experiences, which can foster a sense of community and acceptance.

### Technology-Delivered Interventions for Youth

#### *Potential and Pitfalls*

Although mTDIs have great potential to improve access to evidence-based mental health content, it is important to carefully evaluate their effects when it comes to mental health care, especially for youth. In contrast to computer- and internet-based technology-delivered interventions (TDIs), which are typically developed by clinicians and researchers to incorporate comprehensive and evidence-based treatment methods that parallel professional psychotherapy, *mobile* TDIs often lack such comprehensive, evidence-based principles while also introducing privacy and safety concerns [37]. The rapidly developing and competitive mobile app marketplace also poses some challenges in connecting evidence-based mental health practices with marketable and engaging mobile technology. Commercially available apps are typically designed by technology companies outside of the health care industry [38,39] with the aim of being engaging and attractive, thereby prioritizing appealing features such as design and gamification over evidence-based clinical techniques [40,41]. In contrast, research-developed apps prioritize evidence-based content and rigorous trials, which slows widespread availability as it can take 2 decades from conceptualization to public dissemination [37,42], leaving a gap between research-tested and commercially available mTDIs. Indeed, recent reviews of the content of >10,000 purported mental health apps in the commercial marketplace have noted that a vast majority have not gone through rigorous intervention development and testing [43] and are lacking or inconsistent with evidence-based psychotherapy principles [44], including apps specifically targeting youth mental health [40,45]. Although research trials have demonstrated promising findings for some mTDIs, the heterogeneity and poor quality of certain mTDIs and studies have led to inconclusive evidence across outcomes [46], which makes the role of these interventions in mental health services less clear [47]. Finally, mental health technologies generally have low rates of engagement and adherence [35,43,48-50]. Thus, a key question for this emerging area of research is how mTDIs can best be designed, prescribed, and implemented to harness their benefits for youth.

### *Areas for Further Research*

Previous reviews of TDIs, a broader category that goes beyond mobile interventions, have demonstrated the benefits of computer- and internet-based interventions, most commonly cognitive behavioral interventions, and most commonly examining the outcomes of depression, anxiety, and stress [51-58]. Similar to findings in adult populations, computer- and internet-based TDIs in youth samples have shown benefits, mostly in reducing internalizing symptoms, behavioral concerns, and eating disorders [51,52,54-60]. The literature on the efficacy of *mobile* TDIs, or mTDIs, is growing, with multiple meta-analytic reviews indicating positive impacts on a range of psychological outcomes in adults [42,53,61-63]. The emerging meta-analytic literature on mTDIs for *youth* is encouraging but limited, including generally beneficial results across (1) reviews blending a few trials of mTDIs together with mostly *nonmobile* TDIs in youth [52,64,65]; (2) a review combining 4 trials of mTDIs in children and adolescents with 21 trials of mTDIs in *adults* (mean age up to 59 years) [66]; (3) a recent meta-analysis of 12 youth trials, both with or *without* comparison groups, of smartphone apps exclusively on *internalizing* disorders [67]; and (4) another recent meta-analysis of 11 randomized controlled trials of smartphone apps for *depression*, *anxiety*, and *stress* in youth (aged 10-35 years) [68]. Although these initial findings are encouraging, a fitting next step for this emerging area of research is to meta-analytically review TDIs that are exclusively mobile in *youth* samples while including a *broad* array of youth clinical presentations and outcomes. Moreover, given the diverse designs of mTDIs for distinct youth characteristics and presenting problems, exploring the moderating influence of the interventions’ mobile technologies, theoretical orientations, technological and support features, and varying dosages would advance our ability to harness the full potential of mTDIs in improving youth well-being.

## The Current Meta-analysis: Goals and Hypotheses

The current 3-level meta-analysis evaluated the impact of wellness- and mental health-focused mTDIs (including smartphones and tablets, other types of mobile phones, and other handheld and wearable devices, including mobile VR) for youth, broadly defined as children, adolescents, and intentional (eg, university student) young adult samples or those with a mean age of ≤26 years. Improving upon some limitations of previous reviews, we included published and unpublished reports, only included controlled (either randomized or quasi-experimental) designs, and evaluated a broad range of participant clinical presentations (eg, unselected, at-risk, or clinical samples), intervention theoretical orientations, and outcomes. Drawing on evidence from prior reviews, we predicted that these mTDIs would yield significant benefits at postintervention on diverse indicators of youth well-being relative to comparison conditions. In addition, we examined the role of several potential moderators of intervention impact within the categories of methodological, youth, and intervention characteristics.

### *Methodological Characteristics*

#### Timing of Outcomes

Prior reviews note that there are a limited number of studies assessing the long-term effects of TDIs [51,54] and mTDIs [68] on youth. The reviews that compare the effects at postintervention versus later follow-up periods have been mixed, with some finding that effects are stable into follow-up periods (eg, parenting TDIs [59]) and others finding that some or all effects diminish over time (eg, adult mTDIs [69] and parenting TDIs [70]). Given that youth might have added challenges in implementing long-term gains [57], we tentatively predicted that the timing of the outcome assessment (posttest vs follow-up) would moderate the strength of the mTDIs’ effects such that the effects of mTDIs would wane over time.

#### Outcome Type

Prior reviews have established the benefits of mTDIs in reducing depression, anxiety, and stress, mostly in adults [42,62,69], with emerging evidence in youth [67,68]. In addition, mTDIs have been shown to be effective in improving life satisfaction, quality of life, and psychological well-being [69]. To broadly evaluate the potential impact of mTDIs on youth, we examined a broad range of youth outcomes, including those that have not yet been examined in prior reviews. Therefore, we expected mTDIs to have a beneficial impact on depression, anxiety, stress, and well-being, and explored whether mTDIs would also have beneficial effects on other outcome types, such as psychosocial strategies and skills, interpersonal relationship factors, academic functioning, health-related behaviors, or knowledge.

#### Comparison Group Type

Reviews of TDIs and mTDIs have generally found that effects are largest when they are compared with no-intervention or wait-list groups and smaller when compared with groups that are more active and clinically potent [42,52,58,60,62,69-71]. Thus, we predicted that the comparison group type would moderate the effects of mTDIs. Specifically, we expected that mTDIs would demonstrate the strongest benefits compared with no intervention (eg, wait-list), followed by inert interventions (including information-only and attentional or placebo controls), and demonstrate noninferiority compared with clinical comparisons, including usual clinical care and established clinical interventions.

## *Youth Characteristics*

### Age

Several previous reviews of TDIs in youth have demonstrated that older participants experience a greater reduction in symptoms than their younger counterparts [51,55,65,72]; however, others have found no effect of age [57,59], and preliminary evidence on a small sample of mTDIs in adolescents and young adults also failed to find an effect of age [68]. In an exploratory fashion, we examined whether age moderates the effects of mTDIs.

### Gender

The few reviews on the mental health benefits of smartphone apps that explored the role of participants' gender have revealed nonsignificant effects in adult [42] and youth [59] samples. Nevertheless, given the differences in the rates of various mental health problems as a function of gender across development [73,74], we tested the effects of gender in an exploratory fashion.

### Risk Level

Some prior reviews of mTDIs with adult samples have indicated that higher pretreatment severity (ie, clinical diagnosis or elevated mental health symptoms) is related to a greater reduction in symptoms and, therefore, produces a larger effect size (ES) [60,63]. In contrast, Pennant et al [54] found greater effects of computerized therapies for youth with subclinical symptoms versus a clinical diagnosis of anxiety; however, this effect was not found for depression. Two more recent reviews suggest inconclusive evidence regarding whether TDIs or mTDIs are more effective for youth who present with diagnoses or severe symptomatology [47,65]. Therefore, we examined participant risk level (ie, clinical diagnosis, elevated symptoms, and nonclinical sample) as an exploratory moderator.

## *Intervention Characteristics*

### Type of Technology

A prior review on the impact of mTDIs on both youth and adults did not find significant differences in effects by type of technology (ie, *smart* mobile phones/tablets vs other types of mobile phones, PDAs, wearable devices, or VR headsets) but noted that smartphone apps produced (nonsignificantly) larger ESs than both PDA and SMS text messaging interventions [66]. Given the limited information, we tested the impact of technology type as an exploratory moderator.

### Guiding Theoretical Framework

Cognitive behavioral–based and mindfulness- or acceptance-based interventions are the most commonly examined theoretical frameworks among TDIs for youth and adults [51,54-58] and mTDIs for adults [42,53,62,69] and have generally yielded positive effects for problems such as anxiety, depression, externalizing behaviors, and quality of life. Thus, we hypothesized that mTDIs with cognitive behavioral–based and mindfulness- or acceptance-based theoretical frameworks would produce significant effects and examined the impact of additional theoretical orientations (eg, motivational or positive psychology) in a more exploratory fashion.

### Technological and Support Features

Prior reviews have suggested that specific features of mTDIs, such as tailoring (ie, content shifting based on responses), gamification, and automatic reminders, may increase engagement and yield more robust effects [53,75]; however, much of this research has been conducted with adult samples. Therefore, we explored the potential moderating impact of various technological features such as personalization, tailoring, gamification, and social elements (eg, peer mentoring).

Research also points to the possibility that support from a human or virtual ("bot") professional, who can provide guidance, coaching, accountability, and, in some cases, supervised skills practice, may lead to increased adherence to mTDIs [47,76,77]. Indeed, several reviews have indicated that self-guided interventions are generally more effective with some access to human support or guidance, in part because of increased engagement and adherence [53,78,79]. Similarly, research on psychotherapy and other in-person youth interventions have highlighted the benefits of supervised practice in contributing to youths' psychological skill development, especially when delivered over multiple sessions [79-84]. Nevertheless, the overall evidence on human support for mTDI use is mixed, with a handful of reviews *not* finding added benefits for interventions incorporating human support as compared with those that do not [42,54,55,85]. Thus, we explored the potential impact of human or bot support elements, such as coaching, supportive accountability, and supervised skills practice.

### Dosage: Frequency and Duration

Previous studies have rarely investigated the dose-response relationship between TDI use and outcomes, and those that have done so tend to yield mixed results [47,51,54]. Some reviews have found that higher dosages, or longer durations, predicted a greater reduction in symptoms with particular outcomes (eg, problem behavior or depression) [59,86]. However, other studies were unable to establish such a relationship [57,85]. Therefore, we explored whether the prescribed or completed frequency (eg, weekly) or duration (eg, minutes, sessions, or weeks) of the intervention moderated the benefits of mTDIs.

# Methods

## Search Strategy and Study Selection

We used 5 systematic search strategies to assemble an unbiased, representative sample of published and unpublished controlled trials. First, we conducted searches for reports appearing through March 2021 in 5 academic databases: PsycINFO, ERIC, ProQuest Digital Dissertations, MEDLINE (Web of Science), and PubMed. We used a combination of several groups of search terms to find studies meeting our criteria for (1) participants (eg, *child**, *adolescen**, *teen*, *youth*, *young adult*, *university student*), (2) interventions (eg, *mental health*, *psychological*, *intervention*, *cognitive behavioral*, *mindfulness*), (3) mental health (eg, *depress**, *anxiety**, *well-being*), (4) technology (eg, *smartphone app**, *mobile app**, *tablet-based*), and (5) research design (eg, *RCT*, *controlled trial*, *clinical trial*, *quasi*, *comparison group*, *PRISMA*). Second, we also inspected the reference lists of each study meeting our criteria and of relevant previous reviews. Third, we hand-searched the contents of 16 selected journals most likely to publish studies on mobile mental health interventions involving youth. Fourth, we hand-searched the contents of proceedings for the recent years of 7 relevant academic conferences. Finally, we contacted authors of prior reviews, reports, and conference proceedings relevant to our sample to inquire about additional published or unpublished evaluations of fitting trials. Further details on these search strategies are provided in Multimedia Appendix 1.

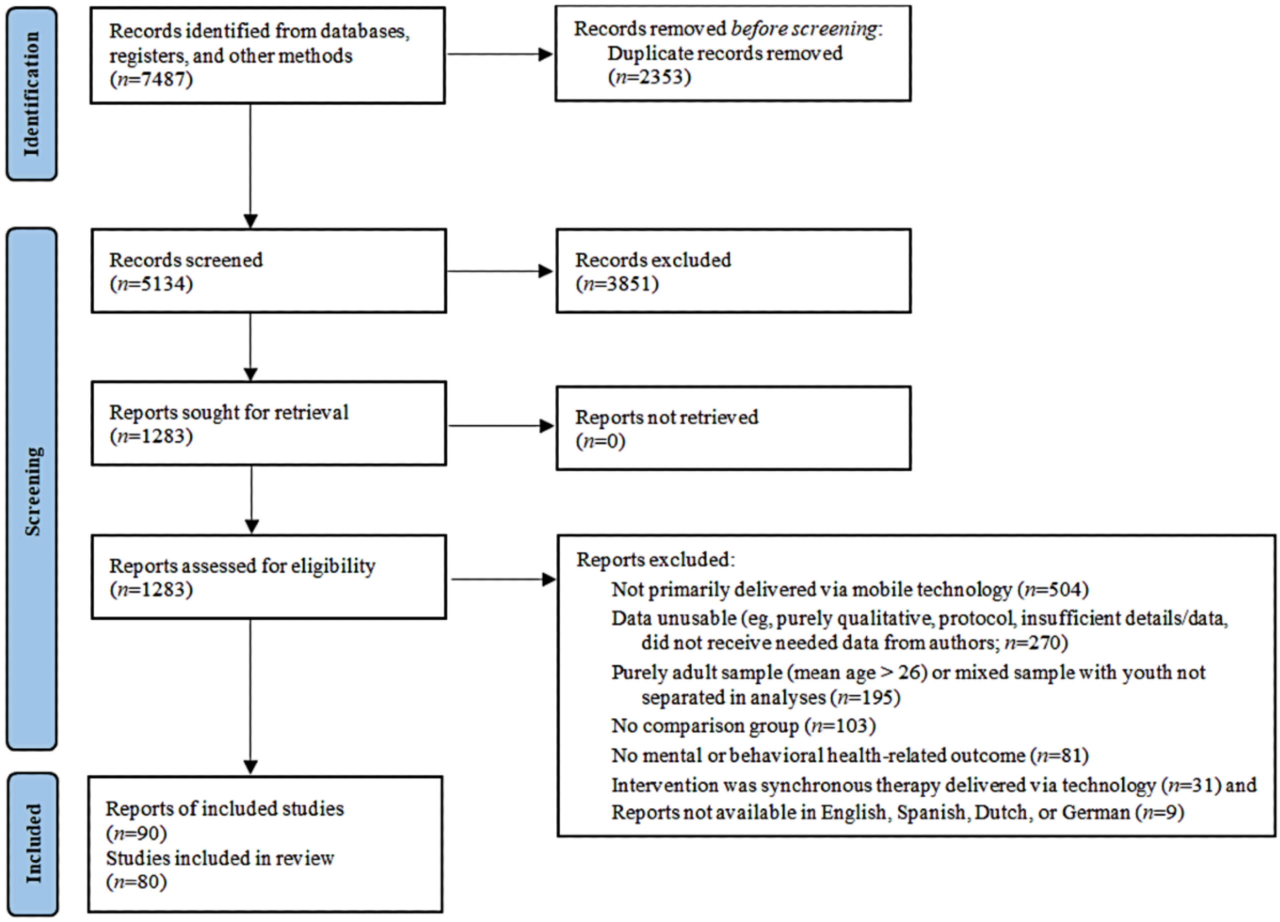
To be included in our final sample, the studies had to meet six criteria: (1) examine an *automated* psychological or behavioral intervention, either selecting participants based on a diagnosis or risk factor or targeting an unselected sample to promote mental health and wellness; (2) deliver the intervention primarily via *mobile* (handheld or wearable) technology, including pre–cellular technology handheld computers (eg, palm pilot and PDA), mobile cellular phones or tablets (eg, iPad and iPod touch)—using SMS text messaging, instant messaging, or more current mobile mental health apps—and wearable devices (eg, smart watch, smart glasses, VR headsets that are fully self-contained or linked with a mental health app on a smartphone or tablet device that is portable and able to be used in the participant's home); (3) contain at least one quantitatively assessed mental or behavioral health *outcome* measure (described in the following sections) for which ESs could be calculated; (4) target *youth*, broadly defined as children, adolescents, and intentional (eg, university student) young adult samples or those with a mean age ≤26 years (including interventions delivered to parents that targeted youth outcomes); (5) include a *comparison group* with at least 10 participants assigned to each condition; and (6) be reported in English, Spanish, Dutch, or German.

We excluded interventions with a primary focus on academics or physical health (eg, nutrition, weight loss, or diabetes management) but included studies that focused on psychobehavioral health such as smoking cessation, insomnia, and disordered eating. We did not include interventions delivered through audio or video tapes or videodiscs, a local computer program, or a website only. In addition, we did not include mobile interventions that were primarily reliant on human support (eg, therapists sending messages). Finally, we excluded interventions comprising solely medication reminders.

Figure 1 shows a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of sample searching, selection, and inclusion. The aforementioned search procedures identified 7487 potentially relevant reports, including 2353 (31.43%) duplicates that were removed. An additional 51.44% (3851/7487) of reports were eliminated as they did not meet our inclusion criteria. Among the 1283 eligible reports, some contained variants of the same intervention (eg, 2 interventions with the same active component but varying lengths), and we only included the intervention that was more comprehensive (ie, contained more elements or was longer in duration) or completely technology-based. However, if conceptually distinct interventions (eg, 2 different apps using different techniques) were evaluated in the same report, each intervention was coded separately. Data from multiple reports on the same sample and intervention were combined into a single report, reducing 16 overlapping reports to a sample size of 6.

In cases where means and SDs were not included in the original reports or effects could not be calculated because of insufficient data, we attempted to contact study authors to secure missing data. On the basis of a lack of author response, we excluded 16 studies for which no ESs could be calculated for any relevant outcome measure. This screening process led to a final sample of 83 interventions reported in 80 studies between 2005 and 2021.

**Figure 1.** PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of the study selection process.



## Study Coding (Data Extraction)

### Methodological Characteristics

For each report, we coded the year of the report, publication status, country in which the intervention took place, type of experimental design and comparison group, sample size, outcome types, and additional codes described in the following sections.

### Timing of Outcomes

We coded the number of weeks between pre- and postintervention outcome assessments and between postintervention and each follow-up assessment period.

### Outcome Type

We coded a broad array of outcomes to capture the various psychosocial and related aspects of functioning that might be affected by mTDIs. The relevant outcomes assessed in our sample of studies were classified into 14 possible categories, some of which were conceptually nested under higher-order categories, as noted in Textbox 1.

**Textbox 1.** Outcome types coded.

- General psychological well-being or distress included 2 subcategories:
  - ***Stress*** (eg, perceived and physiological indices of stress)
  - ***General or global psychological distress*** and well-being (eg, distress, positive and negative affect, mood states, quality of life, happiness, or life satisfaction)
- Psychosocial strategies or skills included 2 subcategories:
  - ***Social-cognitive strategies or styles*** (eg, different types of affective, cognitive, and social skills related to effective coping strategies, help-seeking behaviors, or mindfulness practices; overcoming dysfunctional beliefs, rumination, or hostility; resilience; or emotional self-awareness and regulation)
  - ***Self-perceptions*** (eg, self-esteem or self-efficacy)
- Internalizing symptoms included 2 subcategories:
  - ***Depression***
  - ***Anxiety***
- ***Other (noninternalizing) mental health problems*** (eg, autism, attention deficit hyperactivity disorder, or eating disorders)
- ***Health and health behavior*** (eg, substance use, sleep, physical activity, pain, or eating behaviors)
- ***Interpersonal relationships*** (eg, conflict, perceived social support, belongingness, loneliness, or social skills)
- ***Academics*** (eg, academic performance or adjustment)
- ***Psychology or health-related knowledge*** (eg, knowledge about topics such as substance use norms and consequences or sleep hygiene)
- ***Psychosocial outcomes in someone other than the target youth*** (eg, parent stress, warmth, use of punishment)
- ***Other*** (eg, perceptions of productivity, stigma, or close friend's smoking behavior)
- ***Intervention (ie, app) ratings*** (eg, intervention feasibility and social validity, acceptability of the mobile technology–delivered intervention (mTDI), and its uptake or use)

### Comparison Group Type

Studies were coded as having 1 of 3 different comparison groups. The majority of studies included a *no-intervention* (eg, wait-list control) condition in which the comparison group only completed assessment procedures. Some studies compared the intervention of interest with an *inert* comparison group, whether information-only (eg, pamphlets or website links to general health-related information), attention-placebo (eg, passive SMS text messages), or minimal treatment-as-usual (eg, standard protocol before a medical procedure) conditions that did not contain the therapeutic elements of the evaluated intervention. These comparison groups generally attempted to control for nonspecific factors such as attention or social interaction. Finally, some studies included a *clinical* comparison group, whether a usual clinical care comparison or some other established (validated or otherwise intended to be beneficial) intervention. In some studies with a clinical comparison group, both the mTDI and comparison groups received a similar base intervention (eg, counseling vs counseling+app) and thus tested the *added* or incremental benefit of the mTDI of interest.

## *Youth Characteristics*

### Age, Gender, Race, and Ethnicity

When the information was available, studies were coded for the sample age (mean, SD, and range), gender, race, and ethnicity.

### Risk Level and Type

We coded whether researchers selected participants based on particular symptoms or risk factors into the following categories: (1) *psychological clinical sample* (ie, symptoms indicative of a Diagnostic and Statistical Manual of Mental Disorders diagnosis) [87,88], (2) *psychological or mental health at-risk sample* (ie, subclinical symptoms of psychopathology), (3) *nonmental health risk* (ie, medical risk, diagnosis, or procedure), and (4) *general (unselected) community sample* not selected for any particular risk factor.

## *Intervention Characteristics*

### Type of Technology

We coded each intervention's primary and secondary (if relevant) type of technology into one of the following categories: (1) *smartphone or tablet* (eg, iPhone, iPad, or iPod touch), (2) *presmartphone mobile device* (eg, presmart mobile phone, palm pilot, or PDA), (3) *mobile VR* (eg, headsets) or *video game* (ie, handheld), and (4) *other wearable devices* (eg, smart watch, biosensor or activity monitor, and smart glasses). VR headsets and other wearable devices were typically used in conjunction with smartphones or tablets. Finally, some interventions were also able to be accessed on a (5) *computer* as a *secondary* type of technology.

### Guiding Theoretical Framework

Interventions were coded as having one of the following primary guiding theoretical frameworks: (1) *cognitive behavioral*, (2) *mindfulness- or acceptance-based* (eg, mindfulness-based stress reduction, mindfulness-based cognitive therapy, acceptance and commitment therapy, or dialectical behavior therapy), (3) *blended cognitive behavioral and mindfulness*, (4) *other or multiple* (ie, positive psychology, interpersonal, motivational or stages of change, and transtheoretical), or (5) *atheoretical or not specified*. When mTDIs were available in the commercial market, they were consulted directly to supplement the information obtained from the research reports.

### Technological and Support Features

We coded whether the intervention included *personalization* (ie, the ability to alter the app environment through features such as personal preferences; personal dashboards; or use of photos, music, or contacts), *tailoring* (ie, the use of algorithms that alter intervention content based on contact sensing, prior responses, feedback, or other input), a *social component* (eg, forum or social media use or mentoring), or *gamification* (eg, rewards, badges, points, levels, or quests).

We also coded several intervention features designed to support participants in using the mTDI: (1) *training or orientation* for the participants about using the mTDI (eg, virtual training within the app or via email or video chat, in-person training, or a paper manual); (2) *in-person element* besides orientation or training (eg, simultaneous counseling); (3) *reminders* sent to encourage the use of the mTDI, either automatically through the app (eg, push or banner notifications) or outside of the app (eg, emails, texts, or calendar reminders); (4) *human or bot (automatic) support* (eg, supportive SMS text messages, phone calls, or personalized feedback) specifically around the mTDI; (5) targeted guidance indicative of *supportive accountability*, designed to increase adherence to an intervention via support and accountability from a trustworthy coach who assists with setting process-oriented expectations and goals [77]; and (6) targeted guidance in the form of *supervised skills practice* [79].

### Dosage: Frequency and Duration

When the information was available, studies were coded for the intervention's prescribed and completed (both objectively determined and self-reported) frequency and duration. Specifically, the frequency of use was coded as one of the following categories: at least 4 days per week or as much as feasible, 2 to 3 days per week, once per week, less than once per week, one-time session, and not stated or at user discretion. The intervention duration was coded in terms of minutes, weeks, and sessions. When data were available, we also calculated the percentage of the completed duration of the intervention by dividing the completed duration by the prescribed duration.

## Risk of Bias (Quality) Assessment

For study quality, we followed the approach of an integrative study quality coding scheme [89] designed to draw upon the strengths of several previously validated quality indices, including the Cochrane Collaboration's tool for assessing the risk of bias [90-92]. This coding scheme rates each study on 10 features: peer review and impact factor, experimental design, sample size, attrition, reliability of measures, validity of measures, adjustment for pretest differences, intent-to-treat analysis, reporting of sample characteristics, and involvement of study authors in mTDI development. Each feature is rated on a 4-point scale (from 0, indicating the lowest quality, to 3, indicating the highest quality). The 10 item scores are then summed, resulting in a score for which a score of 20 represents average or normal research practices.

## Reliability of Coding

A team of 5 trained postbaccalaureate and graduate students assessed the studies for eligibility and inclusion in the meta-analysis and met weekly to review any questions for consensus. A team of 6 graduate students with advanced clinical and quantitative training then reviewed and coded eligible reports for descriptive features, moderators, quality indicators, and outcome data. The coders were supervised by 3 faculty members with expertise in clinical psychology, mTDIs, and meta-analytic procedures. After the iterative training phase, coders had ongoing opportunities for consensus checks through a consultation system and weekly faculty supervision. From a subsample of 18 to 31 studies (depending on the code) containing 44 interventions, 46 comparisons, and 108 ESs, any code that did not reach adequate reliability (ie, >0.80 κ, 85% agreement, or 0.95 intraclass correlation coefficient, as fitting) [93,94] was reviewed by at least one other coder in the entire sample. Lead authors provided an additional review of randomly selected articles throughout the coding process. Any questions or discrepancies were resolved through discussions.

## Meta-analytic Strategy

### *ES Calculation*

Cohen *d* was calculated for each outcome to reflect the effect of mTDIs relative to the comparison condition, with positive ESs representing outcomes in which the intervention group outperformed the comparison group. If *d* values could not be obtained directly from primary studies, the formulas by Borenstein et al [95] and Lipsey and Wilson [96] were used to transform the reported statistical information into Cohen *d*. Whenever possible, *d* values were calculated using means and SDs, frequencies or proportions, odds ratios, or results from *F* or *t* tests. If a primary study did not report sufficient information to extract or calculate the ES, the study authors were contacted for additional information. When the only information available indicated that an ES was nonsignificant, we conservatively set that ES to zero, following Mullen [97]. This procedure was preferred above excluding primary studies from the review, as the latter would reduce the statistical power in the analyses. To correct for pretreatment differences, we adjusted the postintervention and follow-up effects for preintervention baseline outcome levels (using subtraction, similar to procedures in other meta-analyses) [89,98,99] when pretreatment data were available. Finally, before analysis, all ESs were converted to Hedges *g* to account for potential bias in small sample sizes.

### *The 3-Level Meta-analytic Model*

Most primary studies included in this review reported on multiple intervention effects, typically because multiple outcomes were tested or multiple comparison conditions were part of the study design. The resulting dependency in ESs (ie, the fact that ESs extracted from the same study are more alike than the ESs extracted from different studies) violates the assumption of independent ESs underlying traditional meta-analytic techniques [96].

Therefore, a 3-level random-effects model was used for all analyses [100-104]. In this 3-level model, 3 sources of variance were modeled: sampling variance of the observed ESs (ie, sampling variance; level 1), variance between ESs derived from the same study (ie, within-study variance; level 2), and variance in ESs derived from different studies (ie, between-study variance; level 3). The sampling variance at level 1 of the model is not estimated but considered known and calculated using the formula given by Cheung [101].

To determine whether testing select moderators would be informative, we first examined the ES heterogeneity by testing the significance of the within-study variance (level 2) and the between-study variance (level 3). We performed 2 one-sided log-likelihood ratio tests in which the deviance of the full model was compared with the deviance of the model without one of these variance parameters. If the within-study variance or the between-study variance were significant, we proceeded with the moderator analyses. The coded variables were only tested as moderators when (categories of) these variables were based on at least three studies or three ESs. In some cases, we consolidated categories with <3 studies or ESs into another (or the *other*) category.

### *Software and Parameters*

We used the function *rma.mv* of the *metafor* package [105] in the R statistical environment (version 3.6.1; R Foundation for Statistical Computing) [106], following the setup and R syntax by Assink and Wibbelink [100], to model the 3 sources of ES variance [103,104]. The overall effect was estimated using an intercept-only model, and potential moderators were examined by adding these variables as covariates to the intercept-only model. The $t$ distribution was used in testing individual regression coefficients of the models and for calculating the corresponding CIs [107]. When models were extended with categorical moderators comprising >3 categories, the omnibus test followed an $F$ distribution. The restricted maximum likelihood estimation method was used to estimate the model parameters. Before conducting moderator analyses, continuous variables were centered on their means, and dichotomous dummy variables were created for categorical variables. The log-likelihood ratio tests were conducted as 1-tailed, whereas all other significance tests were conducted as 2-tailed. The significance level was set to 0.05 in all analyses, and 95% CIs were estimated.

### *Publication Bias*

A problem that may arise in meta-analysis is the *file drawer problem* [108], in that studies with nonsignificant or negative results are less likely to be published than studies that produced significant and positive results. To reduce this problem, we attempted to be exhaustive in our search strategy to retrieve both published and unpublished primary studies. To further assess bias in our data set of ESs, 2 analyses were conducted. First, we performed the trim-and-fill analysis by Duval and Tweedie [109,110] to examine the symmetry of a funnel plot in which ESs were plotted against their SEs. In the case of publication bias, the plot is asymmetrical, as the ESs are missing to the left of the estimated mean. The trim-and-fill algorithm estimates these *missing* ESs using an iterative nonparametric method. After imputing these ESs, the symmetry of the plot is restored, and an *adjusted* overall effect can be estimated. We also examined bias by performing the Egger test, in which ESs are regressed on their SEs [111]. This was performed by adding the SE as a covariate to an intercept-only, 3-level meta-analytic model. In this model, a significant positive slope indicated the presence of publication bias.

## *Results*

### Study Sample and Descriptive Characteristics

Multimedia Appendix 2 [112-200] provides a table with details about each of the 80 studies eligible for this meta-analysis, 3 (4%) of which contained 2 eligible interventions and 10 (13%) that contained 2 eligible comparison groups, yielding 83 interventions, 93 comparisons, and a combined sample size of 19,748 youth. Of these 80 studies, 76 (95%) provided estimates of 484 postintervention ESs, and 29 (36%) studies provided estimates of 225 follow-up ESs.

Several aspects of the 80 included studies are worthy of comment. First, most (68/80, 85%) of the studies were published in peer-reviewed scientific journals, and the remainder were unpublished dissertations or in-preparation manuscripts that extended prior peer-reviewed work published as a pilot trial or presented at an academic conference. In addition, most of the studies were published within the past decade or so, with 96% (77/80) published since 2010 and 28% (22/80) since 2020. Of the 80 studies, 33 (41%) were conducted in the United States, with 36 (45%) reports from the broader North American continent, 23 (29%) from Europe, 11 (14%) from Australia, 9 (11%) from Asia or the Middle East, and 1 (1%) from South America.

In terms of participants, across the 93% (74/80) of studies reporting relevant demographic information (and among the 67/80, 84% of studies reporting SD), the average age ranged from 2.93 to 26.25 (weighted mean 15.92; SD 2.86) years, and on average, 63.83% of study samples were female (but notably, most studies did not report on, or likely assess, gender other than female or male). Only 38% (30/80) of studies provided a full breakdown of participant race and ethnicity, and 23% (18/80) provided no information on these demographics at all. Furthermore, 63% (50/80) of studies selected participants based on one or more risk factors versus recruiting a general community sample. The most common risk factor used for participant recruitment and screening was subclinical psychological risk (eg, substance use or elevated depression; 30/80, 38% of studies), followed by some nonmental health risk (12/80, 15% of studies; in all cases within this sample, this was a medical diagnosis such as spina bifida or obesity or a medical procedure such as surgery or dental work), and, finally, participants with a clinical psychological or psychiatric diagnosis (eg, anxiety or autism; 8/80, 10% of studies).

In terms of the 83 interventions, 74 (89%) used smartphones or tablets (1 used an iPod touch); 4 (5%) used presmartphone mobile devices (all phones, including Motorola A925, Sony Ericsson, and Vodafone); 4 (5%) used VR headsets, either freestanding or in conjunction with a mobile phone app; and 1 (1%) used a handheld video game. Most (70/83, 84%) of the interventions took place in participants' daily environments; however, several (13/83, 16%) took place in a medical setting (eg, to address anxiety or pain related to a medical procedure). The most prevalent guiding theoretical framework of the mTDIs was cognitive behavioral (36/83, 43% of the interventions), followed by other or transtheoretical frameworks (eg, positive psychology and motivational; 20/83, 24%), mindfulness- or acceptance-based (17/83, 20%), and a few atheoretical or unspecified frameworks (3/83, 4%). Furthermore, in nonexclusive categories, the interventions' technological features included personalization (18/83, 22%), tailoring (36/83, 43%), a social component (10/83, 12%), and gamification (20/83, 24%). In terms of support features, of the 83 interventions, 30 (36%) included some sort of orientation or training (either virtual or in person), 12 (14%) contained one or more other in-person element, 40 (48%) incorporated reminders to encourage the use of the intervention, and 22 (27%) included some form of human or bot support or guidance, with 20 (24%) containing supportive accountability and 9 (11%) containing supervised skills practice.

All (80/80, 100%) studies provided some information about the prescribed or completed dosage (or both) of their interventions, whether objectively pulled from the mTDI or self-reported by the participants; however, the specific details reported were variable. Of the 83 interventions, 13 (16%) were single-session interventions and the remainder were prescribed to range from 4 to 2505 sessions (weighted mean 89.86, SD 374.29; *k*=42 studies reporting on 44 interventions) across a time span of 2 days to 43.45 weeks (weighted mean 7.48, SD 7.46; *k*=65 studies reporting on 68 interventions). Of these 70 interventions (contained in 67 studies), 43 (61%) were prescribed for daily use, 9 (13%) for 2 to 3 days per week, 7 (10%) for once a week, and the remaining 11 (16%) were either prescribed to be used as needed or at the user's discretion or not stated in the report. In terms of duration of use, the prescribed minutes of use for interventions ranged from 5 to 3650 minutes (mean 345.25, SD 789.95; *k*=32 studies reporting on 32 interventions). Notably, only 47 out of 80 (59%) studies provided some sort of objective information about how much participants *actually* engaged in the intervention (eg, number of sessions, minutes, or weeks). Using all available information, we calculated the intervention *completion* percentage and found the average to be 85.05% of the researchers' prescribed sessions (*k*=29 studies reporting on 30 interventions), 87.19% of the prescribed intervention minutes (*k*=13 studies reporting on 13 interventions), and 86.7% of the prescribed intervention weeks (*k*=37 studies reporting on 43 interventions).

Studies assessed a variety of psychosocial outcomes, which we originally coded in 14 categories (see the *Methods* section) and then consolidated into 6 categories because of conceptually similar content or small numbers of studies or effects (see the final list of consolidated categories in the note below Multimedia Appendix 2).

Notably, 25% (20/80) of studies also provided information about the intervention group's ratings on measures of the intervention's social validity (eg, user satisfaction, perceived usefulness, quality, usability, or acceptability). As these data were generally only available at the posttest time points and for intervention but not comparison groups, we do not report ESs on these types of outcomes. However, to analyze trends in diverse measures of mTDI social validity across all studies with such data available, we standardized all available Likert scale ratings for these constructs onto a single scale, with 0 representing the lowest and 100 the highest possible rating of social validity. On this standardized scale, the average rating (weighted by included sample size) for self-report scores of the mTDIs' social validity was 58.24 and ranged quite widely across studies (30.20-100).

There was variability in the types of comparison groups as well. Slightly less than half (37/80, 46%) of the comparisons involved groups such as wait-lists that contained no active intervention, whereas the remainder of the comparisons involved either passive information-only or placebo groups (28/80, 35%) or, less commonly, clinical comparisons that were intended to have therapeutic benefits (15/80, 19%). For the studies (43/80, 54%) that used an active (inert or clinical) comparison, the modality of the comparison group was distributed fairly evenly across in-person (20/43, 47%) and other technology-based interventions (18/43, 42%), with just a few (5/43, 12%) having some other modality (ie, blended interventions containing both technology and in-person elements or paper-and-pencil materials).

## Average Effect of mTDIs

The average ES across all possible comparisons within the 80 studies (yielding 709 ESs across posttest and follow-up assessments) was $g=0.27$ ($P<.001$; 95% CI 0.20-0.33). There was significant heterogeneity across studies ($\sigma^2$ level 3=0.06; $P<.001$; 51.76% of the variance among ESs), as well as between ESs extracted from the same study ($\sigma^2$ level 2=0.03; $P<.001$; 27.50% of the variance among ESs). Random sampling error accounted for 20.74% of the variance. To explore the substantial variability between and within studies, a number of moderators were considered. These analyses are described in the following 3 sections and detailed in Table 1.

**Table 1.** Moderators of the effectiveness of mobile technology–delivered interventions for youth[a].

| Characteristics | $k$[b] | Effect sizes, n | $B_0$ (intercept), *g* (95% CI) | $B_1$ (slope), *g* (95% CI) | *F* ($df_1$, $df_2$) | *P* value |
|---|---|---|---|---|---|---|
| **Methodological characteristics** | | | | | | |
| Study quality | 80 | 709 | 0.25 (0.19 to 0.31)$^{***}$ | –0.02 (–0.04 to –0.01)$^{**}$ | 7.03 (1, 707) | .01[c] |
| **Timing of outcome** | | | | | 0.07 (1, 707) | .79 |
| Posttest (RC[d]) | 76 | 484 | 0.27 (0.21 to 0.33)$^{***}$ | N/A[e] | | |
| Follow-up | 29 | 225 | 0.26 (0.19 to 0.34)$^{***}$ | –0.01 (–0.06 to 0.05) | | |
| **Outcome type** | | | | | 2.70 (5, 703) | .02 |
| General psychological well-being or distress (RC) | 35 | 98 | 0.28 (0.20 to 0.37)$^{***}$ | N/A | | |
| Internalizing (depression, anxiety) | 44 | 145 | 0.30 (0.22 to 0.39)$^{***}$ | 0.02 (–0.06 to 0.10) | | |
| Other (noninternalizing) mental health | 7 | 42 | 0.21 (0.04 to 0.38)$^{*}$ | –0.07 (–0.25 to 0.10) | | |
| Psychosocial strategies and skills | 26 | 161 | 0.34 (0.25 to 0.42)$^{***}$ | 0.05 (–0.02 to 0.13) | | |
| Health (behavior; eg, substance use) | 35 | 190 | 0.24 (0.16 to 0.32)$^{***}$ | –0.04 (–0.15 to 0.06) | | |
| Other (eg, knowledge or relationships) | 20 | 73 | 0.15 (0.04 to 0.25)$^{**}$ | –0.14 (–0.25 to –0.03)$^{**}$ | | |
| **Comparison group type** | | | | | 0.17 (2, 706) | .84[c] |
| No intervention (RC; eg, wait-list) | 41 | 376 | 0.28 (0.20 to 0.36)$^{***}$ | N/A | | |
| Inert (eg, placebo or information-only) | 30 | 186 | 0.26 (0.18 to 0.35)$^{***}$ | –0.02 (–0.11 to 0.08) | | |
| Clinical (eg, established intervention) | 18 | 147 | 0.24 (0.12 to 0.36)$^{***}$ | –0.04 (–0.17 to 0.09) | | |
| Mean age (years) | 74 | 669 | 0.26 (0.19 to 0.32)$^{***}$ | 0.003 (–0.01 to 0.01) | 0.17 (1, 667) | .68[c] |
| Gender (percentage female) | 77 | 699 | 0.26 (0.20 to 0.33)$^{***}$ | 0.002 (–0.002 to 0.01) | 1.06 (1, 697) | .30 |
| **Risk level and type** | | | | | 3.15 (3, 705) | .02[c] |
| General sample not selected for risk (RC) | 30 | 247 | 0.19 (0.09 to 0.29)$^{***}$ | N/A | | |
| Nonmental health (ie, medical) risks | 12 | 46 | 0.52 (0.33 to 0.72)$^{***}$ | 0.33 (0.11 to 0.55)$^{**}$ | | |
| Psychological or mental health at-risk sample | 30 | 332 | 0.29 (0.19 to 0.39)$^{***}$ | 0.09 (–0.05 to 0.23) | | |
| Psychological clinical sample (diagnosis) | 8 | 84 | 0.20 (0.01 to 0.39)$^{*}$ | 0.01 (–0.21 to 0.22) | | |
| **Intervention characteristics** | | | | | | |
| **Primary type of technology** | | | | | 0.82 (2, 706) | .44 |
| Smartphone or tablet (RC) | 71 | 665 | 0.25 (0.19 to 0.32)$^{***}$ | N/A | | |
| Presmartphone mobile device | 4 | 23 | 0.37 (0.09 to 0.65)$^{*}$ | 0.11 (–0.18 to 0.40) | | |
| Mobile VR[f] headset or handheld video game | 5 | 21 | 0.41 (0.13 to 0.68)$^{**}$ | 0.15 (–0.13 to 0.44) | | |
| **Guiding theoretical framework** | | | | | 1.04 (4, 704) | .38 |
| Cognitive or behavioral (RC) | 35 | 249 | 0.31 (0.21 to 0.40)$^{***}$ | N/A | | |
| Mindfulness or acceptance | 16 | 238 | 0.28 (0.15 to 0.42)$^{***}$ | –0.02 (–0.19 to 0.15) | | |

| Characteristics | *k*[b] | Effect sizes, n | $B_0$ (intercept), *g* (95% CI) | $B_1$ (slope), *g* (95% CI) | *F* ($df_1$, $df_2$) | *P* value |
|---|---|---|---|---|---|---|
| Cognitive behavioral and mindfulness | 6 | 46 | 0.35 (0.12 to 0.58)$^{**}$ | 0.05 (–0.02 to 0.29) | | |
| Other or multiple (eg, motivational) | 20 | 171 | 0.16 (0.04 to 0.29)$^{**}$ | –0.14 (–0.30 to 0.01)[g] | | |
| Atheoretical or not specified | 3 | 5 | 0.41 (–0.04 to 0.85)[g] | 0.10 (–0.35 to 0.56) | | |
| **Intervention technological features** | | | | | | |
| **Personalization** | | | | | 1.37 (1, 706) | .24 |
| Absent (RC) | 63 | 575 | 0.28 (0.21 to 0.35)$^{***}$ | N/A | | |
| Present | 17 | 133 | 0.19 (0.07 to 0.32)$^{**}$ | –0.09 (–0.23 to 0.06) | | |
| **Tailoring** | | | | | 0.84 (1, 706) | .36 |
| Absent (RC) | 45 | 399 | 0.28 (0.20 to 0.37)$^{***}$ | N/A | | |
| Present | 34 | 309 | 0.23 (0.13 to 0.32)$^{***}$ | –0.06 (–0.18 to 0.07) | | |
| **Social component** | | | | | 0.17 (1, 706) | .68 |
| Absent (RC) | 70 | 628 | 0.26 (0.20 to 0.33)$^{***}$ | N/A | | |
| Present | 9 | 80 | 0.22 (0.05 to 0.40)$^{*}$ | –0.04 (–0.23 to 0.15) | | |
| **Gamification** | | | | | 0.78 (1, 706) | .38 |
| Absent (RC) | 60 | 531 | 0.24 (0.17 to 0.31)$^{***}$ | N/A | | |
| Present | 19 | 177 | 0.31 (0.18 to 0.44)$^{***}$ | 0.07 (–0.08 to 0.21) | | |
| **Intervention support features** | | | | | | |
| **Orientation to or training on mTDI**[h] | | | | | 0.09 (1, 678) | .77 |
| Absent (RC) | 48 | 361 | 0.27 (0.19 to 0.35)$^{***}$ | N/A | | |
| Present | 29 | 319 | 0.25 (0.15 to 0.35)$^{***}$ | –0.02 (–0.15 to 0.11) | | |
| **Other in-person element** | | | | | 0.004 (1, 707) | .95 |
| Absent (RC) | 68 | 637 | 0.27 (0.20 to 0.33)$^{***}$ | N/A | | |
| Present | 12 | 72 | 0.27 (0.11 to 0.43)$^{***}$ | 0.006 (–0.16 to 0.17) | | |
| **Reminders** | | | | | 1.91 (1, 696) | .17 |
| Absent (RC) | 40 | 233 | 0.31 (0.21 to 0.40)$^{***}$ | N/A | | |
| Present | 37 | 465 | 0.22 (0.13 to 0.30)$^{***}$ | –0.09 (–0.22 to 0.04) | | |
| **Guidance, coaching, and feedback** | | | | | 0.24 (1, 707) | .62 |
| Absent (RC) | 59 | 531 | 0.28 (0.20 to 0.35)$^{***}$ | N/A | | |
| Present | 21 | 178 | 0.24 (0.12 to 0.36)$^{***}$ | –0.04 (–0.18 to 0.11) | | |
| **Supportive accountability** | | | | | 0.14 (1, 707) | .71 |
| Absent (RC) | 62 | 538 | 0.27 (0.20 to 0.35)$^{***}$ | N/A | | |
| Present | 18 | 171 | 0.25 (0.12 to 0.38)$^{***}$ | –0.03 (–0.18 to 0.12) | | |
| **Supervised skills practice** | | | | | 0.62 (1, 707) | .43 |
| Absent (RC) | 72 | 651 | 0.26 (0.19 to 0.33)$^{***}$ | N/A | | |
| Present | 8 | 58 | 0.34 (0.14 to 0.54)$^{***}$ | 0.08 (–0.13 to 0.30) | | |
| **Dosage: prescribed frequency of use** | | | | | 4.39 (4, 704) | .002[c] |

| Characteristics | *k*[b] | Effect sizes, n | $B_0$ (intercept), *g* (95% CI) | $B_1$ (slope), *g* (95% CI) | *F* ($df_1$, $df_2$) | *P* value |
|---|---|---|---|---|---|---|
| As much as feasible; ≥4 days per week (RC) | 43 | 507 | 0.23 (0.15 to 0.30)*** | N/A | | |
| Some days or more than once a week | 9 | 53 | 0.27 (0.09 to 0.44)*** | 0.04 (–0.15 to 0.23) | | |
| About once a week | 7 | 58 | 0.53 (0.33 to 0.73)*** | 0.30 (0.09 to 0.52)** | | |
| One-time session | 13 | 45 | 0.46 (0.28 to 0.63)*** | 0.23 (0.04 to 0.42)* | | |
| Not stated, when needed, or at user discretion | 8 | 46 | 0.08 (–0.09 to 0.24) | –0.15 (–0.33 to 0.03) | | |
| **Dosage: prescribed duration of intervention** | | | | | | |
| Weeks | 78 | 697 | 0.27 (0.21 to 0.34)*** | –0.003 (–0.01 to 0.01) | 0.35 (1, 695) | .56 |
| Sessions | 54 | 414 | 0.33 (0.24 to 0.41)*** | –0.0001 (–0.0003 to 0.0002) | 0.44 (1, 412) | .51 |

[a] The right columns list the omnibus *F* test and *P* value for each moderation test. The middle columns list the intercept ($B_0$), or mean effect size, and slope ($B_1$), an estimated unstandardized regression coefficient, of the relevant Hedges *g* statistics, with CIs around each. Effects and slopes that differ significantly from 0 are denoted with asterisks in the intercept ($B_0$) and slope ($B_1$) columns, respectively. For categorical moderators, each intercept represents the mean effect of a category, whereas each slope represents the difference in the mean effect between the category and reference category. Depending on its sign, the slope of a continuous moderator represents an increase or decrease in the effect size with each unit increase in the variable.

[b] Number of studies with relevant effect size data for a given row. In cases of multiple interventions or comparisons, some studies were counted in multiple rows; thus, these numbers sometimes exceeded 80. Owing to missing data, some counts fall short of 80. Further details on what was included in different categories of the included moderators are provided in the *Methods* section.

[c] Significance of moderation analysis changed when conducted on a subsample of the highest-quality studies (*k*=38; Table 2).

[d] RC: reference category.

[e] N/A: not applicable (as the slope represents a comparison with the reference category).

[f] VR: virtual reality.

[g] *P*<.10.

[h] mTDI: mobile technology–delivered intervention.

**P*<.05.

***P*<.01.

****P*<.001.

## Differences in Effects of mTDIs Based on Methodological Characteristics

### *Study Quality and Publication Bias*

Overall study quality significantly moderated the overall effect of mTDIs in such a way that ESs decreased as the quality index increased (Table 1). The slope indicated that for every 1-point increase in quality index score, the effect decreased by 0.02. Given this moderation effect, we also ran all analyses with only the higher-quality studies—that is, studies that achieved a total quality index >20, which denotes studies that, on average, surpassed benchmarks for average-quality research methods [89]. Unless otherwise noted in the relevant presentation of results in the following sections, the pattern and significance of the results with this reduced, higher-quality sample of studies were identical to those of the full sample of studies. However, in cases where the statistical significance of results shifted when tested with only higher-quality studies, the results from analyses with only the higher-quality studies are presented separately in Table 2 (full set of results available from authors upon request). The average ES across all possible comparisons within the 38 higher-quality studies (yielding 428 ESs across posttest and follow-up assessments) was *g*=0.20 (*P*<.001; 95% CI 0.13-0.27). A funnel plot analysis revealed that publication bias was unlikely, with no studies missing on the left side of the funnel plot (Figure 2). Indeed, the trim-and-fill algorithm suggested that, if anything, 78 ESs from 34 studies were missing at the *right* side of the funnel plot, suggesting a possible selection bias that excluded studies with *larger* ESs. After imputation of these missing ESs, an adjusted overall effect was estimated, which produced an average ES of *g*=0.40 (*P*<.001; 95% CI 0.33-0.46), somewhat larger than our initially estimated overall effect (Δ*g*=0.13). Nevertheless, it is important to note that the trim-and-fill analysis does not take the dependency in ESs into account. An Egger regression test, which better models dependencies among ESs, revealed that SE was a significant and positive predictor of ESs ($\beta_1$=1.65, *P*<.001; 95% CI 1.07-2.23), which may indicate publication bias rather than selection bias.

**Table 2.** Moderators of the effectiveness of mobile technology–delivered interventions for youth: higher-quality studies only[a].

| Characteristics | *k*[b] | Effect sizes, n | B$_0$ (intercept), *g* (95% CI) | B$_1$ (slope), *g* (95% CI) | *F* test (df$_1$, df$_2$) | *P* value |
|---|---|---|---|---|---|---|
| **Methodological characteristics** | | | | | | |
| Study quality | 38 | 428 | 0.20 (0.11 to 0.29)*** | –0.0005 (–0.03 to 0.03) | 0.001 (1, 426) | .97 |
| **Comparison group type** | | | | | 3.25 (2, 425) | .04 |
| No intervention (RC[c]; eg, wait-list) | 21 | 215 | 0.26 (0.18 to 0.35)*** | N/A[d] | | |
| Inert (eg, placebo or information-only) | 15 | 112 | 0.14 (0.05 to 0.23)** | –0.13 (–0.23 to –0.03)* | | |
| Clinical (eg, established intervention) | 7 | 101 | 0.12 (–0.01 to 0.26)[e] | –0.14 (–0.29 to 0.01)[e] | | |
| **Youth Characteristics** | | | | | | |
| Mean age (years) | 36 | 407 | 0.19 (0.11 to 0.26)*** | 0.01 (–0.0002 to 0.03)[e] | 3.75 (1, 405) | .05 |
| **Risk level and type** | | | | | 1.18 (3, 424) | .32 |
| General sample not selected for risk (RC) | 17 | 175 | 0.13 (0.03 to 0.24)** | N/A | | |
| Nonmental health (ie, medical) risks | 1 | 2 | 0.11 (–0.50 to 0.72) | –0.03 (–0.64 to 0.59) | | |
| Psychological or mental health at-risk sample | 16 | 229 | 0.25 (0.15 to 0.36)*** | 0.12 (–0.03 to 0.27) | | |
| Psychological clinical sample (diagnosis) | 4 | 22 | 0.31 (0.07 to 0.55)* | 0.18 (–0.09 to 0.44) | | |
| **Intervention characteristics** | | | | | | |
| **Prescribed frequency of use** | | | | | 1.79 (4, 423) | .13 |
| As much as feasible; ≥4 days per week (RC) | 22 | 306 | 0.23 (0.14 to 0.32)*** | N/A | | |
| Some days, or more than once a week | 5 | 27 | 0.22 (0.02 to 0.42)* | –0.01 (–0.22 to 0.21) | | |
| About once a week | 3 | 36 | 0.40 (0.15 to 0.66)*** | 0.18 (–0.09 to 0.44) | | |
| One-time session | 2 | 15 | 0.13 (–0.20 to 0.46) | –0.10 (–0.44 to 0.25) | | |
| Not stated, when needed, or at user discretion | 6 | 44 | 0.03 (–0.14 to 0.20) | –0.20 (–0.39 to –0.01)* | | |

[a] This table presents moderation results for higher-quality studies (*k*=38) only in cases where the statistical significance of the moderation effect differs from the full-sample (*k*=80) results presented in Table 1. The right columns list the omnibus *F* test and *P* value for each moderation test. The middle columns list the intercept (B$_0$), or mean effect size, and slope (B$_1$), an estimated unstandardized regression coefficient, of the relevant Hedges *g* statistics, with CIs around each. Effects and slopes that differ significantly from 0 are denoted with asterisks in the intercept (B$_0$) and slope (B$_1$) columns, respectively. For categorical moderators, each intercept represents the mean effect of a category, whereas each slope represents the difference in the mean effect between the category and reference category. Depending on its sign, the slope of a continuous moderator represents an increase or decrease in the effect size with each unit increase in the variable.

[b] Number of studies with relevant ES data for a given row. In cases of multiple interventions or comparisons, some studies were counted in multiple rows; thus, these numbers sometimes exceeded 38. Owing to missing data, some counts fell short of 38. Further details on what was included in different categories of the included moderators are provided in the *Methods* section.

[c] RC: reference category.

[d] N/A: not applicable (as the slope represents a comparison with the reference category).

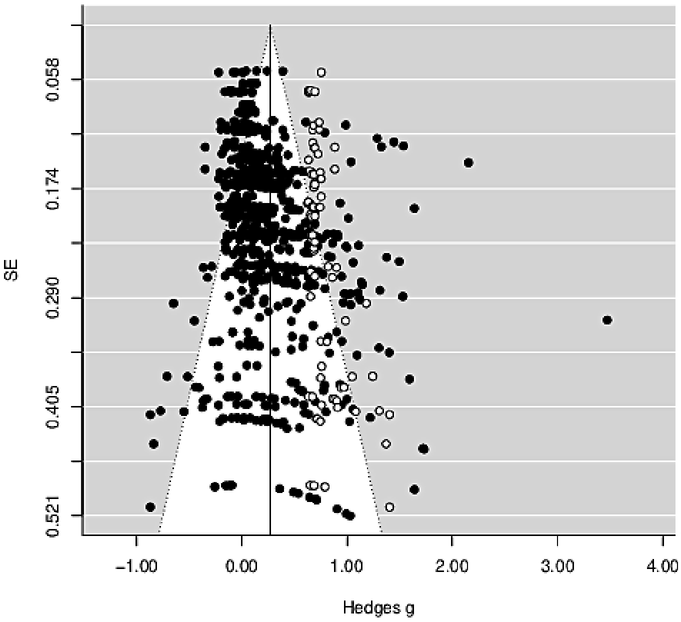
[e] *P*<.10.

*P<.05.

**P<.01.

***P<.001.

**Figure 2.** Funnel plot of observed mTDI effects (solid circles) and imputed effects (open circles) plotted against their standard error. mTDI: mobile technology–delivered intervention.



### Timing of Outcome Assessment

There were no significant differences in ESs immediately after the intervention versus those at longer-term follow-up assessments (Table 1).

### Outcome Type

The effectiveness of mTDIs varied as a function of the type of youth outcome that was targeted or assessed. The results in Table 1 indicate that there were statistically significant, positive effects of mTDIs on all of the coded outcome categories: general psychological distress or well-being, internalizing distress, noninternalizing mental health concerns, psychosocial strategies and skills, health-related outcomes, and *other* outcomes (see the *Methods* section). However, the *other* outcomes showed significantly lower ESs than the reference category, on average.

### Comparison Group Type

Contrary to expectations, in the full sample, the comparison group type did not moderate ES, such that effects were not statistically different across studies using no-intervention (eg, wait-list) or inert (eg, placebo or information-only) comparison groups, as well as studies using clinical treatments as their comparison group (Table 1). However, among the higher-quality studies, the results were more in line with our hypotheses, in that studies using inert comparison groups produced lower ESs than studies using no-intervention control groups, and studies using a clinical comparison no longer showed statistically significant effects on youth outcomes (Table 2).

## Differences in Effects of mTDIs Based on Youth Characteristics

The results showed that the mean age of the youth participants did not moderate the impact of the mTDIs (Table 1). Among the higher-quality studies, there was an effect right at the $P=.05$ threshold, such that the older the mean age of participants, the stronger the effect (Table 2). There were no differences in the study ESs as a function of the youth gender breakdown in the sample. Missing data on race and ethnicity limited our ability to analyze this variable as a moderator.

Youth level and type of risk significantly moderated intervention effects in the full sample (Table 1), such that samples with nonmental health (ie, medical) risks showed larger effects of mTDIs than general, unselected youth samples. However, in the subsample of higher-quality studies, this moderating effect was not found; in fact, the medical risk category dropped to one study and was no longer significantly different from zero (Table 2).

## Differences in Effects of mTDIs Based on Intervention Characteristics

### Primary Type of Technology

Moderation analysis did not detect statistically significant differences in the impact of mTDIs based on the primary type of technology: mTDIs were effective—and similar in their impact on youth outcomes—whether delivered on a smartphone or tablet, a presmartphone mobile device, or a mobile VR or handheld video game (Table 1).

### Guiding Theoretical Framework

As hypothesized, both cognitive behavioral and mindfulness- or acceptance-based interventions (as well as interventions that blended these 2 orientations) had significant effects on youth outcomes. Interventions grounded in one or multiple other theoretical frameworks also yielded significant effects and did not appear to differ systematically in their effects from cognitive behavioral interventions. Those mTDIs that were atheoretical or did not specify a guiding theoretical framework did not significantly differ from zero in their impact on youth outcomes, and the CI around their intercept (mean effect) was quite wide, indicating considerable heterogeneity. Of note, these studies were rare ($k$=3), and all 3 studies were dropped from the analysis of higher-quality studies; however, the overall pattern of results remained the same.

### *Technological and Support Features*

Exploratory analyses of the impact of intervention features and support failed to detect significant moderation of intervention effects based on the presence or absence of various technological features of the mTDI, including personalization, tailoring, social components, or gamification elements. Similarly, there were no differences in effects for mTDIs that integrated various support features, such as a training or orientation to the mTDI; some other in-person element; reminders to use the app; human or bot guidance, coaching, feedback in mTDI use; provision of supportive accountability; or supervised practice of skills taught by the mTDI. Although no significant differences were found between the absence and presence of any of these features and support types, it is notable that mTDIs both with and without each of these features had significant and positive mean effects (Table 1).

### *Dosage: Prescribed Frequency and Duration of Use*

There was a significant moderation effect for the prescribed frequency of mTDI use (Table 1). All prescribed use frequencies, except for leaving use to user discretion (including unstated use prescriptions), had a statistically significant impact on youth outcomes. Those mTDIs that involved prescribed use about once per week or were a single session yielded higher ESs than the reference category, which involved prescriptions of more frequent mTDI use (ie, at least 4 days per week or as much as feasible). However, this effect was not retained in the sample of higher-quality studies: The 2 remaining studies that prescribed a one-time session no longer yielded ESs that differed from zero statistically, and the prescribed use frequency was no longer a significant moderator of ESs (Table 2).

Additional moderation analyses probing the number of prescribed weeks or sessions of mTDI use did not detect significant effects for the prescribed duration of the intervention, whether the number of intervention sessions or weeks (Table 1). Although we also intended to examine the moderating effect of completed dosage (frequency, weeks, and sessions), there were substantial missing data, precluding meaningful analysis of these moderators.

# Discussion

## Principal Findings and Comparisons With Prior Work

To our knowledge, this study represents the first review and meta-analysis of mTDIs for a wide variety of youth well-being outcomes, an area of research that has grown rapidly in the past decade. Rigorous searches of the published and unpublished literature in this area yielded 80 studies evaluating 83 mTDIs for youth. A 3-level meta-analysis revealed an overall Hedges *g* of 0.27 across all youth outcomes and follow-up assessments, indicating a small effect that is generally consistent with the observed impact of mTDIs in other meta-analyses [42,62,67-69]. This finding addresses a critical gap in the existing literature in that most previous meta-analyses have focused solely on the effects of mTDIs in adult populations [42,53,63], and the few studies focusing specifically on youth have aggregated across diverse types of mobile and nonmobile technologies [64,65] or limited their scope to a specific subset of youth disorders (eg, internalizing disorders [67]).

It is worth noting that our sample included many studies of mTDIs that were still relatively early in their development and were, therefore, primarily interested in evaluating the feasibility and acceptability of the technology, although they also included measures of more distal mental health outcomes that they ultimately aimed to influence. Therefore, our analyses may underestimate, to some extent, the impact that these mTDIs would have had in larger or longer efficacy trials more specifically designed to influence youth mental health outcomes. As this literature continues to mature, it will be important to focus the inclusion criteria more specifically on studies that measure the effects of mTDIs on more distal mental health outcomes as their primary focus.

Our publication bias analyses yielded conflicting findings that were difficult to interpret, given the lack of conventions for analyzing publication bias in 3-level models. However, it is worth noting that despite rigorous screening criteria for the methodology of included studies, our coding revealed significant variability in study quality, with less than half of the studies in our sample comprising effects that surpassed our defined standards for typical research practices [89]. Study quality appeared to significantly influence the observed ESs such that as the score for study quality increased, the ESs generally decreased. Only a few prior meta-analyses of TDIs have directly assessed the influence of study quality on ES and found no impact [57,201]. However, our findings are consistent with some previous findings linking study quality to observed ESs for other psychological interventions [202,203] and suggest that attention to rigorous experimental methods, such as the reporting of intent-to-treat analyses and the use of well-validated and reliable assessment tools, are essential to accurately identify the impact of mTDIs for youth.

### *Methodological Characteristics Affecting Outcomes*

Given the substantial heterogeneity across ESs, both within and between studies, we explored several moderators as predictors of this variability. Interestingly, the ESs were similar in the immediate posttest assessments and longer-term follow-up assessments. This was true despite the fact that follow-up assessments occurred, on average, at 11.52 weeks, and ranged in length up to 43.14 weeks, after the active intervention period concluded. This finding is in contrast to the decrease in the effectiveness over time of some in-person mental health treatments [204-206] and suggests that the impact of mTDIs endures over time. It is possible that these enduring effects are because mTDIs are more easily integrated into youths' lives, therefore leading to either greater generalizability of the intervention effects, more lasting engagement with the mTDI, or both.

Somewhat contrary to expectations, moderator analyses in the overall sample also revealed that mTDIs had a similar impact on youth outcomes regardless of whether they were compared with a no-intervention control, such as a wait-list, or a more active comparison group, such as an information-only condition or usual clinical care. This helps to rule out the effects of expectancies, demand characteristics, or nonspecific effects accounting for the benefits of mTDIs in youth. This finding contributes to the somewhat mixed literature on this topic, with some past reviews of TDIs with both youth and adult samples finding that, more generally, ESs tend to differ based on comparison type (eg, higher for wait-list vs more active comparisons) [45,52,58,69,70], whereas others indicate that TDIs tend to be similarly effective across various types of study designs (eg, the study by Farrer et al [85]). Indeed, even in this meta-analysis, some findings shifted when only higher-quality studies were analyzed, such that studies with information-only or placebo comparisons yielded lower ESs than studies with no-intervention control groups, and studies with clinical comparison groups no longer showed statistically significant effects. Moreover, it is worth noting that the specific nature of the comparison group varied quite widely across studies, even within a particular coded category. As such, future research should continue to explore the marginal benefits of mTDIs over other available interventions.

The positive impact of mTDIs was observed across the diverse youth outcome categories assessed by each study, with the largest ESs for psychosocial strategies and skills (eg, emotional self-awareness, self-efficacy, and coping) and internalizing symptoms such as depression and anxiety, followed by general psychological distress and well-being, health concerns and health-related behaviors, and other noninternalizing mental health concerns (eg, attention difficulties, aggression, or delinquency). The smallest ES was observed for *other* outcomes (eg, knowledge, peer relationship quality, and stereotype threat), which showed significantly smaller effects than the reference category of general psychological distress. However, it is difficult to interpret this finding, given that our *other* category contained a diverse set of outcomes, many of which were coded very infrequently. As such, the overall ES for this *other* category is not necessarily reflective of the lower impact of mTDIs on each of these less commonly coded outcome types, and future research should continue to explore the scope of the impact of mTDIs on diverse youth problems. Nevertheless, these findings regarding outcome types generally suggest that mTDIs can be effective in treating a wide array of problems across youth development, including diverse areas of psychopathology (ie, both internalizing and externalizing domains), in addition to a number of cognitive, behavioral, and social risk factors that are often associated with poor mental health. This is a significant contribution to the literature, which has previously focused on narrower sectors of outcomes when analyzing the effectiveness of TDIs and mTDIs for youth [51,67,68].

### Youth Characteristics Affecting Outcomes

In the overall sample, there was no association between average youth age and the impact of mTDIs. However, when lower-quality studies were excluded from the analysis, the effects of youth age emerged more strongly, with ESs increasing as the mean participant age increased. This effect was right at the threshold for statistical significance ($P$=.05) and should, therefore, be interpreted with caution and replicated in future studies; however, this finding that suggests stronger effects of mTDIs for older youth is consistent with some past literature on TDIs more generally [51,54,55,65,72]. As many mental health interventions, including TDIs, were originally developed with adults in mind and only later adapted for youth at various stages of development, it is perhaps not surprising that mTDIs could have a more robust impact on older adolescents and young adults. For example, these youth may have greater internal motivation to engage with the intervention and be better able to interact with and adhere to the cognitive or behavioral skills taught by the mTDIs. However, given the complex ways in which developmental stages interact with risk for diverse mental health problems, as well as the effectiveness of mental health interventions, future research should continue to probe interactions between youth age and other dimensions of mTDIs (eg, level of human support, guiding theoretical framework, and availability of a social component) in predicting the impact of mTDIs.

Youth risk characteristics significantly moderated intervention effects in the full sample, with studies in which youth were selected for indicators of medical risk (eg, youth diagnosed with spina bifida or about to undergo surgery or another medical procedure) showing an average ES more than double that of studies with general, unselected samples of youth. Samples of youth with psychological risk (either clinically significant or subclinical risk) fell somewhere in the middle. However, in the analyses that excluded the lowest-quality studies, this moderating effect was no longer observed. The lack of differential findings for the impact of mTDIs on outcomes for youth with clinically significant versus subclinical risk, even when compared with general or unselected samples, is somewhat consistent with previous findings on TDIs, which tend to be quite mixed in terms of the impact of TDIs for youth with a variety of risk profiles [47,53,60,65]. Future research should continue to explore the presenting problems and risk indicators that are the best fit for referral to mTDIs versus more or less intensive interventions.

Other youth characteristics, in addition to age and risk factors, were not demonstrated to predict observed ESs. Youth race and ethnicity were reported inconsistently and according to widely varying conventions and, therefore, could not be tested as moderators of ESs. Furthermore, these identities and lived experiences are intertwined with structural inequalities and systems of discrimination and oppression that are more important for research to assess and relate to well-being outcomes. Consistent with several previous reviews [42,59], the breakdown of youth gender in the study sample did not predict ESs, although it is worth noting that studies infrequently made a note of nonbinary gender categories. These findings point to the need for a more careful and nuanced assessment of youth identities and lived experiences, including those connected to race, cultural identity, gender, sexual identity, and socioeconomic background, in studies testing the impact of mTDIs on youth.

### Intervention Characteristics Affecting Outcomes

Smartphone- or tablet-based mTDIs were by far the most commonly reported primary technology in our sample of studies relative to presmartphone mobile devices or other (mobile VR and handheld video game) technologies. The type of technology did not appear to moderate ESs, with each of these types of mTDIs yielding average ESs that were statistically significant and of a similar size. Although our sample was limited in number, mobile VR technologies are promising avenues for further research, especially given the effectiveness of these technologies for conditions such as posttraumatic stress disorder, depression, and pediatric pain and anxiety during medical procedures [207,208]. Although no studies in our sample used wearable devices as the primary type of technology, a handful used a smartphone along with some sort of wearable biosensor such as a sleep monitor or a physical activity wristwatch [113,183,188,195]. As these technologies are likely to become more common over time, research should continue to explore their effectiveness as a primary or supplemental feature of mTDIs for youth well-being.

Consistent with the previous literature on TDIs and mTDIs for youth and adults [51,53,54,69], both cognitive behavioral and mindfulness- or acceptance-based interventions (as well as interventions that blended these 2 orientations) had significant effects on youth outcomes. In traditional in-person treatment settings, cognitive behavioral interventions, including third-wave cognitive behavioral treatments that include components of mindfulness and acceptance, have become increasingly popular as empirical support has grown for their effectiveness in treating a wide range of childhood disorders, including anxiety, depression, conduct or aggression problems, and attention difficulties [209]. However, a growing body of literature shows that many youths and families are not able to access these *gold standard* evidence-based treatments, whether because of lack of availability in their community or issues with accessing mental health care in general, such as cost and stigma [26,210,211]. Therefore, it is encouraging to see that the effectiveness of these evidence-based interventions can be translated into low-cost, mobile technology–delivered formats that can reach far larger numbers of youth, and perhaps in a way that is more generalizable to the naturalistic environments of their lives. Interventions in our sample that were grounded in one or multiple *other* theoretical frameworks, such as positive psychology or motivational interviewing, also yielded significant effects and did not appear to differ systematically in their effects from cognitive behavioral interventions. Although few studies have examined (m)TDIs using these theoretical approaches, these findings are consistent with previous research that has evaluated the impact of these specific theoretical orientations [202,212]. Our sample also included 3 mTDIs that did not specify a guiding theoretical framework [154,166,171], and collectively, they did not significantly differ from zero in their impact on youth outcomes. These findings should be interpreted with caution, given the small number of studies, wide CIs around their intercepts (mean effects), and the fact that these studies were excluded from the analysis of higher-quality studies. Future research should continue to explore the impact of mTDIs grounded in diverse theoretical frameworks. With the vast and rapidly growing number of available mTDIs purporting to support the well-being and mental health of youth, it is critical to ascertain the theoretical frameworks that may lend themselves best to developing active interventions with strong empirical support for their effectiveness in a mobile technology–delivered format.

As mTDIs have become increasingly popular, many have begun to incorporate additional technological features intended to better leverage the technology-based format to engage and sustain users' attention. For example, some apps may personalize features of the intervention to the user's personal preferences or tailor the intervention based on a user's in-the-moment responses [75]. Others may include a social component, such as integration with social media platforms or a chat forum, or incorporate aspects of gamification, such as challenges or quests associated with points or badges [213]. Our moderator analyses showed no influence of these features on the ESs. However, it should be noted that many of these features are still relatively uncommon in mTDIs tested by research. For example, only 12% (10/83) of mTDIs in our meta-analysis mentioned a social component, and only 22% (18/83) described elements of personalization. Thus, the importance of these features may become more apparent as mTDIs targeting youth well-being begin to incorporate them more regularly and with greater proficiency. It is also likely that the most impactful mTDIs use an effective *combination* of these features to engage youth rather than simply incorporating one or another single design feature.

Given extensive theories [77,214,215] and some prior research on the benefits of outside guidance on engagement with mTDIs [53,57,216], we were somewhat surprised to see that the incorporation of support features, such as the provision of supportive accountability for technology use or supervised practice of the skills introduced by the mTDI, did not significantly moderate the ESs for mTDIs for youth. However, research in this area has been quite mixed, with several other studies finding little or no benefit from the inclusion of coaching or human support [42,54,55]. It is possible that mTDIs that do not rely on any component of human or bot support tend to be designed in a more comprehensive and self-contained way to offset this lack [42]. Moreover, the lack of significant moderation findings for these features in our meta-analysis does not necessarily indicate that these features are unimportant to the success of mTDIs in youth. Our meta-analysis captured an unusual sample of mTDI users, given that all effects were evaluated within the context of researcher-guided studies. As such, *all* participants were likely exposed to greater-than-usual accountability and support for their technology use as a function of taking part in a research study. This level of baseline accountability may have made it difficult to observe the added benefits of other forms of guidance or support. In addition, there are likely several kinds of informal human support—such as guidance or support from caregivers, teachers, or peers—that influence youth but were not typically assessed or reported in our sample of studies. Future research should continue to explore the kinds of support that are needed to maximize engagement with mTDIs for youth mental health and the ways in which these supports interact with factors such as youth age or risk characteristics (eg, younger or more clinically at-risk youth may require greater support).

Finally, moderation analyses provided tentative evidence that mTDIs can be effective regardless of the prescribed frequency or duration of use. In the overall sample, the number of prescribed weeks or sessions of use did not moderate the ESs.

Moreover, mTDIs yielded significant ESs across all prescribed use frequencies, except for mTDIs that did not prescribe a use frequency or left use up to the user's discretion. In the higher-quality sample, single-session interventions no longer yielded a significant ES; however, this finding should be interpreted with caution as only 2 single-session interventions remained in the higher-quality analytic sample.

These findings add to a growing body of mixed findings regarding the impact of prescribed and actual mTDI use on intervention outcomes [57,85,86]. Reviews of technology-based mental health interventions often highlight significant problems with treatment initiation and dropout [217-219], particularly for self-guided treatments that involve lower levels of structure and prescriptive guidance [220,221]. As such, it is critical for meta-analytic research to continue to explore trends in whether and how the prescribed dosage of mTDIs influences youth outcomes, with a particular focus on how different types of prescriptive guidance fit best with users' specific needs. For example, youth with more severe clinical diagnoses may require a different dosage of mTDI than those engaging with a prevention-oriented mTDI designed to improve general well-being. Notably, there was wide variability in how studies reported on mTDI dosage and adherence. We chose to analyze the *prescribed* dosage and frequency of mTDI use, given that these statistics were most consistently reported. It would be ideal to analyze the actual *completed* use or uptake of mTDIs among participants as well; however, this was reported inconsistently among the studies. Several studies reported participant use only among study completers or dropped unengaged or low-engaged users from the analysis [121,127,152], whereas others used financial incentives for protocol compliance [151], potentially introducing bias into the use statistics recorded in research studies that have additional levels of accountability built into the protocol.

## Limitations and Future Directions

The results of this meta-analysis should be interpreted with several limitations in mind. First, the quality of any meta-analysis is limited by the quality of the available primary studies. Only 60% of our coded ESs, which came from 48% (38/80) of the included studies, met our high-quality standards. We attempted to address this issue by running all analyses on both the full sample of studies and a subsample of higher-quality studies. Nevertheless, as a substantial number of studies and ESs of lower quality were dropped from the higher-quality subsample analysis, the statistical power declined, and some moderator categories could not be examined. Future research in this area should attend to existing procedures for designing and reporting on high-quality clinical intervention research. For example, there is a need for more studies that use larger sample sizes and retain larger percentages of their participants (regardless of their mTDI engagement), include more reliable outcome measurements, and use intent-to-treat analyses, as well as studies conducted by authors who were not involved with app development and are, therefore, able to provide a more unbiased assessment of the mTDI's impact.

Relatedly, our coding scheme yielded incomplete data for many of our hypothesized moderators, given the variability in reports on characteristics of the tested mTDI (eg, human support features, duration and frequency of use), as well as youth characteristics (eg, race and ethnicity, gender, and risk characteristics). Therefore, additional studies that carefully document these kinds of data are needed to more thoroughly test the various moderators of the overall effects of mTDIs.

Given the lack of standard approaches for assessing publication bias within a 3-level meta-analysis, we applied 2 different techniques that produced conflicting findings. The trim-and-fill analysis pointed toward a potential underestimation of the true overall effect, whereas the Egger test pointed toward a potential overestimation of the true overall effect (and thus publication bias). The Egger test—in which ES dependencies are modeled—is likely more valid than the trim-and-fill results for our multilevel study. Nevertheless, the results of both techniques should be interpreted with caution, as neither was developed for a 3-level meta-analysis, and both rely on an assumption of homogeneity in ESs, which is often not met in meta-analytic studies, including this study.

As noted previously, another limitation in interpreting the present findings is that youth mTDI use likely differs within versus outside of a research study. At the very least, research participants tend to be much more informed about intervention goals and receive more structured support in the process of using an mTDI than users outside the research context. In more naturalistic settings, such as clinical practice or completely self-guided use, mTDIs are likely to be used much more flexibly and may be adapted to the needs and circumstances of individual youths.

When considering the implications of the present findings for the use of mTDIs in clinical and other naturalistic settings, it is also essential to consider issues such as accessibility and cultural sensitivity of mTDIs, which were not examined in this meta-analysis. Some types of mTDI technologies, such as VR headsets, may be prohibitively expensive and less readily accessible to some youth and families, particularly those living or seeking treatment in low-resource settings. Moreover, wide variability in families' digital literacy and cultural norms around using technology to improve well-being is likely to play a key role in the effectiveness of these types of interventions. The studies in this meta-analysis were largely limited to Western cultural contexts, with very little research emerging from certain parts of the world (eg, our search did not yield any research from the African continent). In addition, reporting on youth characteristics such as socioeconomic status, race, and ethnicity was quite limited and followed widely varying conventions. Taken together, these issues limited our ability to test questions related to the cultural responsivity or tailoring of particular mTDIs based on youth cultural backgrounds—important questions that future research will need to explore to fulfill the promise of mTDIs for youth living in communities traditionally underserved by available in-person prevention and intervention programs.

In addition, participant ratings of social validity (ie, acceptability of mTDIs and user satisfaction), among the 25% (20/80) of studies for which these data were available, averaged under 60%, indicating that these technologies likely have room for improvement in user interface and user experience design. Further research should more diligently assess for different aspects of social validity, including qualitative feedback, and relate these elements to outcomes and potential moderators (eg, age of users, in-app features, and clinical severity of users), with the ultimate goal of improving the engagement and uptake, and thus impact, of mTDIs for youth.

Finally, future research should continue to explore the ideal setting and level of support for various mTDIs. The optimal approaches to integrating mTDIs with other mental health tools, as well as face-to-face interaction with mental health providers, remain largely an open question at this time. Although in some cases, mTDIs may serve as low-cost and accessible substitutions or adjunctive supports for face-to-face intervention programs in areas with limited access to health professionals, mTDIs may also be valuable as a way of socializing some youth to psychosocial interventions, with the goal of eventually connecting families to more traditional face-to-face services. In other cases, mTDIs may be most useful when accompanied by the support of a clinician or paraprofessional, such as a teacher, mentor, or academic advisor who guides the youth through the technology-based intervention [118,120,158].

## Study Strengths and Conclusions

This is the first comprehensive 3-level meta-analysis to evaluate the effects of mTDIs on diverse aspects of well-being in youth. We built on prior work by taking an inclusive but rigorous approach to testing the impact of interventions using various types of mobile technologies on diverse outcomes across childhood, adolescence, and young adulthood. Using a 3-level approach to meta-analysis, we were able to synthesize all relevant ESs while accounting for both within- and between-study heterogeneity in ESs and maximizing the statistical power of the analyses. Moreover, we coded a comprehensive set of more than 3 dozen potential moderators of study effects and found sufficient information to analyze the moderating role of >20 of these variables, including technological and support features (eg, human support and availability of in-app reminders) that are hypothesized to be critical to the success of these interventions but have rarely been tested as predictors of effects in previous meta-analyses.

Our synthesis of primary research confirms the significant benefits of mTDIs across a variety of psychosocial outcomes, comparison types (ie, no intervention, inert, and clinical), and time points (both immediate postintervention and longer-term follow-up effects). Although additional high-quality research on which kinds of mTDIs are most effective and under what conditions is clearly needed, we conclude that mTDIs have the potential to improve multiple aspects of youth well-being, and may confer significant, durable benefits in a broad array of domains, particularly for youth who are not otherwise getting their mental health needs met.

## Acknowledgments

The authors would like to thank Danielle Arntson, Jeremy Astesano, Maria Bandriwsky, Turner Block, Abigail Blum, Marie Chamberlain, Loan Ho, Cherrelle Jones, Lauren Nowakowski, Jean Rhodes, Emily Romero, Geert-Jan Stams, Allison Tetzlaff, and Charlotte Utschig. This research was supported by funding from the Center for Evidence-Based Mentoring.

## Authors' Contributions

CC and ER conceptualized the study, oversaw sample selection and coding, and drafted the manuscript. KB, SB, MH, NF, and KC coded studies for analysis and contributed to the manuscript. MA ran analyses and contributed to manuscript writing.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional search strategy details.
[PDF File (Adobe PDF File), 196 KB - mental_v9i7e34254_app1.pdf ]

Multimedia Appendix 2
Selected characteristics of 80 studies evaluating 83 mobile technology-delivered interventions for youth.
[PDF File (Adobe PDF File), 402 KB - mental_v9i7e34254_app2.pdf ]

## References

1. Bitsko RH, Holbrook JR, Ghandour RM, Blumberg SJ, Visser SN, Perou R, et al. Epidemiology and impact of health care provider-diagnosed anxiety and depression among US children. J Dev Behav Pediatr 2018 Jun;39(5):395-403 [FREE Full text] [doi: 10.1097/DBP.0000000000000571] [Medline: 29688990]
2. Twenge JM, Joiner TE, Rogers ML, Martin GN. Increases in depressive symptoms, suicide-related outcomes, and suicide rates among U.S. Adolescents after 2010 and links to increased new media screen time. Clin Psychol Sci 2017 Nov 14;6(1):3-17. [doi: 10.1177/2167702617723376]

3. Kessler RC, Amminger GP, Aguilar-Gaxiola S, Alonso J, Lee S, Ustün TB. Age of onset of mental disorders: a review of recent literature. Curr Opin Psychiatry 2007 Jul;20(4):359-364 [FREE Full text] [doi: 10.1097/YCO.0b013e32816ebc8c] [Medline: 17551351]
4. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):593-602. [doi: 10.1001/archpsyc.62.6.593] [Medline: 15939837]
5. Merikangas KR, He J, Burstein M, Swanson SA, Avenevoli S, Cui L, et al. Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication--Adolescent Supplement (NCS-A). J Am Acad Child Adolesc Psychiatry 2010 Oct;49(10):980-989 [FREE Full text] [doi: 10.1016/j.jaac.2010.05.017] [Medline: 20855043]
6. American College Health Association. American College Health Association-National College Health Assessment II: Reference Group Executive Summary Spring 2019. American College Health Association. 2019. URL: https://www.acha.org/documents/ncha/NCHA-II_SPRING_2019_US_REFERENCE_GROUP_EXECUTIVE_SUMMARY.pdf [accessed 2021-10-01]
7. Belden AC, Thomson NR, Luby JL. Temper tantrums in healthy versus depressed and disruptive preschoolers: defining tantrum behaviors associated with clinical problems. J Pediatr 2008 Jan;152(1):117-122 [FREE Full text] [doi: 10.1016/j.jpeds.2007.06.030] [Medline: 18154912]
8. Bucchianeri MM, Arikian AJ, Hannan PJ, Eisenberg ME, Neumark-Sztainer D. Body dissatisfaction from adolescence to young adulthood: findings from a 10-year longitudinal study. Body Image 2013 Jan;10(1):1-7 [FREE Full text] [doi: 10.1016/j.bodyim.2012.09.001] [Medline: 23084464]
9. Feld LD, Shusterman A. Into the pressure cooker: student stress in college preparatory high schools. J Adolesc 2015 Jun 14;41(1):31-42. [doi: 10.1016/j.adolescence.2015.02.003] [Medline: 25780894]
10. Leadbeater B, Thompson K, Gruppuso V. Co-occurring trajectories of symptoms of anxiety, depression, and oppositional defiance from adolescence to young adulthood. J Clin Child Adolesc Psychol 2012 Nov;41(6):719-730 [FREE Full text] [doi: 10.1080/15374416.2012.694608] [Medline: 22742519]
11. Eisenberg D, Lipson SK, Ceglarek P, Zhou S, Talaski A, Patterson A. The Healthy Minds Study Fall Data Report. Healthy Minds Network. URL: https://healthymindsnetwork.org/wp-content/uploads/2020/08/f2019_HMS_national_final.pdf [accessed 2021-10-01]
12. Ghandour RM, Sherman LJ, Vladutiu CJ, Ali MM, Lynch SE, Bitsko RH, et al. Prevalence and treatment of depression, anxiety, and conduct problems in US children. J Pediatrics 2019 Mar;206:256-67.e3 [FREE Full text] [doi: 10.1016/j.jpeds.2018.09.021] [Medline: 30322701]
13. Merikangas KR, He J, Burstein M, Swendsen J, Avenevoli S, Case B, et al. Service utilization for lifetime mental disorders in U.S. adolescents: results of the National Comorbidity Survey-Adolescent Supplement (NCS-A). J Am Acad Child Adolesc Psychiatry 2011 Jan;50(1):32-45 [FREE Full text] [doi: 10.1016/j.jaac.2010.10.006] [Medline: 21156268]
14. Whitney DG, Peterson MD. US national and state-level prevalence of mental health disorders and disparities of mental health care use in children. JAMA Pediatr 2019 Apr 01;173(4):389. [doi: 10.1001/jamapediatrics.2018.5399] [Medline: 30742204]
15. Hodgkinson S, Godoy L, Beers LS, Lewin A. Improving mental health access for low-income children and families in the primary care setting. Pediatrics 2017 Jan;139(1):e20151175 [FREE Full text] [doi: 10.1542/peds.2015-1175] [Medline: 27965378]
16. Avenevoli S, Swendsen J, He J, Burstein M, Merikangas KR. Major depression in the National Comorbidity Survey-adolescent supplement: prevalence, correlates, and treatment. J Am Acad Child Adolesc Psychiatry 2015 Jan;54(1):37-44.e2 [FREE Full text] [doi: 10.1016/j.jaac.2014.10.010] [Medline: 25524788]
17. Hom MA, Stanley IH, Joiner TE. Evaluating factors and interventions that influence help-seeking and mental health service utilization among suicidal individuals: a review of the literature. Clin Psychol Rev 2015 Aug;40:28-39. [doi: 10.1016/j.cpr.2015.05.006] [Medline: 26048165]
18. Gallagher R. National survey of college counseling centers 2014. The International Association of Counseling Services (IACS). 2015. URL: http://d-scholarship.pitt.edu/28178/ [accessed 2021-10-06]
19. LeViness P, Bershad C, Gorman K, Braun L, Murray T. The association for university and college counseling center directors annual survey – Public version 2018. The Association for University and College Counseling Center Directors. 2018. URL: https://www.aucccd.org/assets/documents/Survey/2018%20AUCCCD%20Survey-Public-June%2012-FINAL.pdf [accessed 2021-10-02]
20. Martínez-Hernáez A, DiGiacomo SM, Carceller-Maicas N, Correa-Urquiza M, Martorell-Poveda MA. Non-professional-help-seeking among young people with depression: a qualitative study. BMC Psychiatry 2014 Apr 28;14:124 [FREE Full text] [doi: 10.1186/1471-244X-14-124] [Medline: 24774644]
21. Radez J, Reardon T, Creswell C, Lawrence PJ, Evdoka-Burton G, Waite P. Why do children and adolescents (not) seek and access professional help for their mental health problems? A systematic review of quantitative and qualitative studies. Eur Child Adolesc Psychiatry 2021 Feb 21;30(2):183-211 [FREE Full text] [doi: 10.1007/s00787-019-01469-4] [Medline: 31965309]

22. Rickwood D, Deane FP, Wilson CJ, Ciarrochi J. Young people's help-seeking for mental health problems. Aus e J Advancement Mental Health 2014 Dec 17;4(3):218-251. [doi: 10.5172/jamh.4.3.218]
23. Salaheddin K, Mason B. Identifying barriers to mental health help-seeking among young adults in the UK: a cross-sectional survey. Br J Gen Pract 2016 Sep 29;66(651):e686-e692. [doi: 10.3399/bjgp16x687313] [Medline: 27688518]
24. van Vulpen KS, Habegar A, Simmons T. Rural school-based mental health services: parent perceptions of needs and barriers. Children Schools 2018;40(2):104-111. [doi: 10.1093/cs/cdy002]
25. Gronholm PC, Nye E, Michelson D. Stigma related to targeted school-based mental health interventions: a systematic review of qualitative evidence. J Affect Disord 2018 Nov;240:17-26. [doi: 10.1016/j.jad.2018.07.023] [Medline: 30041074]
26. Planey AM, Smith SM, Moore S, Walker TD. Barriers and facilitators to mental health help-seeking among African American youth and their families: a systematic review study. Child Youth Services Rev 2019 Jun;101:190-200. [doi: 10.1016/j.childyouth.2019.04.001]
27. Babatunde GB, van Rensburg AJ, Bhana A, Petersen I. Barriers and facilitators to child and adolescent mental health services in low-and-middle-income countries: a scoping review. Glob Soc Welf 2019 Jun 07;8(1):29-46. [doi: 10.1007/s40609-019-00158-z]
28. Ingoldsby EM. Review of interventions to improve family engagement and retention in parent and child mental health programs. J Child Fam Stud 2010 Oct 01;19(5):629-645. [doi: 10.1007/s10826-009-9350-2] [Medline: 20823946]
29. Salloum A, Johnco C, Lewin AB, McBride NM, Storch EA. Barriers to access and participation in community mental health treatment for anxious children. J Affect Disord 2016 May 15;196:54-61. [doi: 10.1016/j.jad.2016.02.026] [Medline: 26901657]
30. Sarikhani Y, Bastani P, Rafiee M, Kavosi Z, Ravangard R. Key barriers to the provision and utilization of mental health services in low-and middle-income countries: a scope study. Community Ment Health J 2021 Jul 13;57(5):836-852. [doi: 10.1007/s10597-020-00619-2] [Medline: 32285371]
31. Anderson M, Jiang J. Teens, social media and technology 2018. Pew Research Center. 2018. URL: https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/ [accessed 2021-10-02]
32. Launching the new influence central trend report!. Influence Central. 2018. URL: https://influence-central.com/trendspotting/launching-the-new-influence-central-trend-report [accessed 2021-10-08]
33. Rideout V, Fox S, Well Being Trust. Digital health practices, social media use, and mental well-being among teens and young adults in the U.S. Articles, Abstract, and Reports. 2018. URL: https://digitalcommons.psjhealth.org/publications/1093 [accessed 2021-10-02]
34. Kubb C, Foran HM. Online health information seeking by parents for their children: systematic review and agenda for further research. J Med Internet Res 2020 Aug 25;22(8):e19985 [FREE Full text] [doi: 10.2196/19985] [Medline: 32840484]
35. Struthers A, Charette C, Bapuji SB, Winters S, Ye X, Metge C, et al. The acceptability of E-mental health services for children, adolescents, and young adults: a systematic search and review. Can J Community Mental Health 2015 Jul 01;34(2):1-21. [doi: 10.7870/cjcmh-2015-006]
36. Pretorius C, Chambers D, Coyle D. Young people's online help-seeking and mental health difficulties: systematic narrative review. J Med Internet Res 2019 Nov 19;21(11):e13873 [FREE Full text] [doi: 10.2196/13873] [Medline: 31742562]
37. Wright JH, Mishkind M, Eells TD, Chan SR. Computer-assisted cognitive-behavior therapy and mobile apps for depression and anxiety. Curr Psychiatry Rep 2019 Jun 27;21(7):62. [doi: 10.1007/s11920-019-1031-2] [Medline: 31250242]
38. Ahmed I, Ahmad NS, Ali S, Ali S, George A, Saleem Danish H, et al. Medication adherence apps: review and content analysis. JMIR Mhealth Uhealth 2018 Mar 16;6(3):e62 [FREE Full text] [doi: 10.2196/mhealth.6432] [Medline: 29549075]
39. Bauer M, Glenn T, Geddes J, Gitlin M, Grof P, Kessing LV, et al. Smartphones in mental health: a critical review of background issues, current status and future concerns. Int J Bipolar Disord 2020 Jan 10;8(1):2 [FREE Full text] [doi: 10.1186/s40345-019-0164-x] [Medline: 31919635]
40. Bry LJ, Chou T, Miguel E, Comer JS. Consumer smartphone apps marketed for child and adolescent anxiety: a systematic review and content analysis. Behav Ther 2018 Mar;49(2):249-261 [FREE Full text] [doi: 10.1016/j.beth.2017.07.008] [Medline: 29530263]
41. Wu A, Scult MA, Barnes ED, Betancourt JA, Falk A, Gunning FM. Smartphone apps for depression and anxiety: a systematic review and meta-analysis of techniques to increase engagement. NPJ Digit Med 2021 Feb 11;4(1):20 [FREE Full text] [doi: 10.1038/s41746-021-00386-8] [Medline: 33574573]
42. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct;16(3):287-298 [FREE Full text] [doi: 10.1002/wps.20472] [Medline: 28941113]
43. Torous J, Andersson G, Bertagnoli A, Christensen H, Cuijpers P, Firth J, et al. Towards a consensus around standards for smartphone apps and digital mental health. World Psychiatry 2019 Feb;18(1):97-98 [FREE Full text] [doi: 10.1002/wps.20592] [Medline: 30600619]
44. Huguet A, Rao S, McGrath PJ, Wozney L, Wheaton M, Conrod J, et al. A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. PLoS One 2016;11(5):e0154248 [FREE Full text] [doi: 10.1371/journal.pone.0154248] [Medline: 27135410]

45. Grist R, Porter J, Stallard P. Mental health mobile apps for preadolescents and adolescents: a systematic review. J Med Internet Res 2017 May 25;19(5):e176 [FREE Full text] [doi: 10.2196/jmir.7332] [Medline: 28546138]
46. Colbert S, Thornton L, Richmond R. Smartphone apps for managing alcohol consumption: a literature review. Addict Sci Clin Pract 2020 May 07;15(1):17 [FREE Full text] [doi: 10.1186/s13722-020-00190-x] [Medline: 32381062]
47. Hollis C, Falconer CJ, Martin JL, Whittington C, Stockton S, Glazebrook C, et al. Annual research review: digital health interventions for children and young people with mental health problems - a systematic and meta-review. J Child Psychol Psychiatry 2017 Apr 10;58(4):474-503. [doi: 10.1111/jcpp.12663] [Medline: 27943285]
48. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. J Med Internet Res 2019 Sep 25;21(9):e14567 [FREE Full text] [doi: 10.2196/14567] [Medline: 31573916]
49. Lattie EG, Kashima K, Duffecy JL. An open trial of internet-based cognitive behavioral therapy for first year medical students. Internet Interv 2019 Dec;18:100279 [FREE Full text] [doi: 10.1016/j.invent.2019.100279] [Medline: 31534911]
50. Linardon J, Shatte A, Messer M, Firth J, Fuller-Tyszkiewicz M. E-mental health interventions for the treatment and prevention of eating disorders: an updated systematic review and meta-analysis. J Consult Clin Psychol 2020 Nov;88(11):994-1007. [doi: 10.1037/ccp0000575] [Medline: 32852971]
51. Ebert DD, Zarski A, Christensen H, Stikkelbroek Y, Cuijpers P, Berking M, et al. Internet and computer-based cognitive behavioral therapy for anxiety and depression in youth: a meta-analysis of randomized controlled outcome trials. PLoS One 2015;10(3):e0119895 [FREE Full text] [doi: 10.1371/journal.pone.0119895] [Medline: 25786025]
52. Harrer M, Adam SH, Baumeister H, Cuijpers P, Karyotaki E, Auerbach RP, et al. Internet interventions for mental health in university students: a systematic review and meta-analysis. Int J Methods Psychiatr Res 2019 Jun 26;28(2):e1759 [FREE Full text] [doi: 10.1002/mpr.1759] [Medline: 30585363]
53. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry 2019 Oct 09;18(3):325-336 [FREE Full text] [doi: 10.1002/wps.20673] [Medline: 31496095]
54. Pennant ME, Loucas CE, Whittington C, Creswell C, Fonagy P, Fuggle P, Expert Advisory Group. Computerised therapies for anxiety and depression in children and young people: a systematic review and meta-analysis. Behav Res Ther 2015 Apr;67:1-18. [doi: 10.1016/j.brat.2015.01.009] [Medline: 25727678]
55. Podina IR, Mogoase C, David D, Szentagotai A, Dobrean A. A meta-analysis on the efficacy of technology mediated CBT for anxious children and adolescents. J Rat Emo Cognitive Behav Ther 2015 Nov 17;34(1):31-50. [doi: 10.1007/s10942-015-0228-5]
56. Rooksby M, Elouafkaoui P, Humphris G, Clarkson J, Freeman R. Internet-assisted delivery of cognitive behavioural therapy (CBT) for childhood anxiety: systematic review and meta-analysis. J Anxiety Disord 2015 Jan;29:83-92. [doi: 10.1016/j.janxdis.2014.11.006] [Medline: 25527900]
57. Vigerland S, Lenhard F, Bonnert M, Lalouni M, Hedman E, Ahlen J, et al. Internet-delivered cognitive behavior therapy for children and adolescents: a systematic review and meta-analysis. Clin Psychol Rev 2016 Dec;50:1-10 [FREE Full text] [doi: 10.1016/j.cpr.2016.09.005] [Medline: 27668988]
58. Ye X, Bapuji SB, Winters SE, Struthers A, Raynard M, Metge C, et al. Effectiveness of internet-based interventions for children, youth, and young adults with anxiety and/or depression: a systematic review and meta-analysis. BMC Health Serv Res 2014 Jul 18;14:313 [FREE Full text] [doi: 10.1186/1472-6963-14-313] [Medline: 25037951]
59. Florean IS, Dobrean A, Păsărelu CR, Georgescu RD, Milea I. The efficacy of internet-based parenting programs for children and adolescents with behavior problems: a meta-analysis of randomized clinical trials. Clin Child Fam Psychol Rev 2020 Dec 08;23(4):510-528. [doi: 10.1007/s10567-020-00326-0] [Medline: 32897527]
60. Grist R, Croker A, Denne M, Stallard P. Technology delivered interventions for depression and anxiety in children and adolescents: a systematic review and meta-analysis. Clin Child Fam Psychol Rev 2019 Jun 18;22(2):147-171 [FREE Full text] [doi: 10.1007/s10567-018-0271-8] [Medline: 30229343]
61. Donker T, Van Esveld S, Fischer N, Van Straten A. 0Phobia - towards a virtual cure for acrophobia: study protocol for a randomized controlled trial. Trials 2018 Aug 09;19(1):433 [FREE Full text] [doi: 10.1186/s13063-018-2704-6] [Medline: 30092844]
62. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
63. Weisel KK, Fuhrmann LM, Berking M, Baumeister H, Cuijpers P, Ebert DD. Standalone smartphone apps for mental health-a systematic review and meta-analysis. NPJ Digit Med 2019 Dec 2;2(1):118 [FREE Full text] [doi: 10.1038/s41746-019-0188-8] [Medline: 31815193]
64. Conley CS, Durlak JA, Shapiro JB, Kirsch AC, Zahniser E. A meta-analysis of the impact of universal and indicated preventive technology-delivered interventions for higher education students. Prev Sci 2016 Aug;17(6):659-678. [doi: 10.1007/s11121-016-0662-3] [Medline: 27225631]

65. Domhardt M, Steubl L, Baumeister H. Internet- and mobile-based interventions for mental and somatic conditions in children and adolescents. Z Kinder Jugendpsychiatr Psychother 2020 Jan 13;48(1):33-46 [FREE Full text] [doi: 10.1024/1422-4917/a000625] [Medline: 30422059]
66. Lindhiem O, Bennett CB, Rosen D, Silk J. Mobile technology boosts the effectiveness of psychotherapy and behavioral interventions: a meta-analysis. Behav Modif 2015 Nov 17;39(6):785-804 [FREE Full text] [doi: 10.1177/0145445515595198] [Medline: 26187164]
67. Buttazzoni A, Brar K, Minaker L. Smartphone-based interventions and internalizing disorders in youth: systematic review and meta-analysis. J Med Internet Res 2021 Jan 11;23(1):e16490 [FREE Full text] [doi: 10.2196/16490] [Medline: 33427682]
68. Leech T, Dorstyn D, Taylor A, Li W. Mental health apps for adolescents and young adults: a systematic review of randomised controlled trials. Children and Youth Services Review 2021 Aug;127:106073. [doi: 10.1016/j.childyouth.2021.106073]
69. Gál É,  tefan S, Cristea IA. The efficacy of mindfulness meditation apps in enhancing users' well-being and mental health related outcomes: a meta-analysis of randomized controlled trials. Journal of Affective Disorders 2021 Jan;279:131-142. [doi: 10.1016/j.jad.2020.09.134]
70. Flujas-Contreras JM, García-Palacios A, Gómez I. Technology-based parenting interventions for children's physical and psychological health: a systematic review and meta-analysis. Psychol Med 2019 Aug;49(11):1787-1798. [doi: 10.1017/S0033291719000692] [Medline: 30977462]
71. Välimäki M, Anttila K, Anttila M, Lahti M. Web-based interventions supporting adolescents and young people with depressive symptoms: systematic review and meta-analysis. JMIR Mhealth Uhealth 2017 Dec 08;5(12):e180 [FREE Full text] [doi: 10.2196/mhealth.8624] [Medline: 29222079]
72. Bennett SD, Cuijpers P, Ebert DD, McKenzie Smith M, Coughtrey AE, Heyman I, et al. Practitioner review: unguided and guided self-help interventions for common mental health disorders in children and adolescents: a systematic review and meta-analysis. J Child Psychol Psychiatry 2019 Aug 18;60(8):828-847. [doi: 10.1111/jcpp.13010] [Medline: 30775782]
73. Auerbach RP, Mortier P, Bruffaerts R, Alonso J, Benjet C, Cuijpers P, WHO WMH-ICS Collaborators. WHO World mental health surveys international college student project: prevalence and distribution of mental disorders. J Abnorm Psychol 2018 Oct;127(7):623-638 [FREE Full text] [doi: 10.1037/abn0000362] [Medline: 30211576]
74. Campbell OL, Bann D, Patalay P. The gender gap in adolescent mental health: a cross-national investigation of 566,829 adolescents across 73 countries. SSM Popul Health 2021 Mar;13:100742 [FREE Full text] [doi: 10.1016/j.ssmph.2021.100742] [Medline: 33748389]
75. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. JMIR Ment Health 2016 Mar 01;3(1):e7 [FREE Full text] [doi: 10.2196/mental.4984] [Medline: 26932350]
76. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on internet-based mental health interventions — a systematic review. Internet Interv 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]
77. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011 Mar 10;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]
78. Heber E, Ebert DD, Lehr D, Cuijpers P, Berking M, Nobis S, et al. The benefit of web- and computer-based interventions for stress: a systematic review and meta-analysis. J Med Internet Res 2017 Feb 17;19(2):e32 [FREE Full text] [doi: 10.2196/jmir.5774] [Medline: 28213341]
79. Conley CS, Durlak JA, Kirsch AC. A Meta-analysis of universal mental health prevention programs for higher education students. Prev Sci 2015 May 7;16(4):487-507. [doi: 10.1007/s11121-015-0543-1] [Medline: 25744536]
80. Bakırcı-Taylor AL, Reed DB, McCool B, Dawson JA. mHealth improved fruit and vegetable accessibility and intake in young children. J Nutr Educ Behav 2019 May;51(5):556-566. [doi: 10.1016/j.jneb.2018.11.008] [Medline: 30638880]
81. Elliott SN, Frey JR, Davies M. Systems for assessing and improving students' social skills to achieve academic competence. In: Durlak JA, Domitrovich CE, Weissberg RP, Gullotta TP, editors. Handbook of social and emotional learning: Research and practice. New York, United States: The Guilford Press; 2015:301-319.
82. Gottfredson DC, Cook TD, Gardner FE, Gorman-Smith D, Howe GW, Sandler IN, et al. Standards of evidence for efficacy, effectiveness, and scale-up research in prevention science: next generation. Prev Sci 2015 Oct;16(7):893-926 [FREE Full text] [doi: 10.1007/s11121-015-0555-x] [Medline: 25846268]
83. Payton JW, Wardlaw DM, Graczyk PA, Bloodworth MR, Tompsett CJ, Weissberg RP. Social and emotional learning: a framework for promoting mental health and reducing risk behavior in children and youth. Journal of School Health 2000 May;70(5):179-185. [doi: 10.1111/j.1746-1561.2000.tb06468.x] [Medline: 10900594]
84. Salas E, Cannon-Bowers JA. The science of training: a decade of progress. Annu Rev Psychol 2001 Feb;52(1):471-499. [doi: 10.1146/annurev.psych.52.1.471] [Medline: 11148314]
85. Farrer L, Gulliver A, Chan JK, Batterham PJ, Reynolds J, Calear A, et al. Technology-based interventions for mental health in tertiary students: systematic review. J Med Internet Res 2013 May 27;15(5):e101 [FREE Full text] [doi: 10.2196/jmir.2639] [Medline: 23711740]

86. Linardon J, Fuller-Tyszkiewicz M. Attrition and adherence in smartphone-delivered interventions for mental health problems: a systematic and meta-analytic review. J Consult Clin Psychol 2020 Jan;88(1):1-13. [doi: 10.1037/ccp0000459] [Medline: 31697093]
87. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 4th Edition. Washington, DC: American Psychiatric Association; 1994.
88. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th Edition. Washington, DC: American Psychiatric Association; 2013.
89. van der Stouwe T, Gubbels J, Castenmiller YL, van der Zouwen M, Asscher JJ, Hoeve M, et al. The effectiveness of social skills training (SST) for juvenile delinquents: a meta-analytical review. J Exp Criminol 2020 Mar 03;17(3):369-396. [doi: 10.1007/s11292-020-09419-w]
90. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health 1998 Jun 01;52(6):377-384 [FREE Full text] [doi: 10.1136/jech.52.6.377] [Medline: 9764259]
91. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011 Oct 18;343(oct18 2):d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]
92. Thomas B, Ciliska D, Dobbins M, Micucci S. A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions. Worldviews Evidence Based Nursing 2004 Jun 23;1(3):176-184. [doi: 10.1111/j.1524-475x.2004.04006.x] [Medline: 17163895]
93. Bartko JJ, Carpenter WT. On the methods and theory of reliability. J Nerv Ment Dis 1976 Nov;163(5):307-317. [doi: 10.1097/00005053-197611000-00003] [Medline: 978187]
94. Meyer GJ. Assessing reliability: critical corrections for a critical examination of the Rorschach Comprehensive System. Psychol Assess 1997;9(4):480-489. [doi: 10.1037/1040-3590.9.4.480]
95. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to Meta-analysis, First edition. United Kingdom: John Wiley & Sons, Ltd; 2009.
96. Lipsey MW, Wilson DB. Practical Meta-Analysis. Thousand Oaks, CA: Sage Publications; 2001.
97. Mullen B. Advanced BASIC Meta-Analysis, First edition. Hillsdale, NJ: Lawrence Erlbaum Associates; 1989.
98. de Vries SL, Hoeve M, Assink M, Stams GJ, Asscher JJ. Practitioner review: effective ingredients of prevention programs for youth at risk of persistent juvenile delinquency--recommendations for clinical practice. J Child Psychol Psychiatry 2015 Feb 21;56(2):108-121. [doi: 10.1111/jcpp.12320] [Medline: 25143121]
99. van Loon AW, Creemers HE, Beumer WY, Okorn A, Vogelaar S, Saab N, et al. Can schools reduce adolescent psychological stress? A multilevel meta-analysis of the effectiveness of school-based intervention programs. J Youth Adolesc 2020 Jun 7;49(6):1127-1145 [FREE Full text] [doi: 10.1007/s10964-020-01201-5] [Medline: 32034632]
100. Assink M, Wibbelink CJ. Fitting three-level meta-analytic models in R: a step-by-step tutorial. TQMP 2016 Oct 01;12(3):154-174. [doi: 10.20982/tqmp.12.3.p154]
101. Cheung MW. Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. Psychol Methods 2014 Jun;19(2):211-229. [doi: 10.1037/a0032968] [Medline: 23834422]
102. Hox JJ. Multilevel Analysis: Techniques and Applications, second edition. New York: Routledge; 2010.
103. Van den Noortgate W, López-López JA, Marín-Martínez F, Sánchez-Meca J. Three-level meta-analysis of dependent effect sizes. Behav Res Methods 2013 Jun;45(2):576-594. [doi: 10.3758/s13428-012-0261-6] [Medline: 23055166]
104. Van den Noortgate W, López-López JA, Marín-Martínez F, Sánchez-Meca J. Meta-analysis of multiple outcomes: a multilevel approach. Behav Res Methods 2015 Dec 1;47(4):1274-1294. [doi: 10.3758/s13428-014-0527-2] [Medline: 25361866]
105. Viechtbauer W. Conducting meta-analyses in with the package. J Stat Soft 2010;36(3):1-48. [doi: 10.18637/jss.v036.i03]
106. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2020.
107. Knapp G, Hartung J. Improved tests for a random effects meta-regression with a single covariate. Stat Med 2003 Sep 15;22(17):2693-2710. [doi: 10.1002/sim.1482] [Medline: 12939780]
108. Rosenthal R. Writing meta-analytic reviews. Psychol Bull 1995;118(2):183-192. [doi: 10.1037/0033-2909.118.2.183]
109. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000;56(2):455-463. [doi: 10.1111/j.0006-341x.2000.00455.x] [Medline: 10877304]
110. Duval S, Tweedie R. A nonparametric “trim and fill” method of accounting for publication bias in meta-analysis. J Am Stat Assoc 2000 Mar;95(449):89-98. [doi: 10.1080/01621459.2000.10473905]
111. Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. J Clin Epidemiol 2000 Nov;53(11):1119-1129. [doi: 10.1016/s0895-4356(00)00242-0] [Medline: 11106885]
112. Anastasiadou D, Folkvord F, Brugnera A, Cañas Vinader L, SerranoTroncoso E, Carretero Jardí C, et al. An mHealth intervention for the treatment of patients with an eating disorder: a multicenter randomized controlled trial. Int J Eat Disord 2020 Jul 08;53(7):1120-1131. [doi: 10.1002/eat.23286] [Medline: 32383503]

113. Antle AN, Chesick L, Sridharan SK, Cramer E. East meets west: a mobile brain-computer system that helps children living in poverty learn to self-regulate. Pers Ubiquit Comput 2018 Jun 12;22(4):839-866. [doi: 10.1007/s00779-018-1166-x]
114. Arps ER, Friesen MD, Overall NC. Promoting youth mental health via text-messages: a New Zealand feasibility study. Appl Psychol Health Well Being 2018 Nov 19;10(3):457-480. [doi: 10.1111/aphw.12143] [Medline: 30230691]
115. Baskerville NB, Struik LL, Guindon GE, Norman CD, Whittaker R, Burns C, et al. Effect of a mobile phone intervention on quitting smoking in a young adult population of smokers: randomized controlled trial. JMIR Mhealth Uhealth 2018 Oct 23;6(10):e10893 [FREE Full text] [doi: 10.2196/10893] [Medline: 30355563]
116. Beidel DC, Tuerk PW, Spitalnick J, Bowers CA, Morrison K. Treating childhood social anxiety disorder with virtual environments and serious games: a randomized trial. Behav Ther 2021 Nov;52(6):1351-1363. [doi: 10.1016/j.beth.2021.03.003] [Medline: 34656191]
117. Bohleber L, Crameri A, Eich-Stierli B, Telesko R, von Wyl A. Can we foster a culture of peer support and promote mental health in adolescence using a web-based app? A control group study. JMIR Ment Health 2016 Sep 23;3(3):e45 [FREE Full text] [doi: 10.2196/mental.5597] [Medline: 27663691]
118. Borjalilu S, Mazaheri MA, Talebpour A. Effectiveness of mindfulness-based stress management in the mental health of Iranian university students: a comparison of blended therapy, face-to-face sessions, and mHealth app (Aramgar). Iran J Psychiatry Behav Sci 2019 May 12;13(2):e84726. [doi: 10.5812/ijpbs.84726]
119. Breitenstein SM, Fogg L, Ocampo EV, Acosta DI, Gross D. Parent use and efficacy of a self-administered, tablet-based parent training intervention: a randomized controlled trial. JMIR Mhealth Uhealth 2016 Apr 20;4(2):e36 [FREE Full text] [doi: 10.2196/mhealth.5202] [Medline: 27098111]
120. Broglia E, Millings A, Barkham M. Counseling with guided use of a mobile well-being app for students experiencing anxiety or depression: clinical outcomes of a feasibility trial embedded in a student counseling service. JMIR Mhealth Uhealth 2019 Aug 15;7(8):e14318 [FREE Full text] [doi: 10.2196/14318] [Medline: 31418424]
121. Bruehlman-Senecal E, Hook CJ, Pfeifer JH, FitzGerald C, Davis B, Delucchi KL, et al. Smartphone app to address loneliness among college students: pilot randomized controlled trial. JMIR Ment Health 2020 Oct 20;7(10):e21496 [FREE Full text] [doi: 10.2196/21496] [Medline: 33079071]
122. Choi EK, Jung E, Bae E, Ji Y, Lee A. Two-step integrative education program and mHealth for Korean children with spina bifida: a quasi-experimental pre-post study. J Pediatr Nurs 2020 Mar;51:e92-e99. [doi: 10.1016/j.pedn.2019.12.006] [Medline: 31924471]
123. Chow CH, Van Lieshout RJ, Schmidt LA, Buckley N. Tablet-based intervention for reducing children's preoperative anxiety: a pilot study. J Dev Behav Pediatr 2017;38(6):409-416. [doi: 10.1097/dbp.0000000000000454] [Medline: 28661955]
124. Clarke PJF, Bedford K, Notebaert L, Bucks RS, Rudaizky D, Milkins BC, et al. Assessing the therapeutic potential of targeted attentional bias modification for insomnia using smartphone delivery. Psychother Psychosom 2016;85(3):187-189. [doi: 10.1159/000442025] [Medline: 27043732]
125. Conley CS, Huguenel BM, Duffecy J, Silton R. Benefits of a mobile mindfulness application for depressed college students: preliminary findings from an ongoing RCT. 2019 Presented at: Association for Behavioral and Cognitive Therapies Annual Convention; 2019; Atlanta, GA.
126. Cumino DO, Vieira JE, Lima LC, Stievano LP, Silva RA, Mathias LA. Smartphone-based behavioural intervention alleviates children's anxiety during anaesthesia induction: a randomised controlled trial. Eur J Anaesthesiol 2017 Mar;34(3):169-175. [doi: 10.1097/EJA.0000000000000589] [Medline: 28146459]
127. de Niet J, Timman R, Bauer S, van den Akker E, Buijks H, de Klerk C, et al. The effect of a short message service maintenance treatment on body mass index and psychological well-being in overweight and obese children: a randomized controlled trial. Pediatr Obes 2012 Jun;7(3):205-219. [doi: 10.1111/j.2047-6310.2012.00048.x] [Medline: 22492669]
128. Deluca P, Coulton S, Alam MF, Boniface S, Cohen D, Donoghue K, et al. Brief interventions to prevent excessive alcohol use in adolescents at low-risk presenting to Emergency Departments: three-arm, randomised trial of effectiveness and cost-effectiveness. Int J Drug Policy 2021 Jul;93:103113 [FREE Full text] [doi: 10.1016/j.drugpo.2021.103113] [Medline: 33487528]
129. Earle AM, LaBrie JW, Boyle SC, Smith D. In pursuit of a self-sustaining college alcohol intervention: deploying gamified PNF in the real world. Addict Behav 2018 May;80:71-81 [FREE Full text] [doi: 10.1016/j.addbeh.2018.01.005] [Medline: 29407688]
130. Egilsson E, Bjarnason R, Njardvik U. Usage and weekly attrition in a smartphone-based health behavior intervention for adolescents: pilot randomized controlled trial. JMIR Form Res 2021 Feb 17;5(2):e21432 [FREE Full text] [doi: 10.2196/21432] [Medline: 33481750]
131. Elicherla SR, Bandi S, Nuvvula S, Challa RS, Saikiran KV, Priyanka VJ. Comparative evaluation of the effectiveness of a mobile app (Little Lovely Dentist) and the tell-show-do technique in the management of dental anxiety and fear: a randomized controlled trial. J Dent Anesth Pain Med 2019 Dec;19(6):369-378 [FREE Full text] [doi: 10.17245/jdapm.2019.19.6.369] [Medline: 31942452]
132. Fish MT, Saul AD. The gamification of meditation: a randomized-controlled study of a prescribed mobile mindfulness meditation application in reducing college students’ depression. Simulation Gaming 2019 Jun 04;50(4):419-435. [doi: 10.1177/1046878119851821]

133. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR Ment Health 2017 Jun 06;4(2):e19 [FREE Full text] [doi: 10.2196/mental.7785] [Medline: 28588005]
134. Flett JA, Conner TS, Riordan BC, Patterson T, Hayne H. App-based mindfulness meditation for psychological distress and adjustment to college in incoming university students: a pragmatic, randomised, waitlist-controlled trial. Psychol Health 2020 Sep 12;35(9):1049-1074. [doi: 10.1080/08870446.2019.1711089] [Medline: 32046499]
135. Gajecki M, Berman AH, Sinadinovic K, Rosendahl I, Andersson C. Mobile phone brief intervention applications for risky alcohol use among university students: a randomized controlled study. Addict Sci Clin Pract 2014 Jul 02;9:11 [FREE Full text] [doi: 10.1186/1940-0640-9-11] [Medline: 24985342]
136. Gajecki M, Andersson C, Rosendahl I, Sinadinovic K, Fredriksson M, Berman AH. Skills training via smartphone app for university students with excessive alcohol consumption: a randomized controlled trial. Int J Behav Med 2017 Oct 21;24(5):778-788 [FREE Full text] [doi: 10.1007/s12529-016-9629-9] [Medline: 28224445]
137. Gibson KL. Efficacy of a self-monitoring intervention for college students with attention problems. University of California, Los Angeles. 2015. URL: https://escholarship.org/uc/item/77j3q163 [accessed 2021-10-02]
138. Gipson CS, Chilton JM, Dickerson SS, Alfred D, Haas BK. Effects of a sleep hygiene text message intervention on sleep in college students. J Am Coll Health 2019 Jan 31;67(1):32-41. [doi: 10.1080/07448481.2018.1462816] [Medline: 29652630]
139. Glissmann C. Calm College: testing a brief mobile app meditation intervention among stressed college students. Arizona State University. 2018. URL: https://keep.lib.asu.edu/items/156759 [accessed 2021-10-02]
140. Gonzales R, Ang A, Murphy DA, Glik DC, Anglin MD. Substance use recovery outcomes among a cohort of youth participating in a mobile-based texting aftercare pilot program. J Subst Abuse Treat 2014 Jul;47(1):20-26 [FREE Full text] [doi: 10.1016/j.jsat.2014.01.010] [Medline: 24629885]
141. Gonzales R, Hernandez M, Murphy DA, Ang A. Youth recovery outcomes at 6 and 9 months following participation in a mobile texting recovery support aftercare pilot study. Am J Addict 2016 Jan 21;25(1):62-68 [FREE Full text] [doi: 10.1111/ajad.12322] [Medline: 26689171]
142. Gonzales-Castaneda R, McKay JR, Steinberg J, Winters KC, Yu CH, Valdovinos IC, et al. Testing mediational processes of substance use relapse among youth who participated in a mobile texting aftercare project. Subst Abus 2022 Oct 22;43(1):1-12. [doi: 10.1080/08897077.2019.1671941] [Medline: 31638878]
143. Grassi A, Preziosa A, Villani D, Riva G. A relaxing journey: the use of mobile phones for well-being improvement. Annual Rev CyberTherapy Telemed 2007;5:123-131.
144. Grassi A, Gaggioli A, Riva G. The green valley: the use of mobile narratives for reducing stress in commuters. Cyberpsychol Behav 2009 Apr;12(2):155-161. [doi: 10.1089/cpb.2008.0156] [Medline: 19275475]
145. Grassi A, Gaggioli A, Riva G. New technologies to manage exam anxiety. Annual Review of Cybertherapy and Telemedicine 2011;167:57-62. [doi: 10.3233/978-1-60750-766-6-57]
146. Greer S, Ramo D, Chang Y, Fu M, Moskowitz J, Haritatos J. Use of the chatbot "Vivibot" to deliver positive psychology skills and promote well-being among young people after cancer treatment: randomized controlled feasibility trial. JMIR Mhealth Uhealth 2019 Oct 31;7(10):e15018 [FREE Full text] [doi: 10.2196/15018] [Medline: 31674920]
147. Haug S, Schaub MP, Venzin V, Meyer C, John U. Efficacy of a text message-based smoking cessation intervention for young people: a cluster randomized controlled trial. J Med Internet Res 2013 Aug 16;15(8):e171 [FREE Full text] [doi: 10.2196/jmir.2636] [Medline: 23956024]
148. Haug S. Mobile phone text messaging to reduce alcohol and tobacco use in young people – a narrative review. Smart Homecare Technol TeleHealth 2013 Oct;1:11. [doi: 10.2147/shtt.s43222]
149. Haug S, Paz Castro R, Kowatsch T, Filler A, Dey M, Schaub MP. Efficacy of a web- and text messaging-based intervention to reduce problem drinking in adolescents: results of a cluster-randomized controlled trial. J Consult Clin Psychol 2017 Feb;85(2):147-159. [doi: 10.1037/ccp0000138] [Medline: 27606700]
150. Hides L, Dingle G, Quinn C, Stoyanov SR, Zelenko O, Tjondronegoro D, et al. Efficacy and outcomes of a music-based emotion regulation mobile app in distressed young people: randomized controlled trial. JMIR Mhealth Uhealth 2019 Jan 16;7(1):e11482 [FREE Full text] [doi: 10.2196/11482] [Medline: 30664457]
151. Hilt LM, Swords CM. Mindfulness mobile app for ruminative adolescents: a randomized controlled trial. In: Proceedings of the Anxiety and Depression Association of America Annual Conference. 2022 Presented at: Anxiety and Depression Association of America annual conference; 2022; Denver, CO.
152. Huberty J, Green J, Glissmann C, Larkey L, Puzia M, Lee C. Efficacy of the mindfulness meditation mobile app "Calm" to reduce stress among college students: randomized controlled trial. JMIR Mhealth Uhealth 2019 Jun 25;7(6):e14273 [FREE Full text] [doi: 10.2196/14273] [Medline: 31237569]
153. Johnson N, Bree O, Lalley EE, Rettler K, Grande P, Gani MO, et al. Effect of a social script iPad application for children with autism going to imaging. J Pediatr Nurs 2014 Nov;29(6):651-659. [doi: 10.1016/j.pedn.2014.04.007] [Medline: 24836052]
154. Kajitani K, Higashijima I, Kaneko K, Matsushita T, Fukumori H, Kim D. Short-term effect of a smartphone application on the mental health of university students: a pilot study using a user-centered design self-monitoring application for mental health. PLoS One 2020 Sep 25;15(9):e0239592 [FREE Full text] [doi: 10.1371/journal.pone.0239592] [Medline: 32976515]

155. Kauer SD, Reid SC, Crooke AH, Khor A, Hearps SJ, Jorm AF, et al. Self-monitoring using mobile phones in the early stages of adolescent depression: randomized controlled trial. J Med Internet Res 2012 Jun 25;14(3):e67 [FREE Full text] [doi: 10.2196/jmir.1858] [Medline: 22732135]
156. Reid SC, Kauer SD, Hearps SJ, Crooke AH, Khor AS, Sanci LA, et al. A mobile phone application for the assessment and management of youth mental health problems in primary care: a randomised controlled trial. BMC Fam Pract 2011 Nov 29;12:131 [FREE Full text] [doi: 10.1186/1471-2296-12-131] [Medline: 22123031]
157. Reid SC, Kauer SD, Hearps SJ, Crooke AH, Khor AS, Sanci LA, et al. A mobile phone application for the assessment and management of youth mental health problems in primary care: health service outcomes from a randomised controlled trial of mobiletype. BMC Fam Pract 2013 Jun 19;14(1):84 [FREE Full text] [doi: 10.1186/1471-2296-14-84] [Medline: 23782796]
158. Kennard BD, Goldstein T, Foxwell AA, McMakin DL, Wolfe K, Biernesser C, et al. As Safe As Possible (ASAP): a brief app-supported inpatient intervention to prevent postdischarge suicidal behavior in hospitalized, suicidal adolescents. Am J Psychiatry 2018 Sep 01;175(9):864-872 [FREE Full text] [doi: 10.1176/appi.ajp.2018.17101151] [Medline: 30021457]
159. Kenny R. Is There an App for That? Development and Evaluation of a Mobile App-based Mental Health Intervention for Young People. Dublin, Ireland: University College Dublin; 2016.
160. Kenny R, Dooley B, Fitzgerald A. Ecological momentary assessment of adolescent problems, coping efficacy, and mood states using a mobile phone app: an exploratory study. JMIR Ment Health 2016 Nov 29;3(4):e51 [FREE Full text] [doi: 10.2196/mental.6361] [Medline: 27899340]
161. Kenny R, Fitzgerald A, Segurado R, Dooley B. Is there an app for that? A cluster randomised controlled trial of a mobile app-based mental health intervention. Health Informatics J 2020 Sep 08;26(3):1538-1559 [FREE Full text] [doi: 10.1177/1460458219884195] [Medline: 31702409]
162. Kollei I, Lukas CA, Loeber S, Berking M. An app-based blended intervention to reduce body dissatisfaction: a randomized controlled pilot study. J Consult Clin Psychol 2017 Nov;85(11):1104-1108. [doi: 10.1037/ccp0000246] [Medline: 28857576]
163. Lee J, Jung H, Lee G, Kim H, Park S, Woo S. Effect of behavioral intervention using smartphone application for preoperative anxiety in pediatric patients. Korean J Anesthesiol 2013 Dec;65(6):508-518 [FREE Full text] [doi: 10.4097/kjae.2013.65.6.508] [Medline: 24427456]
164. Lee RA, Jung ME. Evaluation of an mHealth app (DeStressify) on university students' mental health: pilot trial. JMIR Ment Health 2018 Jan 23;5(1):e2 [FREE Full text] [doi: 10.2196/mental.8324] [Medline: 29362209]
165. Levin ME, Hicks ET, Krafft J. Pilot evaluation of the stop, breathe & think mindfulness app for student clients on a college counseling center waitlist. J Am Coll Health 2022 Jan 09;70(1):165-173 [FREE Full text] [doi: 10.1080/07448481.2020.1728281] [Medline: 32150519]
166. Liguori S, Stacchini M, Ciofi D, Olivini N, Bisogni S, Festini F. Effectiveness of an app for reducing preoperative anxiety in children: a randomized clinical trial. JAMA Pediatr 2016 Aug 01;170(8):e160533. [doi: 10.1001/jamapediatrics.2016.0533] [Medline: 27294708]
167. Marx LS. A mindful eating "App" for non-treatment-seeking university women with eating and weight concerns. ProQuest. 2016. URL: https://www.proquest.com/docview/1896581612/abstract/8306C60B5DDF4ADCPQ/1 [accessed 2021-10-02]
168. Mason M, Mennis J, Way T, Zaharakis N, Campbell LF, Benotsch EG, et al. Text message delivered peer network counseling for adolescent smokers: a randomized controlled trial. J Prim Prev 2016 Oct 7;37(5):403-420. [doi: 10.1007/s10935-016-0439-2] [Medline: 27388626]
169. Mason MJ, Zaharakis NM, Russell M, Childress V. A pilot trial of text-delivered peer network counseling to treat young adults with cannabis use disorder. J Subst Abuse Treat 2018 Jun;89:1-10 [FREE Full text] [doi: 10.1016/j.jsat.2018.03.002] [Medline: 29706169]
170. McCloud T, Jones R, Lewis G, Bell V, Tsakanikos E. Effectiveness of a mobile app intervention for anxiety and depression symptoms in university students: randomized controlled trial. JMIR Mhealth Uhealth 2020 Jul 31;8(7):e15418 [FREE Full text] [doi: 10.2196/15418] [Medline: 32735221]
171. Moore SC, Crompton K, van Goozen S, van den Bree M, Bunney J, Lydall E. A feasibility study of short message service text messaging as a surveillance tool for alcohol consumption and vehicle for interventions in university students. BMC Public Health 2013 Oct 25;13(1):1011 [FREE Full text] [doi: 10.1186/1471-2458-13-1011] [Medline: 24160674]
172. Newman MG, Jacobson NC, Rackoff GN, Bell MJ, Taylor CB. A randomized controlled trial of a smartphone-based application for the treatment of anxiety. Psychother Res 2021 Apr 14;31(4):443-454 [FREE Full text] [doi: 10.1080/10503307.2020.1790688] [Medline: 32662323]
173. Nguyen-Feng V. A Randomized Controlled Trial of a Mobile Ecological Momentary Stress Management Intervention for Students With and Without a History of Emotional Abuse. University of Minnesota. 2019. URL: http://conservancy.umn.edu/handle/11299/209119 [accessed 2021-10-02]
174. Nolan J. Effectiveness, Feasibility, and Acceptability of a Mindfulness Based Mobile Application for Undergraduate Health Science Students. ProQuest. 2019. URL: https://www.proquest.com/openview/80ad66e2794d4571056ed01e8e9f2902/1?pq-origsite=gscholar&cbl=51922&diss=y [accessed 2021-10-02]
175. O'Dea B, Han J, Batterham PJ, Achilles MR, Calear AL, Werner-Seidler A, et al. A randomised controlled trial of a relationship-focussed mobile phone application for improving adolescents' mental health. J Child Psychol Psychiatry 2020 Aug;61(8):899-913 [FREE Full text] [doi: 10.1111/jcpp.13294] [Medline: 32683737]

176. Orosa-Duarte A, Mediavilla R, Muñoz-Sanjose A, Palao A, Garde J, López-Herrero V, et al. Mindfulness-based mobile app reduces anxiety and increases self-compassion in healthcare students: a randomised controlled trial. Med Teach 2021 Jun;43(6):686-693. [doi: 10.1080/0142159X.2021.1887835] [Medline: 33645416]
177. Parisod H, Pakarinen A, Axelin A, Löyttyniemi E, Smed J, Salanterä S. Feasibility of mobile health game "Fume" in supporting tobacco-related health literacy among early adolescents: a three-armed cluster randomized design. Int J Med Inform 2018 May;113:26-37. [doi: 10.1016/j.ijmedinf.2018.02.013] [Medline: 29602430]
178. Patel A, Schieble T, Davidson M, Tran MC, Schoenberg C, Delphin E, et al. Distraction with a hand-held video game reduces pediatric preoperative anxiety. Paediatr Anaesth 2006 Oct;16(10):1019-1027. [doi: 10.1111/j.1460-9592.2006.01914.x] [Medline: 16972829]
179. Pbert L, Druker S, Crawford S, Frisard C, Trivedi M, Osganian SK, et al. Feasibility of a smartphone app with mindfulness training for adolescent smoking cessation: craving to quit (C2Q)-teen. Mindfulness (N Y) 2020 Mar;11(3):720-733 [FREE Full text] [doi: 10.1007/s12671-019-01273-w] [Medline: 33343761]
180. Pierce B. Tacting of function in college student mental health: an online and app-based approach to psychological flexibility. Utah State University. 2019. URL: https://digitalcommons.usu.edu/etd/7619 [accessed 2021-10-02]
181. Piskorz J, Czub M. Effectiveness of a virtual reality intervention to minimize pediatric stress and pain intensity during venipuncture. J Spec Pediatr Nurs 2018 Jan;23(1). [doi: 10.1111/jspn.12201] [Medline: 29155488]
182. Piskorz JE, Czub M, Šulžickaja B, Kiliś-Pstrusińska K. Mobile virtual reality distraction reduces needle pain and stress in children? Cyberpsychol 2020 Feb 21;14(1). [doi: 10.5817/cp2020-1-3]
183. Ponzo S, Morelli D, Kawadler JM, Hemmings NR, Bird G, Plans D. Efficacy of the digital therapeutic mobile app BioBase to reduce stress and improve mental well-being among university students: randomized controlled trial. JMIR Mhealth Uhealth 2020 Apr 06;8(4):e17767 [FREE Full text] [doi: 10.2196/17767] [Medline: 31926063]
184. Ranney ML, Pittman SK, Dunsiger S, Guthrie KM, Spirito A, Boyer EW, et al. Emergency department text messaging for adolescent violence and depression prevention: a pilot randomized controlled trial. Psychol Serv 2018 Nov;15(4):419-428 [FREE Full text] [doi: 10.1037/ser0000193] [Medline: 30382737]
185. Rodgers A, Corbett T, Bramley D, Riddell T, Wills M, Lin R, et al. Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging. Tob Control 2005 Aug;14(4):255-261 [FREE Full text] [doi: 10.1136/tc.2005.011577] [Medline: 16046689]
186. Rodgers RF, Donovan E, Cousineau T, Yates K, McGowan K, Cook E, et al. BodiMojo: efficacy of a mobile-based intervention in improving body image and self-compassion among adolescents. J Youth Adolesc 2018 Jul 18;47(7):1363-1372. [doi: 10.1007/s10964-017-0804-3] [Medline: 29349593]
187. Sandrick J, Tracy D, Eliasson A, Roth A, Bartel J, Simko M, et al. Effect of a counseling session bolstered by text messaging on self-selected health behaviors in college students: a preliminary randomized controlled trial. JMIR Mhealth Uhealth 2017 May 17;5(5):e67 [FREE Full text] [doi: 10.2196/mhealth.6638] [Medline: 28526664]
188. Siembor B. Exploring the effectiveness of a mindfulness training app for managing stress in a university student population: a pilot study. Northeastern University. 2017. URL: https://repository.library.northeastern.edu/files/neu:cj82rb241 [accessed 2021-10-02]
189. Suffoletto B, Kristan J, Callaway C, Kim KH, Chung T, Monti PM, et al. A text message alcohol intervention for young adult emergency department patients: a randomized clinical trial. Ann Emerg Med 2014 Dec;64(6):664-72.e4 [FREE Full text] [doi: 10.1016/j.annemergmed.2014.06.010] [Medline: 25017822]
190. Suffoletto B, Kristan J, Chung T, Jeong K, Fabio A, Monti P, et al. An interactive text message intervention to reduce binge drinking in young adults: a randomized controlled trial with 9-month outcomes. PLoS One 2015;10(11):e0142877 [FREE Full text] [doi: 10.1371/journal.pone.0142877] [Medline: 26580802]
191. Suffoletto B, Callaway C, Kristan J, Kraemer K, Clark DB. Text-message-based drinking assessments and brief interventions for young adults discharged from the emergency department. Alcohol Clin Exp Res 2012 Mar;36(3):552-560. [doi: 10.1111/j.1530-0277.2011.01646.x] [Medline: 22168137]
192. Tennant M, Youssef GJ, McGillivray J, Clark T, McMillan L, McCarthy MC. Exploring the use of immersive virtual reality to enhance psychological well-being in pediatric oncology: a pilot randomized controlled trial. Eur J Oncol Nurs 2020 Oct;48:101804. [doi: 10.1016/j.ejon.2020.101804] [Medline: 32949941]
193. Tighe J, Shand F, Ridani R, Mackinnon A, De La Mata N, Christensen H. Ibobbly mobile health intervention for suicide prevention in Australian Indigenous youth: a pilot randomised controlled trial. BMJ Open 2017 Jan 27;7(1):e013518 [FREE Full text] [doi: 10.1136/bmjopen-2016-013518] [Medline: 28132007]
194. Vu A. Randomized controlled trial of Pacifica, a CBT and mindfulness-based app for stress, depression, and anxiety management with health monitoring. University of Minnesota. 2018. URL: http://conservancy.umn.edu/handle/11299/216811 [accessed 2021-10-02]
195. Wang SS, Teo WZ, Teo WZ, Chai YW. Virtual reality as a bridge in palliative care during COVID-19. J Palliat Med 2020 Jun 01;23(6):756-756. [doi: 10.1089/jpm.2020.0212] [Medline: 32324080]
196. Wantanakorn P, Harintajinda S, Chuthapisith J, Anurathapan U, Rattanatamrong P. A new mobile application to reduce anxiety in pediatric patients before bone marrow aspiration procedures. Hosp Pediatr 2018 Oct;8(10):643-650. [doi: 10.1542/hpeds.2018-0073] [Medline: 30213798]

197. Whitehouse AJ, Granich J, Alvares G, Busacca M, Cooper MN, Dass A, et al. A randomised controlled trial of an iPad-based application to complement early behavioural intervention in Autism Spectrum Disorder. J Child Psychol Psychiatry 2017 Sep 25;58(9):1042-1052. [doi: 10.1111/jcpp.12752] [Medline: 28543302]
198. Whittaker R, Merry S, Dorey E, Maddison R. A development and evaluation process for mHealth interventions: examples from New Zealand. J Health Commun 2012;17 Suppl 1:11-21. [doi: 10.1080/10810730.2011.649103] [Medline: 22548594]
199. Whittaker R, Stasiak K, McDowell H, Doherty I, Shepherd M, Chua S, et al. MEMO: an mHealth intervention to prevent the onset of depression in adolescents: a double-blind, randomised, placebo-controlled trial. J Child Psychol Psychiatry 2017 Sep 02;58(9):1014-1022. [doi: 10.1111/jcpp.12753] [Medline: 28573672]
200. Yap AG, Roy RE, Lasala JR, Tan D, Hechanova MR, Diy WD, et al. Evaluation of a cognitive-behavioral game design-based mobile game on alcohol use for adolescents. Games Health J 2020 Oct 01;9(5):353-357. [doi: 10.1089/g4h.2019.0091] [Medline: 33054488]
201. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. Cogn Behav Ther 2018 Jan;47(1):1-18. [doi: 10.1080/16506073.2017.1401115] [Medline: 29215315]
202. Bolier L, Haverman M, Westerhof GJ, Riper H, Smit F, Bohlmeijer E. Positive psychology interventions: a meta-analysis of randomized controlled studies. BMC Public Health 2013 Feb 08;13:119 [FREE Full text] [doi: 10.1186/1471-2458-13-119] [Medline: 23390882]
203. Cuijpers P, van Straten A, Bohlmeijer E, Hollon SD, Andersson G. The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size. Psychol Med 2009 Jun 03;40(2):211-223. [doi: 10.1017/s0033291709006114]
204. Keles S, Idsoe T. A meta-analysis of group cognitive behavioral therapy (CBT) interventions for adolescents with depression. J Adolesc 2018 Aug 26;67(1):129-139. [doi: 10.1016/j.adolescence.2018.05.011] [Medline: 29957492]
205. Shea MT, Elkin I, Imber SD, Sotsky SM, Watkins JT, Collins JF, et al. Course of depressive symptoms over follow-up. Findings from the national institute of mental health treatment of depression collaborative research program. Arch Gen Psychiatry 1992 Oct 01;49(10):782-787. [doi: 10.1001/archpsyc.1992.01820100026006] [Medline: 1417430]
206. van Aar J, Leijten P, Orobio de Castro B, Overbeek G. Sustained, fade-out or sleeper effects? A systematic review and meta-analysis of parenting interventions for disruptive child behavior. Clin Psychol Rev 2017 Feb;51:153-163. [doi: 10.1016/j.cpr.2016.11.006] [Medline: 27930935]
207. Deng W, Hu D, Xu S, Liu X, Zhao J, Chen Q, et al. The efficacy of virtual reality exposure therapy for PTSD symptoms: a systematic review and meta-analysis. J Affect Disord 2019 Oct 01;257:698-709. [doi: 10.1016/j.jad.2019.07.086] [Medline: 31382122]
208. Eijlers R, Utens EM, Staals LM, de Nijs PFA, Berghmans JM, Wijnen RM, et al. Systematic review and meta-analysis of virtual reality in pediatrics. Anesthesia Analgesia 2019;129(5):1344-1353. [doi: 10.1213/ane.0000000000004165] [Medline: 31136330]
209. Weisz JR, McCarty CA, Valeri SM. Effects of psychotherapy for depression in children and adolescents: a meta-analysis. Psychol Bull 2006 Jan;132(1):132-149 [FREE Full text] [doi: 10.1037/0033-2909.132.1.132] [Medline: 16435960]
210. Gopalan G, Goldstein L, Klingenstein K, Sicher C, Blake C, McKay MM. Engaging families into child mental health treatment: updates and special considerations. J Can Acad Child Adolesc Psychiatry 2010 Aug;19(3):182-196 [FREE Full text] [Medline: 20842273]
211. Murry VM, Heflinger CA, Suiter SV, Brody GH. Examining perceptions about mental health care and help-seeking among rural African American families of adolescents. J Youth Adolesc 2011 Sep 23;40(9):1118-1131. [doi: 10.1007/s10964-010-9627-1] [Medline: 21259067]
212. Burke BL, Arkowitz H, Menchola M. The efficacy of motivational interviewing: a meta-analysis of controlled clinical trials. J Consult Clin Psychol 2003 Oct;71(5):843-861. [doi: 10.1037/0022-006X.71.5.843] [Medline: 14516234]
213. Wasil AR, Venturo-Conerly KE, Shingleton RM, Weisz JR. A review of popular smartphone apps for depression and anxiety: assessing the inclusion of evidence-based content. Behav Res Ther 2019 Dec;123:103498. [doi: 10.1016/j.brat.2019.103498] [Medline: 31707224]
214. Lattie EG, Schueller SM, Sargent E, Stiles-Shields C, Tomasino KN, Corden ME, et al. Uptake and usage of Intellicare: a publicly available suite of mental health and well-being apps. Internet Interv 2016 May;4(2):152-158 [FREE Full text] [doi: 10.1016/j.invent.2016.06.003] [Medline: 27398319]
215. Schueller SM, Tomasino KN, Mohr DC. Integrating human support into behavioral intervention technologies: the efficiency model of support. Clin Psychol Sci Pract 2017 Mar;24(1):27-45. [doi: 10.1037/h0101740]
216. Wright JH, Owen JJ, Richards D, Eells TD, Richardson T, Brown GK, et al. Computer-assisted cognitive-behavior therapy for depression: a systematic review and meta-analysis. J Clin Psychiatry 2019 Mar 19;80(2):18r12188 [FREE Full text] [doi: 10.4088/JCP.18r12188] [Medline: 30900849]
217. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 05;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]

218. Kaltenthaler E, Sutcliffe P, Parry G, Beverley C, Rees A, Ferriter M. The acceptability to patients of computerized cognitive behaviour therapy for depression: a systematic review. Psychol Med 2008 Nov;38(11):1521-1530. [doi: 10.1017/S0033291707002607] [Medline: 18205964]
219. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. Psychol Med 2009 May;39(5):705-712. [doi: 10.1017/S0033291708004224] [Medline: 18812006]
220. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009 Apr 24;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
221. Marcolino MS, Oliveira JA, D'Agostino M, Ribeiro AL, Alkmim MB, Novillo-Ortiz D. The impact of mHealth interventions: systematic review of systematic reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]

## Abbreviations

**ES:** effect size
**mTDI:** mobile technology–delivered intervention
**PRISMA:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses
**TDI:** technology-delivered intervention
**VR:** virtual reality

---

*Edited by J Torous; submitted 21.10.21; peer-reviewed by M Attridge, S Badawy; comments to author 26.12.21; revised version received 04.03.22; accepted 05.03.22; published 29.07.22.*

*Please cite as:*
*Conley CS, Raposa EB, Bartolotta K, Broner SE, Hareli M, Forbes N, Christensen KM, Assink M*
*The Impact of Mobile Technology-Delivered Interventions on Youth Well-being: Systematic Review and 3-Level Meta-analysis*
*JMIR Ment Health 2022;9(7):e34254*
*URL: https://mental.jmir.org/2022/7/e34254*
*doi:10.2196/34254*
*PMID:35904845*

Original Paper

# Peer-Presented Versus Mental Health Service Provider–Presented Mental Health Outreach Programs for University Students: Randomized Controlled Trial

Laurianne Bastien[1], MA; Bilun Naz Boke[1], MEd; Jessica Mettler[1], MA; Stephanie Zito[1], MEd; Lina Di Genova[2], PhD; Vera Romano[2], PhD; Stephen P Lewis[3], PhD; Rob Whitley[4], PhD; Srividya N Iyer[4], PhD; Nancy L Heath[1], PhD

[1]Department of Educational and Counselling Psychology, McGill University, Montreal, QC, Canada
[2]Student Services, McGill University, Montreal, QC, Canada
[3]Department of Psychology, University of Guelph, Guelph, ON, Canada
[4]Department of Psychiatry, McGill University, Montreal, QC, Canada

**Corresponding Author:**
Laurianne Bastien, MA
Department of Educational and Counselling Psychology
McGill University
3700 McTavish
Montreal, QC, H3A 1Y9
Canada
Phone: 1 5143981232
Email: laurianne.bastien@mail.mcgill.ca

## Abstract

**Background:** University students are reporting concerning levels of mental health distress and challenges. University mental health service provider initiatives have been shown to be effective in supporting students' mental health, but these services are often resource-intensive. Consequently, new approaches to service delivery, such as web-based and peer support initiatives, have emerged as cost-effective and efficient approaches to support university students. However, these approaches have not been sufficiently evaluated for effectiveness or acceptability in university student populations.

**Objective:** Thus, the overarching goal of this study was to evaluate a mental health service provider–presented versus peer-presented web-based mental health resilience–building video outreach program against a wait-list comparison group.

**Methods:** Participants were 217 undergraduate students (mean age 20.44, SD 1.98 years; 171/217, 78.8% women) who were randomly assigned to one of the intervention groups (mental health service provider–presented: 69/217, 31.8%; peer-presented: 73/217, 33.6%) or the wait-list comparison group (75/217, 34.6%). Participants in the intervention groups were asked to watch 3 brief skill-building videos addressing strategies for building mental health resilience, whereas the comparison group was wait-listed. The mental health service provider–presented and peer-presented video series were identical in content, with presenters using a script to ensure consistency across delivery methods, but the videos differed in that they were either presented by mental health service providers or university students (peers). All participants were asked to complete web-based self-report measures of stress, coping self-efficacy, social support, social connectedness, mindfulness, and quality of life at baseline (time 1), 6 weeks later (time 2, after the intervention), and 1-month follow-up (time 3).

**Results:** Results from a series of 2-way ANOVAs found no significant differences in outcomes among any of the 3 groups. Surprisingly, a main effect of time revealed that all students improved on several well-being outcomes. In addition, results for program satisfaction revealed that both the mental health service provider–presented and peer-presented programs were rated very highly and at comparable levels.

**Conclusions:** Thus, findings suggest that a web-based mental health resilience–building video outreach program may be acceptable for university students regardless of it being mental health service provider–presented or peer-presented. Furthermore, the overall increases in well-being across groups, which coincided with the onset and early weeks of the COVID-19 pandemic, suggest an unexpected pattern of response among university students to the early period of the pandemic. Limitations and barriers as well as research implications are discussed.

**Trial Registration:** ClinicalTrials.gov NCT05454592; https://clinicaltrials.gov/ct2/show/NCT05454592

***(JMIR Ment Health 2022;9(7):e34168)*** doi:10.2196/34168

**KEYWORDS**

web-based mental health outreach; resilience building; university student; peer-presented; mental health service provider–presented; mental health; outreach; resilience; student; service provider; randomized controlled trial

## *Introduction*

### Background

Over the past decade, mental health difficulties among university students have become a significant concern, with reports of 20% of students experiencing clinical depression and approximately 59% reporting experiencing above-average to tremendous levels of stress over the previous 12 months [1]. Although the traditional mental health service provider support offered through universities has been shown to be effective in increasing well-being among students, it is often costly and resource-intensive and can incur lengthy wait times because of the overwhelming demand [2-4]. Web-based peer initiatives have been suggested as cost-effective and efficient approaches to provide additional support and build capacity for mental health resilience among university students, but studies report mixed findings on the effectiveness of these approaches and a need for an evidence-based skill-building focus in these types of interventions [5-8]. Thus, the main objective of this study was to explore the acceptability and effectiveness of a web-based mental health resilience–building program as well as to evaluate differences between mental health service provider–presented and peer-presented variations of the program.

Evidence shows that university students are experiencing heightened levels of mental health distress. The National College Health Assessment survey across Canadian campuses with 55,284 student respondents revealed that 69% of students reported feeling overwhelming anxiety and 88% felt overwhelmed within the last year [1]. In addition to these heightened levels of stress, the developmental period of emerging adulthood, which is a theoretically and empirically distinct developmental period that takes place between adolescence and adulthood (ie, between the ages of 18 and 29 years), has been associated with a peak in unhealthy coping behaviors such as alcohol and drug abuse [9,10]. Characteristics of this developmental period include instability (ie, feeling like aspects of one's life such as relationships and work are unstable or easily subject to change) and feeling *in-between* (ie, feeling that they are not an adolescent anymore but not yet feeling like an adult), which have been found to have important mental health implications such as feelings of depression and anxiety [11,12]. As such, there is a clear need to provide university students with appropriate and effective support for building resilience and managing stress [13].

Leading organizations in health promotion have indicated a need for preventative programs aimed at enhancing mental health resilience. Specifically, the World Health Organization has identified increasing self-management and self-care ability through skill development as a core area to be addressed in efforts to enhance the mental health of emerging adults [14]. Furthermore, they note that increasing self-management and self-care would, in turn, result in concomitant decreases in demand for more intensive therapeutic interventions.

### Current Mental Health Support

Several evidence-based self-care and stress management strategies (eg, mindfulness strategies, progressive muscle relaxation, diaphragmatic breathing, and emotion regulation strategies) have been shown to promote resilience through effectively reducing stress and increasing well-being in emerging adults [15]. Students will often access these strategies through professional counseling support, which has been linked to a significant decrease in distress symptoms and improvement in academic performance [16,17].

Although many of these strategies and programs can be effective in supporting university students, they are often presented in individualized therapy and counseling sessions (eg, dialectical behavioral therapy and cognitive behavioral therapy), and these programs often operate at a significant financial cost [2,3]. Owing to heightened demands, this reduces the feasibility and access to such programs for all university students experiencing mental health distress (absent of mental illness) in university environments. These findings highlight the potential need for increasing access to these evidence-based skills to improve university students' ability to manage stress and enhance their coping during the challenging developmental period of emerging adulthood.

### New Approaches for Additional Support

Commensurate with the aforementioned strategies, new cost-effective approaches to service delivery are now being explored, including web-based mental health support. Web-based support takes advantage of the *new digital age*, in which increasing numbers of people (especially emerging adults) normatively obtain information, connection, and support via mobile phones, tablets, and home computers [18]. Web-based approaches also provide flexibility as they are self-paced (ie, fit into students' busy schedules as they can be used at any preferred time) [19,20]. Furthermore, another benefit of web-based mental health resources is the anonymity provided to students who may be reluctant to seek support because of stigma related to mental health [19]. Thus, the development of web-based mental health resources can provide improved access to evidence-based support to build mental health resilience on campuses, but new efforts must be based on solid research and scientific evidence [21]. Of note, an emerging body of research demonstrates that web-based mental health support can provide effective, efficient, and cost-effective support for individuals experiencing mental health distress (absent of mental illness), but there is still a need for innovation and evaluation to optimize student-oriented support [21,22]. Specifically, enhancing the relatability of the content or delivery of web-based mental health resources has been suggested to address low engagement or use of resources found in numerous studies of existing web-based support [23,24].

Peer support initiatives have also been identified as a promising approach for providing universal mental health resilience support. These initiatives require fewer professional resources and have been found to promote empowerment among individuals facing mental health challenges [7]. Peer support initiatives serve to fill gaps in official service provision and provide students with informal, peer-focused support with an emphasis on shared experience as opposed to psychopathology [25]. These initiatives allow peers to provide assistance to others drawing on their own lived experience of mental health challenges and to help others in their recovery journey [26]. They can also help decrease stigma and increase help-seeking behaviors through the sharing of information by those with similar experiences [7]. Indeed, evidence shows that, when students experience mental health difficulties, they tend to first turn to their peers for support in discussing these types of challenges [27,28]. Accordingly, peer support and web-based mental health outreach may be interesting to examine as approaches to provide access to evidence-based resilience-building strategies for university students. Although web-based mental health resources and peer support approaches are gaining popularity, many are not evidence-based, have privacy and confidentiality concerns, or have not been sufficiently evaluated for effectiveness or acceptability in university student populations [5,29,30]. Moreover, further research is needed to understand whether there are differences in the acceptability, satisfaction with, and effectiveness of a web-based mental health program as a function of whether the program is delivered by a mental health service provider or peer.

### This Study

Drawing on the aforementioned literature, the overarching goal of this study was to evaluate a mental health service provider–presented versus peer-presented mental health resilience skill–building web-based video outreach program against a wait-list comparison group. To our knowledge, this is the first study to examine acceptability of and satisfaction with a mental health service provider–presented versus peer-presented universal resilience-building web-based program in a sample of university students.

#### *Objective 1*

The first objective was to evaluate the acceptability of and satisfaction with a mental health service provider–presented versus peer-presented web-based skill-building video outreach program for university students. Given the exploratory nature of objective 1 of this study, no specific hypotheses were made.

#### *Objective 2*

The second objective was to compare group differences between a mental health service provider–presented versus peer-presented web-based skill-building video outreach program and a wait-list comparison group in terms of well-being outcomes over 10 weeks. It was hypothesized that the intervention groups (mental health service provider–presented and peer-presented) would demonstrate a greater increase in well-being outcomes (effectiveness) compared with the wait-list comparison group.

## *Methods*

### Participant Eligibility

Participants were eligible for the study if they were aged between 18 and 29 years given the unique stressors associated with the developmental period of emerging adulthood (age 18-29 years [11]). Furthermore, participants were required to have access to the internet (at least weekly) as the study was completed entirely on the web.

### Program Development and Description

The web-based mental health outreach program for university students was developed using an approach inspired by principles of the Participatory Action Research model, defined as “a partnership among equals with complementary knowledge and expertise” in which three key elements are collaboration, education, and action [31,32]. Consistent with the Participatory Action Research model, the program was developed using the expert knowledge of evidence-based strategies and best-practice applications of a multidisciplinary team of researchers (n=4), student service users with lived experience of mental health challenges (approximately 8-10 core team members who were consistently involved throughout the study and 15 team members whose participation in the project was fluid), mental health service providers (n=3), and decision makers (n=2). All stakeholders were actively involved throughout the project and consulted for project-related decisions (eg, study design and conceptualization and program development and dissemination). The multidisciplinary team met twice per month on average and reached consensus on all aspects of the program after lengthy discussion. In addition, meeting minutes were sent following each meeting, and all members were encouraged to reply via email directly to the project coordinator if they felt that (1) there was any discrepancy between the meeting minutes and what had been agreed upon, (2) there was any missing information, and (3) they had any questions or additional feedback they would like to provide regarding the decisions that were made. Alternatively, team members could edit the meeting minutes web-based document directly if they were more comfortable that way.

The web-based outreach program focused on four key areas of mental health resilience–building identified by the multistakeholder team’s expertise and review of the literature: dealing with stress, decreasing self-criticism, improving self-care and help-seeking behaviors, and enhancing social connections and social support [33-37]. Using videos, web-based infographics, guided audio recordings, and podcasts, students were provided with clear descriptions of each area of mental health resilience as well as a variety of evidence-based strategies (Table 1) specifically targeting one or more of these areas. The program was hosted entirely on the web, and students were encouraged to access the materials most relevant to their needs. A first video was sent to the students describing the web-based program, its overall focus, and how to access the skill-building strategies on the website’s resource library. At a 2-week interval, 2 subsequent videos were sent to (1) help students with problem solving for common challenges to strategy practice and (2) maintain long-term strategy practice habits. To assess differences in terms of preference for deliverer, two series of videos were created: one in which the deliverers were mental health service providers and one in which they were undergraduate students (ie, peers). Presenters were introduced at the beginning of the videos identifying themselves as a university student or mental health service provider working with university students. Their names and titles (mental health service provider or university student) also appeared at the bottom of the screen in the introduction. Attention was paid to ensure continuity of presenter characteristics with representation in terms of gender and race or ethnicity in both the mental health service provider–presented and peer-presented videos. The videos were identical in content in that the presenters used a script to ensure consistency across delivery methods; however, the videos differed in that they were either mental health service provider–presented or peer-presented. Students in the intervention groups (peer-presented and mental health service provider–presented) had access to the resource library throughout the study (ie, 9 weeks). The resource library was a website where students could select a strategy based on the area of resilience building that they wanted to work on, and they would be directed to an infographic or an audio recording to walk them through the different strategies. Students could access the resource library anywhere at any time as these resources were completely self-paced.

**Table 1.** Strategies presented in the resource library with the relevant key areas for resilience building that they address.

| Strategy | Key areas for resilience building that the strategy addresses | | | |
|---|---|---|---|---|
| | Dealing with stress | Decreasing self-criticism | Improving self-care and help seeking | Enhancing social connections and support |
| Mindfulness on the go (infographic) | ✓[a] | ✓ | ✓ | ✓ |
| Thought challenge (infographic) | ✓ | ✓ | ✓ | ✓ |
| Sitting meditation (audio recording) | ✓ | ✓ | ✓ | ✓ |
| Self-compassion meditation (audio recording) | ✓ | ✓ | ✓ | ✓ |
| Acceptance affirmation (infographic) | ✓ | ✓ | ✓ | ✓ |
| Body scan (audio recording) | ✓ | ✓ | ✓ | |
| Three good things (infographic) | ✓ | ✓ | ✓ | |
| Dealing with breakups (infographic and podcast) | ✓ | | ✓ | ✓ |
| Calming breath (audio recording) | ✓ | | ✓ | ✓ |
| Social network in university (infographic and podcast) | ✓ | | ✓ | ✓ |
| Beyond time management (infographic) | ✓ | | ✓ | ✓ |
| Physical well-being (infographic) | ✓ | | ✓ | |
| Riding the wave (infographic) | ✓ | ✓ | | |
| Self-care assessment (infographic) | ✓ | | ✓ | |
| Sleep hygiene (infographic) | ✓ | | ✓ | |
| Smart nutrition (infographic) | ✓ | | ✓ | |
| Yoga nidra (infographic and audio recording) | ✓ | | ✓ | |
| Progressive muscle relaxation (audio recording) | ✓ | | | |
| Financial wellness (infographic) | ✓ | | | |

[a] ✓: The strategy is presented in the resource library under the relevant area of resilience building.

## Procedure

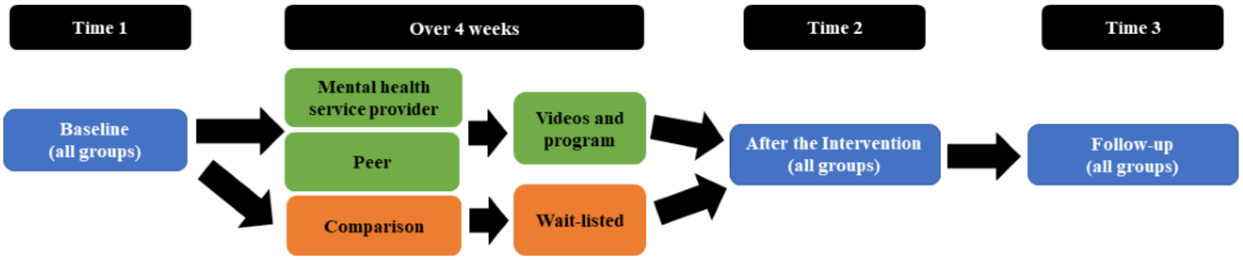
### *Overview*

Participants were recruited using a study flyer distributed to students in person on campus and on the web through email listservs and social media platforms and from an existing database of university students who had participated in previous studies and agreed to be followed up with. Given the self-paced nature of the program, a staggered recruitment approach was used wherein new participants completed the web-based baseline questionnaires between February 2020 and early March 2020. Participants were randomly assigned to one of the three groups (mental health service provider–presented, peer-presented, or a wait-list comparison group) while counterbalancing the three groups based on gender and preference for seeking help from mental health service providers or peers (ie, using results from the General Help-Seeking Questionnaire [GHSQ]) [38]. Specifically, to minimize the risk of imbalance of confounding factors (ie, gender and help-seeking preference) within the different groups, a minimization-based approach was used [39]. Minimization has been recommended for smaller trials (<1000) where specific participant factors (eg, gender) may influence the outcome [40]. Thus, participants were allocated to one of the three groups (mental health service provider–presented, peer-presented, or comparison) based on their scores on (1) formal and (2) informal help-seeking items on the GHSQ (classified as high, medium, or low scores) as well as (3) gender (classified as male, female, or nonbinary). Thus, the participants were randomized on an ongoing basis as they were enrolled in the study. As per minimization, the first participant is allocated to a group completely at random, but subsequent participants' group allocation depends on the characteristics of the participants already enrolled [39]. For example, if the next participant to be allocated had a profile with the following characteristics—female with high formal and medium informal help-seeking preferences—and there were already 5 participants with these characteristics in the mental health service provider–presented group but there were 6 participants in the peer-presented group and 6 participants in the comparison group, then this participant was allocated to the mental health service provider–presented group. The goal was for each allocation to minimize the imbalance across groups based on multiple factors.

A week after the baseline questionnaire was sent, participants in the intervention groups received either the mental health service provider–presented video or the peer-presented video (video 1) depending on which group they were randomly assigned to, as well as a link for access to the resource library. The next 2 videos were sent 2 weeks apart. Participants were encouraged to access the resource library over the duration of the program and were reminded with each video link sent. All participants then received postintervention (time 2) and follow-up (time 3) measures 6 and 10 weeks following baseline completion, respectively (see Figure 1 for the project timeline). Participants in the wait-list comparison group were only asked to complete evaluation measures at the 3 time points (they only received the videos and resource library at the end of the study).

Following completion of the study, the students received an email with a personalized profile indicating their individual scores on various measures, and all participants received full access to the program resources (videos and resource library). The participants were compensated CAD $10 (US $ 7.73) for each survey completed for a total of CAD $30 (US $23.18) and were also entered in a raffle for a 1 in 4 chance to win CAD $50 (US $38.64).

**Figure 1.** Project timeline for the stress and coping web-based outreach program.



### COVID-19 Context

In March 2020, when most students received the web-based mental health outreach program, a state of emergency was declared in the city in which this study was conducted. This resulted in the closure of all recreational centers, public parks and playgrounds, public libraries, bars, restaurants, movie theaters, concert venues, and places of worship as well as the banning of public gatherings. As per public health guidelines, all residents were recommended to stay home unless purchasing necessities (eg, food and supplies), for medical need, for essential work travel, or for 1 form of exercise per day. Strict social distancing guidelines prohibited in-person gatherings, and travel restrictions were implemented. In addition, the university in which this study took place was closed for a period of 2 weeks following the students' reading week (week off for spring break). There was a transition to web-based learning, the university allowed for flexibility for final assignments (students could be provided with extensions, and some final assignments were removed), and students were provided with a pass or fail option rather than a final grade. This was a period of drastic changes and increased social isolation for university students [41]. The data for this study were collected at three time points: before the pandemic (time 1; February 2020), pandemic onset (time 2; March 2020-April 2020), and early pandemic (time 3; April 2020-May 2020).

## Measures

### General Help-Seeking

The GHSQ [38] is a 10-item measure of formal and informal help seeking and uses the following prompt: "If you were having a personal or emotional problem, how likely is it that you would seek help from the following people?" The GHSQ was adapted to include help from classmates, academic advisors, residence supports, professors, research supervisors, and peer support organizations. Scores on willingness to seek support from peers and web-based resources (ie, informal) as well as willingness to seek support from professionals (ie, formal) were examined to randomize participants across program groups (mental health service provider–presented group, peer-presented group, and wait-list comparison group). The GHSQ has good reliability (Cronbach α=.91) and good construct validity [38]. In this study, the GHSQ had good internal consistency, with a Cronbach α at time 1 of .70.

### Training Satisfaction

The *Response to Training* is a researcher-developed measure assessing participants' satisfaction with the program content and delivery. The questions were delivered according to the three levels of the New World Kirkpatrick Model [42] as follows: (1) student viewers' response (ie, satisfaction, engagement, and relevance), (2) learning (ie, knowledge, skills, attitude, confidence, and commitment), and (3) use of skills (ie, willingness to use and frequency of use). All items were scored on a 4- to 6-point Likert scale where higher scores represented a better response to training. Sample items include "I would recommend the program to other university students" or "I am planning to use the program strategies in the future."

### Perceived Stress

The *Perceived Stress Scale* (PSS) [43] is a widely used self-report measure of individuals' perception of stress. This measure contains 10 items in which participants indicate their experience of stress on a 5-point Likert scale (0=*never* to 4=*very often*). The participants were asked to think about their lives over the previous month for baseline (consistent with the original scale) and over the previous 3 weeks (to assess the appropriate period after program use) for postintervention and follow-up measurements. Sample items include "In the past month/3 weeks, how often have you felt difficulties were piling up so high that you could not overcome them?" Ratings were averaged across items such that higher scores represented greater perceived stress. The PSS has good reliability (Cronbach α=.89), construct validity, and predictive validity with reports of psychological and physical symptoms [44,45]. In this study, the PSS had good internal consistency, with a Cronbach α of .86, .82, and .85 at time points 1, 2, and 3, respectively.

### Coping Self-efficacy

The *Coping Self-Efficacy Scale* (CSE) [46] is a measure of one's confidence in effectively engaging in coping behaviors in the face of challenges. This measure contains 26 items in which participants indicate confidence in their coping strategies when it comes to handling challenges and stressors on an 11-point Likert scale (0=*cannot do at all* to 10=*certain can do*). The participants were asked to think about their lives over the previous month for baseline (consistent with the original scale) and over the previous 3 weeks (to assess the appropriate period after program use) for postintervention and follow-up measurements. The CSE states that "When things aren't going well for you, or when you're having problems how confident or certain are you that you can do the following" and includes statements such as "find solutions to your most difficult problems" and "see things from the other person's point of view during a heated argument." Higher scores on the CSE represent higher coping self-efficacy. The CSE has good internal consistency (Cronbach α=.91) and test-retest reliability [46]. In this study, the CSE had good internal consistency, with a Cronbach α at time points 1, 2, and 3 of .92, .93, and .94, respectively.

### Social Support

The *Multidimensional Scale of Perceived Social Support* (MSPSS) [47] is a 12-item self-report questionnaire developed to assess the subjective perception of social support adequacy from family, friends, and significant others. Items are rated on a 7-point Likert scale (1=*strongly disagree* to 7=*strongly agree*). Participants were asked to think about their lives over the previous month for baseline (consistent with the original scale) and over the previous 3 weeks (to assess the appropriate period after program use) for postintervention and follow-up measurements. Sample items include "There is a special person who is around when I am in need" and "My family really tries to help me." Higher scores on the MSPSS represent higher perception of social support. The MSPSS has good reliability (Cronbach α ranging from .81 to .98) and good convergent and construct validity [48]. In this study, the MSPSS had good internal consistency, with a Cronbach α at time points 1, 2, and 3 of .89, .91, and .92, respectively.

### Social Connectedness

The *Social Connectedness Scale-Revised* (SCS-R) [49] is a 20-item self-report questionnaire that assesses emotional distance of the self from both friends and society along with maintaining a sense of closeness. Items are rated on a 6-point Likert scale (1=*strongly disagree* to 6=*strongly agree*). Participants were asked to think about their lives over the previous month for baseline (consistent with the original scale) and over the previous 3 weeks (to assess the appropriate period after program use) for postintervention and follow-up measurements. Sample items include "I feel distant from people" and "I am able to relate to my peers." Higher scores on the SCS-R represent higher perception of social connectedness. The SCS-R has good internal reliability (Cronbach α=.92) and good convergent and discriminant validity [49]. In this study, the SCS-R had good internal consistency, with a Cronbach α at time points 1, 2, and 3 of .90, .89, and .91, respectively.

### Mindfulness

The *Mindful Attention Awareness Scale* (MAAS) [50] measures individuals' dispositional mindfulness (ie, general tendency to be mindful) by assessing the frequency of mindful states over time. The MAAS consists of 15 items asking participants to report the frequency with which they have certain experiences on a 6-point scale (1=*almost always* to 6=*almost never*). Participants were asked to think about their lives over the previous month for baseline (consistent with the original scale) and over the previous 3 weeks (to assess the appropriate period after program use) for postintervention and follow-up measurements. Sample items include descriptions of experiences such as "I find myself preoccupied with the future or the past" and "I find myself doing things without paying attention." Scores for this measure are such that higher scores indicate higher levels of mindfulness. The MAAS has demonstrated strong internal consistency (Cronbach α=.89) as well as high test-retest reliability and convergent and discriminant validity [51]. In this study, the MAAS had good internal consistency, with a Cronbach α at time points 1, 2, and 3 of .80, .79, and .91, respectively.

### Quality of Life

The *World Health Organization Quality of Life Brief* questionnaire (WHOQOL-BREF) [52] is a 26-item measure assessing individuals' perception of their life quality within the following domains: physical health, psychological health, social relationships, and their environment. Participants are asked to rate items related to their experience of their own quality of life (QoL) on a 5-point Likert scale (1=*not at all* to 5=*extreme amount*). Participants were asked to think about their lives over the previous month for baseline (consistent with the original scale) and over the previous 3 weeks (to assess the appropriate period after program use) for postintervention and follow-up measurements. Sample items include "To what extent do you feel that physical pain prevents you from doing what you need to do?" and "How satisfied are you with the conditions of your living place?" The WHOQOL-BREF shows decent reliability (Cronbach α values for physical health, psychological health, social relationships, and environmental health were .65, .77, .52, and .79, respectively) and good internal consistency [53]. In this study, the WHOQOL-BREF had acceptable internal consistency. Specifically, the Cronbach α for the 4 domains ranged from .49 to .79 at time points 1, 2, and 3.

### Data Analysis

All data were analyzed using SPSS (version 26; IBM Corporation). The data were checked for patterns of missingness, univariate and multivariate outliers, and violations of assumptions before running the main analyses. A series of chi-square tests were used to test the first objective, which was to compare group differences on the web-based outreach program's acceptability between the types of deliverers (mental health service provider vs peer). A series of 2-way mixed ANOVAs were used to test the second objective, which was to compare group differences between a mental health service provider–presented and peer-presented web-based skill-building video outreach program and a wait-list comparison group in terms of well-being outcomes. To account for multiple pairwise comparisons (a total of three: mental health service provider, peer, and comparison) throughout the data analysis, the *P* value cutoff for statistical significance was set at .02 (.05 divided by 3) as per the Bonferroni correction.

### Ethics Approval

This study was approved by the McGill University Review Ethics Board (19-11-031).

## Results

### Participants

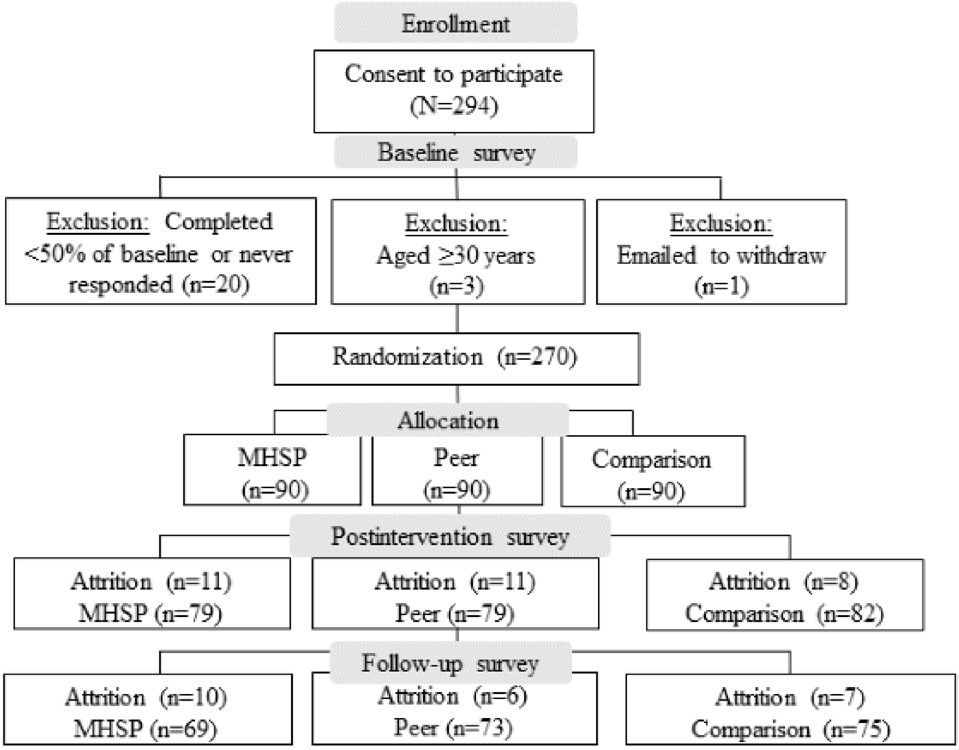
On the basis of data analysis requirements, a priori power analyses conducted with G*Power (version 3.1.9.6; Faul) with a medium effect size [54-56] and a power of 0.80 suggested minimum sample sizes of 186 [57]. Therefore, to account for attrition, a total of 294 undergraduate students were recruited (mean age 20.50, SD 2.35 years; 171/217, 78.8% women). However, of those 294 students who consented to participate, following data cleaning and participant withdrawal or dropout, the total sample of participants who completed all 3 time points was 217 (73.8%) undergraduate students (mean age 20.44, SD 1.98 years). Of this final sample of 217 undergraduate students, 171 (78.8%) self-identified as women, 42 (19.4%) self-identified as men, and 4 (1.8%) self-identified as nonbinary. The participants self-identified as White (132/217, 60.8%), Asian (54/217, 24.9%), Hispanic or Latino (11/217, 5.1%), African American or Black (10/217, 4.6%), Middle Eastern (7/217, 3.2%), and Indigenous (3/217, 1.4%) and, of these participants, 35.9% (78/217) reported being international students (the proportion of international students is comparable with the proportion at the university, which is approximately 30%) [58]. The participants were enrolled in different academic faculties, including Arts (88/217, 40.6%), Science (39/217, 18%), Agricultural and Environmental Studies (21/217, 9.7%), Engineering (20/217, 9.2%), Education (19/217, 8.8%), Management (10/217, 4.6%), and others (20/217, 9.2%). Of this sample, 74.7% (162/217) of students reported having experienced stress or mental health or well-being difficulties at a level that interfered with their ability to engage in the activities of everyday life (eg, school, work, relationships, and health-promoting behaviors) within the previous year. Furthermore, 25.3% (55/217) of the participants reported currently accessing mental health services such as counseling or therapy.

### Data Cleaning

A total of 294 individuals consented to participate in this study (230/294, 78.2% women; mean age 20.50, SD 2.35 years). Of this total sample, 0.3% (1/294) of the participants withdrew before being randomized to an intervention group providing a lack of time to participate in the study as a reasoning for the withdrawal, and 1% (3/294) were excluded as they were aged ≥30 years, which was an a priori exclusion criterion for the trial. In addition, 18% (53/294) of the participants were lost because of attrition (see Figure 2 for further details on the study sample).

Before running primary analyses, a missing values analysis was conducted and revealed that data were missing completely at random given that <5% of data points were missing per variable [59]. To preserve the sample size, the Expectation Maximization imputation method was used, where missing values were imputed within each subscale of measures in the mental health service provider–presented, peer-presented, and wait-list comparison groups separately to maximize prediction accuracy. The data were then screened for potential univariate outliers within each of the dependent variables. Cases 3 SDs above or below the mean were identified as potential outliers. A total of 14 potential univariate outliers were identified and Winsorized to a score with a 1-unit difference from the next most extreme score within each variable to maintain rank order. No multivariate outliers or violations of normality were found within any of the 3 groups. All assumptions for the 2-way mixed ANOVAs were met satisfactorily.

**Figure 2.** Attrition and exclusion of participants. MHSP: mental health service provider.



## Preliminary Analyses

To determine whether randomization effectively balanced the groups across well-being outcomes, multiple 1-way ANOVAs examining group (mental health service provider–presented, peer-presented, and wait-list comparison) differences in baseline stress, coping self-efficacy, social support, social connectedness, mindfulness, and QoL were conducted. Results from the 1-way ANOVAs revealed no significant differences, indicating that none of the groups differed on any of the well-being outcomes at baseline and that the groups were comparable. The means and SDs for the well-being outcomes of each group are shown in Table 2. In addition, demographics were comparable across groups. Gender was accounted for to ensure distribution across groups during randomization, and age was comparable across groups (mental health service provider–presented: mean age 20.40, SD 1.86 years; peer-presented: mean age 20.22, SD 1.96 years; wait-list comparison: mean age 20.58, SD 2.06 years).

**Table 2.** Preliminary 1-way ANOVAs for group differences (mental health service provider–presented, peer-presented, and wait-list comparison) at baseline.

| Variable | Mental health service provider–presented, mean (SD) | Peer-presented, mean (SD) | Wait-list comparison, mean (SD) | ANOVA results | |
|---|---|---|---|---|---|
| | | | | *F* test (*df*) | *P* value |
| Stress | 21.37 (5.76) | 22.58 (6.99) | 22.41 (5.77) | 0.84 (2,214) | .43 |
| Coping self-efficacy | 134.23 (28.97) | 137.1 (39.24) | 134.7 (38.37) | 0.20 (2,214) | .82 |
| Social support | 5.19 (0.96) | 5.22 (1.01) | 5.21 (1.07) | 0.07 (2,214) | .93 |
| Social connectedness | 75.71 (15.97) | 80.83 (18.06) | 78.79 (18.08) | 2.14 (2,214) | .12 |
| Mindfulness | 3.57 (0.81) | 3.64 (0.83) | 3.64 (0.74) | 0.37 (2,213) | .70 |
| QoL[a] (physical health) | 100.89 (15.95) | 102.00 (16.35) | 102.08 (16.35) | 0.21 (2,212) | .82 |
| QoL (psychological health) | 74.24 (15.57) | 74.58 (16.72) | 72.37 (17.29) | 0.30 (2,214) | .75 |
| QoL (social relationships) | 38.83 (9.66) | 39.56 (9.77) | 40.59 (9.34) | 0.84 (2,214) | .43 |
| QoL (environment) | 116.12 (17.18) | 116.8 (21.61) | 115.5 (18.61) | 0.10 (2,211) | .90 |

[a] QoL: quality of life.

## Main Analyses

### *Objective 1*

The first objective sought to evaluate the acceptability and satisfaction of a mental health service provider–presented versus peer-presented web-based skill-building video outreach program for university students. A series of chi-square tests were conducted using the training satisfaction survey at time 2 (Table 3). Results from the chi-square tests revealed that there was no significant difference between the mental health service provider–presented and peer-presented acceptability of the program on any of the training satisfaction items selected to represent each level of the Kirkpatrick model (student viewers' response, learning, and use of skills). Overall, the results of the training satisfaction survey demonstrate that most students were satisfied with the program (Table 3). For example, 81% (58/72) of the students in the mental health service provider–presented group and 91% (69/76) of the students in the peer-presented group said that they were planning to use the program strategies *sometimes* to *frequently*. In addition, 96% (69/72) of the students in the mental health service provider–presented group and 99% (75/76) of the students in the peer-presented group said that they *somewhat agreed* to *strongly agreed* that they would recommend the program strategies to other university students. By time 3, 65% (51/79) of the students in the mental health service provider–presented group and 78% (63/81) of the students in the peer-presented group reported having used the program strategies to cope with COVID-19 stress.

**Table 3.** Training satisfaction by group.

| Survey item and response options | Mental health service provider–presented, n (%) | Peer-presented, n (%) | Chi-square (*df*) | *P* value | Cramer *V* | *P* value |
|---|---|---|---|---|---|---|
| **I used the SCOOP[a] strategies** | | | 1.8 (2) | .40 | 0.112 | .22 |
| Never to rarely | 27 (38)[b] | 22 (29)[c] | | | | |
| Sometimes | 40 (56)[b] | 45 (59)[c] | | | | |
| Frequently | 5 (7)[b] | 9 (12)[c] | | | | |
| **I am planning to use the SCOOP strategies in the future.** | | | 4.0 (2) | .14 | 0.164 | .14 |
| Never to rarely | 14 (19)[b] | 7 (9)[c] | | | | |
| Sometimes | 36 (50)[b] | 48 (63)[c] | | | | |
| Frequently | 22 (31)[b] | 21 (28)[c] | | | | |
| **I would recommend the SCOOP strategies to other university students.** | | | 1.1 (2) | .59 | 0.084 | .59 |
| Strongly disagree to somewhat agree | 25 (35)[b] | 24 (32)[c] | | | | |
| Agree | 34 (47)[b] | 33 (43)[c] | | | | |
| Strongly agree | 13 (18)[b] | 19 (25)[c] | | | | |
| **Video 1—after watching this video, I feel I learned...** | | | 2.8 (2) | .24 | 0.139 | .24 |
| Nothing to a small amount | 25 (35)[d] | 23 (30)[c] | | | | |
| A medium amount | 33 (46)[d] | 30 (39)[c] | | | | |
| A lot | 13 (18)[d] | 23 (30)[c] | | | | |
| **Video 2—after watching this video, I feel I learned...** | | | 2.3 (2) | .32 | 0.129 | .32 |
| Nothing to a small amount | 29 (42)[e] | 21 (30)[f] | | | | |
| A medium amount | 33 (48)[e] | 39 (56)[f] | | | | |
| A lot | 7 (10)[e] | 10 (14)[f] | | | | |
| **Video 3—after watching this video, I feel I learned...** | | | 1.4 (2) | .49 | 0.101 | .49 |
| Nothing to a small amount | 37 (54)[e] | 44 (63)[f] | | | | |
| A medium amount | 25 (36)[e] | 19 (27)[f] | | | | |
| A lot | 7 (10)[e] | 7 (10)[f] | | | | |
| **In general, I found that the information and strategies presented in the resource library were useful to me.** | | | 1.9 (2) | .39 | 0.120 | .39 |
| Strongly disagree to somewhat agree | 30 (48)[g] | 28 (42)[h] | | | | |
| Agree | 28 (44)[g] | 30 (45)[h] | | | | |
| Strongly agree | 4 (6)[g] | 9 (13)[h] | | | | |
| **How much of the different material in the resource library did you actually use?** | | | 3.1 (1) | .45 | 0.153 | .45 |
| None of it to very little | 28 (44)[g] | 20 (30)[h] | | | | |
| Some, most, or all | 35 (56)[g] | 47 (70)[h] | | | | |

[a] SCOOP: Stress and Coping Online Outreach Program.

[b] n=72.

[c] n=76.

[d] n=71.

[e] n=69.

[f] n=70.

[g] n=63.

[h] n=67.

### Objective 2

The second objective sought to compare group differences between a mental health service provider–presented versus peer-presented web-based skill-building video outreach program and a wait-list comparison group in terms of well-being outcomes (ie, decreased stress and increased coping self-efficacy, social support, social connectedness, mindfulness, and QoL) at 3 different time points using a series of 2-way mixed ANOVAs. On the basis of the results from the Mauchly test of sphericity indicating that the assumption of sphericity was violated for some of the 2-way mixed ANOVAs, the Greenhouse-Geisser correction was used for all 2-way mixed ANOVAs for a more conservative approach. As presented in Table 4, the results did not reveal any significant 2-way interaction between group (mental health service provider–presented, peer-presented, or wait-list comparison) and time (baseline, after the intervention, and follow-up) on stress, coping self-efficacy, social support, social connectedness, mindfulness, and QoL, indicating that there was no effect of intervention group on any of the well-being outcomes over time. In addition, the results showed that there was no main effect of group for any of the outcomes assessed, which indicates that, regardless of time, there were no group differences on any of the well-being outcomes. However, as reported in Table 4, the main effect of time was statistically significant for coping self-efficacy, social support, mindfulness, the QoL social relationships domain, and the QoL environment domain, which indicates that, overall, regardless of group (mental health service provider–presented, peer-presented, or wait-list comparison), there was a change in these well-being outcomes over time. Pairwise comparisons were conducted using the Bonferroni correction set at a *P* value of .02 (.05 divided by 3 to account for the 3 comparisons used in the analysis) to assess between which time points the time effects occurred. As illustrated in Figures 3 and 4, the results showed that all students increased in coping self-efficacy and mindfulness from time 1 to time 2 and then remained stable at time 3. As illustrated in Figure 5, the results also showed that students increased in the QoL environment domain from time 1 to time 3, although time 2 was not statistically significant with any other time point. Finally, the results of the pairwise comparisons showed that both social support and the QoL social relationships domain significantly increased from time 1 to time 2 and then significantly decreased from time 2 to time 3 (Figures 6 and 7).

**Table 4.** Results of 3 (group: mental health service provider–presented, peer-presented, or wait-list comparison) × 3 (time: baseline, after the intervention, and follow-up) 2-way mixed ANOVAs on well-being outcomes.

| Well-being outcome and measurement | *F* test (*df*) | *P* value | $\eta_p^2$ | 1 – β | Chi-square (*df*) | *P* value |
|---|---|---|---|---|---|---|
| **Stress** | | | | | | |
| Mauchly test of sphericity—time | N/A[a] | N/A | N/A | N/A | 17.8 (2) | <.001 |
| Interaction—Greenhouse-Geisser | 1.55 (3.70,396.26) | .19 | 0.014 | .46 | N/A | N/A |
| Main effect of group (between) | 1.03 (2,214) | .36 | 0.010 | .23 | N/A | N/A |
| Main effect of time (within)—Greenhouse-Geisser | 3.70 (1.85,396.26) | .03 | 0.013 | .68 | N/A | N/A |
| **Coping self-efficacy** | | | | | | |
| Mauchly test of sphericity—time | N/A | N/A | N/A | N/A | 1.5 (2) | .48 |
| Interaction—Greenhouse-Geisser | 0.61 (3.97,423.08) | .70 | 0.006 | .20 | N/A | N/A |
| Main effect of group (between) | 0.40 (2,213) | .67 | 0.004 | .11 | N/A | N/A |
| Main effect of time (within)—Greenhouse-Geisser | 24.52 (1.99,423.08) | <.001 | 0.103 | 1 | N/A | N/A |
| **Social support** | | | | | | |
| Mauchly test of sphericity—time | N/A | N/A | N/A | N/A | 18.3 (2) | .001 |
| Interaction—Greenhouse-Geisser | 1.94 (3.69,396.52) | .11 | 0.018 | .56 | N/A | N/A |
| Main effect of group (between) | 0.70 (2,211) | .50 | 0.007 | .17 | N/A | N/A |
| Main effect of time (within)—Greenhouse-Geisser | 7.04 (1.85,396.56) | .001 | 0.032 | .93 | N/A | N/A |
| **Social connectedness** | | | | | | |
| Mauchly test of sphericity—time | N/A | N/A | N/A | N/A | 11.3 (2) | .004 |
| Interaction—Greenhouse-Geisser | 1.43 (3.80,399.07) | .22 | 0.013 | .43 | N/A | N/A |
| Main effect of group (between) | 2.11 (2,210) | .12 | 0.020 | .43 | N/A | N/A |
| Main effect of time (within)—Greenhouse-Geisser | 1.74 (1.90,399.07) | .18 | 0.008 | .36 | N/A | N/A |
| **Mindfulness** | | | | | | |
| Mauchly test of sphericity—time | N/A | N/A | N/A | N/A | 8.0 (2) | .02 |
| Interaction—Greenhouse-Geisser | 0.88 (3.86,408.89) | .48 | 0.008 | .28 | N/A | N/A |
| Main effect of group (between) | 0.14 (2,214) | .99 | <0.001 | .05 | N/A | N/A |
| Main effect of time (within)—Greenhouse-Geisser | 9.66 (1.93,408.89) | <.001 | 0.044 | .98 | N/A | N/A |
| **QoL[b] (physical health)** | | | | | | |
| Mauchly test of sphericity—time | N/A | N/A | N/A | N/A | 1.5 (2) | .48 |
| Interaction—Greenhouse-Geisser | 0.67 (3.97,417.05) | .62 | 0.006 | .21 | N/A | N/A |
| Main effect of group (between) | 0.35 (2,210) | .96 | <0.001 | .06 | N/A | N/A |
| Main effect of time (within)—Greenhouse-Geisser | 0.63 (1.99,417.05) | .53 | 0.003 | .16 | N/A | N/A |
| **QoL (psychological health)** | | | | | | |
| Mauchly test of sphericity—time | N/A | N/A | N/A | N/A | 0.8 (2) | .69 |
| Interaction—Greenhouse-Geisser | 0.23 (3.99,420.49) | .93 | 0.002 | .10 | N/A | N/A |
| Main effect of group (between) | 0.63 (2,211) | .53 | 0.006 | .15 | N/A | N/A |
| Main effect of time (within)—Greenhouse-Geisser | 0.65 (1.99,420.49) | .53 | 0.003 | .16 | N/A | N/A |
| **QoL (social relationships)** | | | | | | |

| Well-being outcome and measurement | *F* test (*df*) | *P* value | $\eta_p^2$ | 1 – β | Chi-square (*df*) | *P* value |
|---|---|---|---|---|---|---|
| Mauchly test of sphericity—time | N/A | N/A | N/A | N/A | 11.5 (2) | .003 |
| Interaction—Greenhouse-Geisser | 0.29 (3.80,398.68) | .87 | 0.003 | .11 | N/A | N/A |
| Main effect of group (between) | 0.86 (2,210) | .42 | 0.008 | .20 | N/A | N/A |
| Main effect of time (within)—Greenhouse-Geisser | 4.57 (1.90,398.68) | .01 | 0.021 | .76 | N/A | N/A |
| **QoL (environment)** | | | | | | |
| Mauchly test of sphericity—time | N/A | N/A | N/A | N/A | 6.6 (2) | .03 |
| Interaction—Greenhouse-Geisser | 0.27 (3.88,405.37) | .90 | 0.003 | .11 | N/A | N/A |
| Main effect of group (between) | 0.12 (2,209) | .89 | 0.001 | .07 | N/A | N/A |
| Main effect of time (within)—Greenhouse-Geisser | 7.89 (1.94,405.73) | <.001 | 0.036 | .95 | N/A | N/A |

[a] N/A: not applicable.

[b] QoL: quality of life.

**Figure 3.** University students' reported coping self-efficacy over time. Main effect of time represents a significant difference between time 1 and time 2 as well as between time 1 and time 3. CSE: Coping Self-Efficacy Scale; MHSP: mental health service provider.

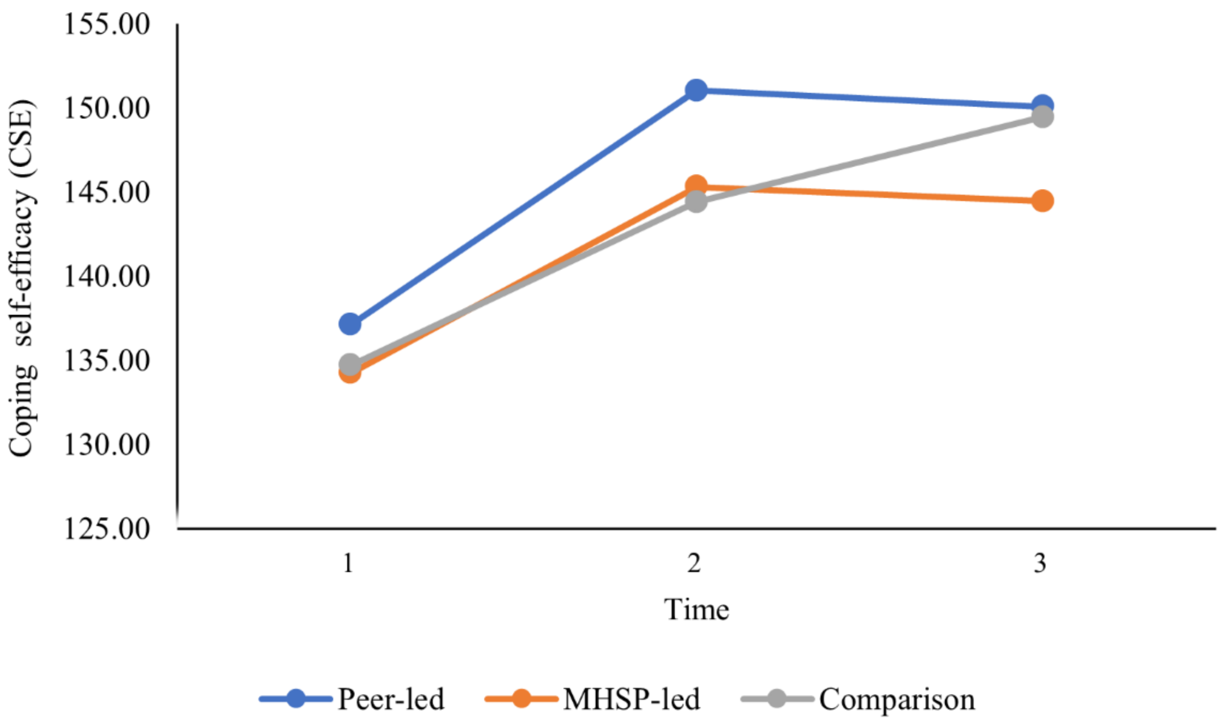


**Figure 4.** University students' reported mindfulness over time. Main effect of time represents a significant difference between time 1 and time 2 as well as between time 1 and time 3. MAAS: Mindful Attention Awareness Scale; MHSP: mental health service provider.

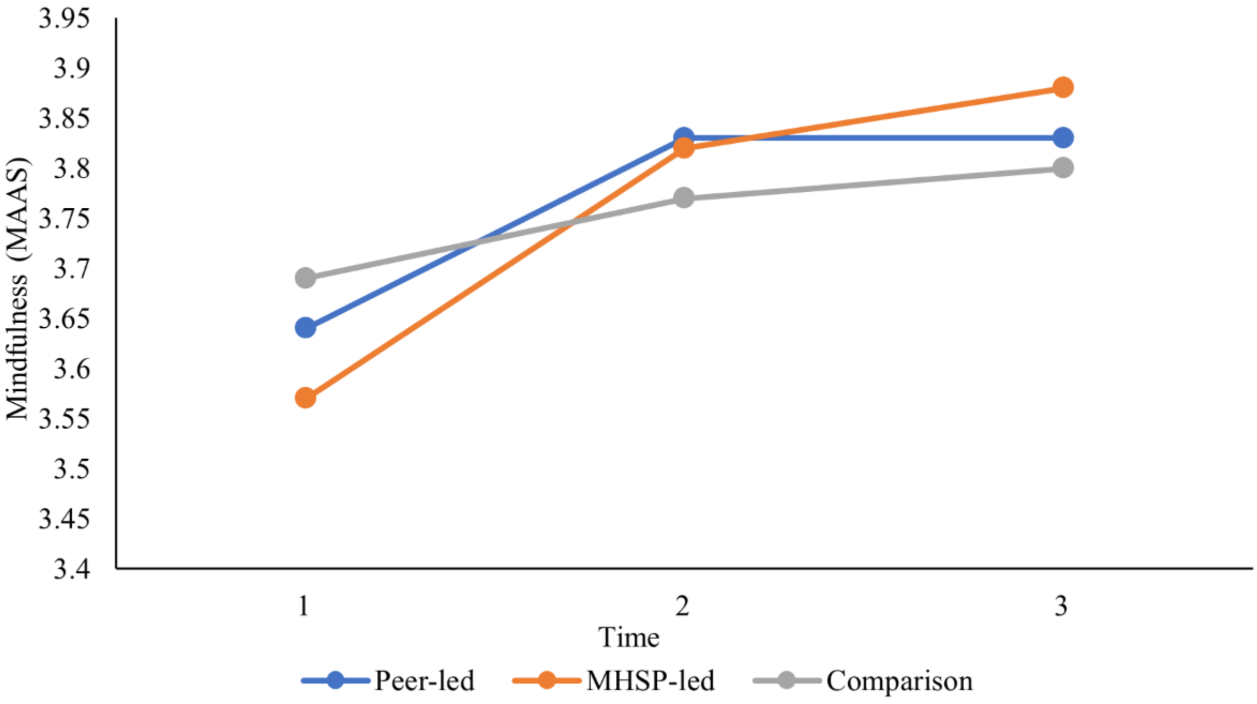


**Figure 5.** University students' reported quality of life (QoL; environment) over time. Main effect of time represents a significant difference between time 1 and time 3. MHSP: mental health service provider.

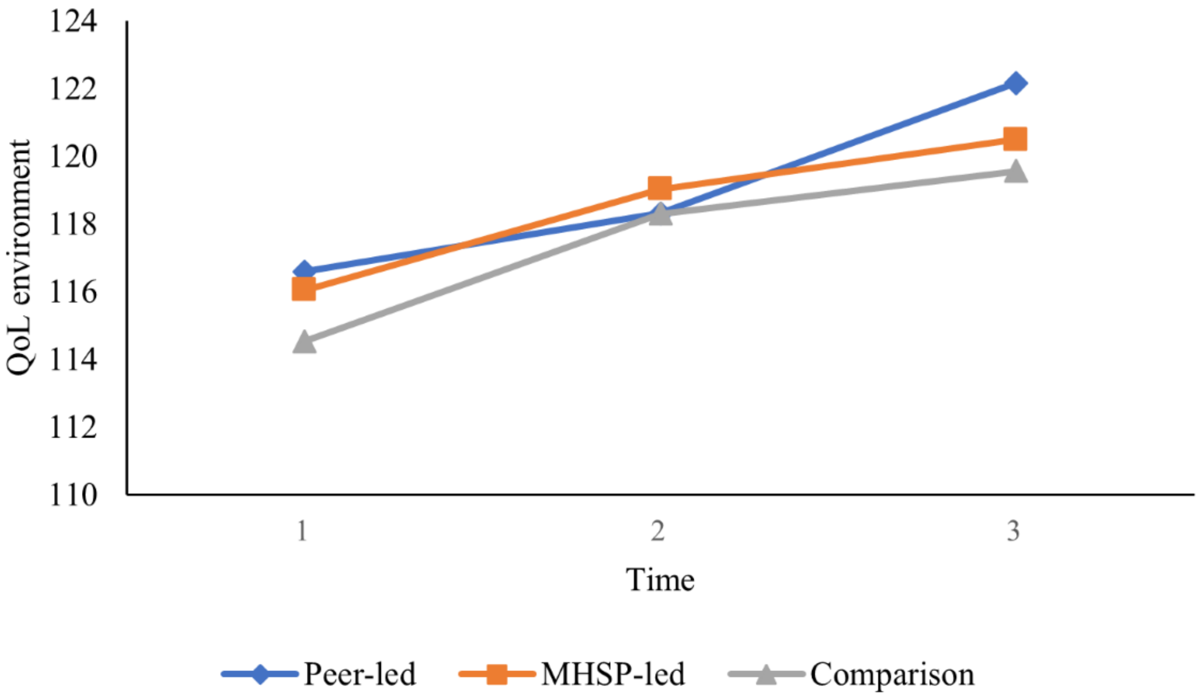


**Figure 6.** University students' reported social support over time. Main effect of time represents a significant difference between time 1 and time 2 as well as between time 2 and time 3. MHSP: mental health service provider; MSPSS: Multidimensional Scale of Perceived Social Support.

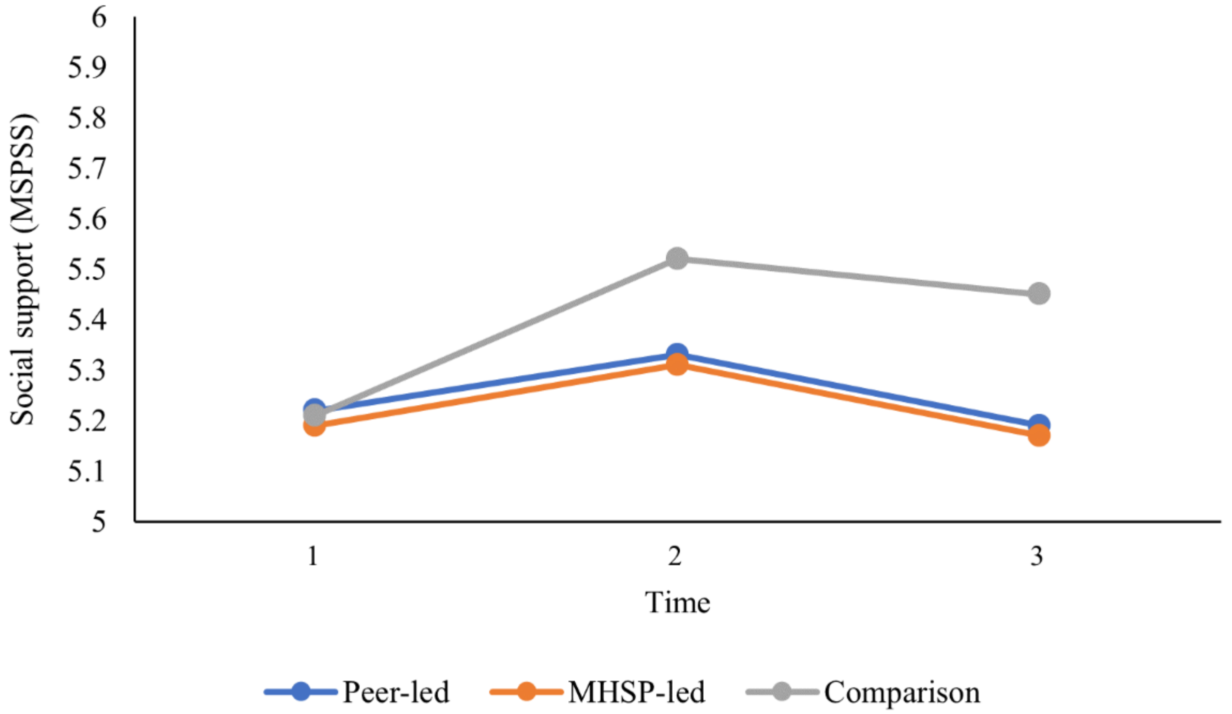
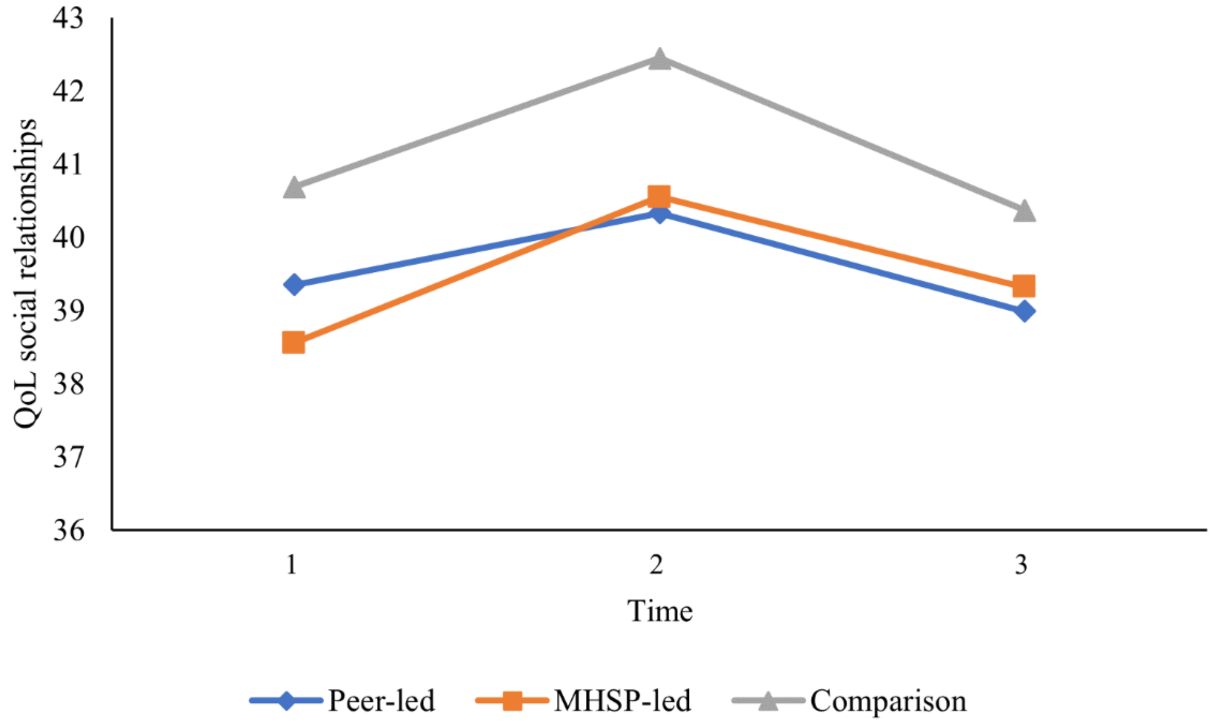


**Figure 7.** University students' reported quality of life (QoL; social relationships) over time. Main effect of time represents a significant difference between time 1 and time 2 as well as between time 2 and time 3. MHSP: mental health service provider.



## Discussion

### Principal Findings

The overarching goal of this study was to evaluate a mental health service provider–presented versus peer-presented mental health resilience skill–building web-based video outreach program against a wait-list comparison group. Specifically, the first objective sought to compare group differences on the web-based outreach program's acceptability between the types of deliverers (mental health service provider vs peer). Building on this, the second objective sought to compare group differences between the intervention groups (mental health service provider–presented and peer-presented) and a wait-list comparison group in terms of well-being outcomes (ie, decreased stress and increased coping self-efficacy, social support, social connectedness, mindfulness, and QoL) over 10 weeks.

Interestingly, the web-based outreach program received similarly high acceptability and satisfaction ratings regardless of whether the program deliverer was a mental health service provider or a peer. In both the mental health service provider–presented and peer-presented programs, most participants (58/72, 81% in the mental health service provider–presented group and 69/76, 91% in the peer-presented group) indicated that they were planning to use the program strategies in the future from *sometimes* to *frequently*. In addition, a large proportion of students in the mental health service provider–presented group (46/71, 65%) and in the peer-presented group (53/76, 70%) indicated that they felt that, after watching video 1, they had learned *a medium amount* to *a lot*. Thus, acknowledging the need to integrate cost-effective and easily accessible mental health programs to build mental health resilience capacity and support students in coping with general stress, these findings provide promising early evidence that a web-based skill-building resource for teaching mental health resilience skills is satisfactory and acceptable for university students. Furthermore, this may be a particularly valuable approach for students who have a preference for peer-led approaches or to complement professional mental health services. This is in line with previous literature reporting high satisfaction with web-based mental health skill-building programs [8,60]. However, to our knowledge, this is the first study to examine such acceptability of and satisfaction with a peer-presented versus mental health service provider–presented universal resilience-building web-based program in a sample of university students. Interestingly, given nonsignificant differences between the groups, the findings suggest that a resilience skill–building video outreach program may be acceptable for university students regardless of service delivery type (mental health service provider–presented or peer-presented).

Although most students reported high overall acceptability and satisfaction with the program, a certain proportion of students (eg, students saying that they were never to rarely planning to use the strategies in the future: 21/148, 14.2%) did indicate less-positive reports regarding the program. However, it is important to note that a random convenience sample was recruited, where students were generously compensated for their time, to avoid a self-selection bias; therefore, a broad sample of students who may or may not have had the need for or interest in these strategies participated in the study. Thus, it is not surprising that a number of the participants reported less interest in the program or willingness to use the strategies in the future as they may not have had a need for these strategies. These findings further confirm that mental health outreach will usually be of interest and relevance specifically to those who are currently feeling a need for this support.

Moreover, interpretation of this study's findings needs to be carried out with a particular focus on the societal context in which the program was delivered. Importantly, while the program was being delivered, a state of emergency because of the COVID-19 pandemic was declared in the province in which this study was conducted. This brought on significant changes for students, such as social distancing restrictions and changes related to remote learning. Thus, such elevated levels of reported acceptability and satisfaction are encouraging as the program was disseminated at the beginning of the COVID-19 pandemic, when there was much uncertainty and students' lifestyles were rapidly changing [44,61]. In addition, the elevated proportion of students who reported having used the program strategies to cope with COVID-19 stress suggests that these types of strategies are feasible to use in times of rapid change and uncertainty. However, considering the societal context, this may have played a role in the nonsignificant group differences between the satisfaction with the mental health service provider–presented program and the satisfaction with the peer-presented program. Given the challenges associated with the pandemic, students may have been eager to access web-based mental health resources regardless of who was delivering the program. Although the findings may have important implications for the development and integration of future outreach programs seeing this program's high acceptability and satisfaction, nonsignificant group differences should be interpreted with caution based on the context, and further investigation may be required.

The second objective was to compare group differences between a mental health service provider–presented versus peer-presented web-based skill-building video outreach program and a wait-list comparison group in terms of well-being outcomes over time (baseline, after the intervention, and follow-up). Although the students rated the program very positively, no difference was found among any of the 3 groups on any of the well-being outcomes over time. Thus, the intervention groups did not demonstrate a greater improvement over time in well-being outcomes relative to the wait-list comparison group, although, as discussed below, there was a general increase in well-being for all groups. Even though previous studies have found that web-based interventions are effective in supporting university students' well-being [62], there are several potential factors that may explain this lack of a detectable intervention benefit. It may be that this null finding indicates that the intervention was ineffective, meaning that perhaps the intervention was not optimized (eg, the time span was too short or the students did not engage with the strategies for a sufficient amount of time). As demonstrated in the literature, intervention dosage is crucial to an intervention [63]; as such, it may be that students need supplemental time to engage with the strategies and for strategy practice.

An alternative hypothesis for this null finding is that all 3 groups of participants may have had increased access to mental health support resources through the university and community given the plethora of web-based student mental health resources offered because of the pandemic. The elevated reports of stress associated with the COVID-19 pandemic and the need to move students to web-based learning resulted in a plethora of web-based resources offered through the university and community to effectively support students during this time. In summary, all 3 groups in this study would have had access to several mental health support resources through the university and community as well as a decrease in academic stressors.

Although no group differences were found, there was a significant change over time in coping self-efficacy, social support, mindfulness, and QoL (social relationships and environment domains) for all 3 groups. Specifically, students increased in coping self-efficacy and mindfulness from time 1 to time 2 and remained stable from time 2 to time 3. Similarly, students reported an increase in the quality of their environment (ie, QoL environment domain) between time 1 and time 3. These findings are in line with a study conducted by Hamza et al [64] demonstrating that students with preexisting mental health concerns reported an increase or similar levels of psychological well-being compared with a year before. Similar to the findings of Hamza et al [64], many students in this sample were already reporting mental health difficulties and may have been better able to cope with the changes associated with the COVID-19 pandemic, such as increased social isolation.

Moreover, mindfulness practice has been gaining popularity as an evidence-based strategy for managing stress. As a result, several means of support offered to deal with the pandemic stress were aimed at enhancing mindfulness [65,66], which may have influenced students' reported mindfulness during this time. Finally, the increase in the QoL environment domain has been hypothesized to have been affected by potential positive experiences of the COVID-19 pandemic, as noted in several recent COVID-19 studies [67,68]. The QoL environment domain measures facets such as "Opportunities for acquiring new information and skills" and "Participation in and opportunities for recreation" and may have increased as a result of the reduction of academic stressors and increase in time available for leisure activities.

Interestingly, the pattern of change for social support and the QoL social relationships domain was different, with an increase from time 1 to time 2 before returning to baseline levels at time 3. The increase in social support and the QoL social relationships domain is consistent with previous literature on natural or societal disasters, where there is an increase in social and general mental health support directly following these tragic events [69,70]. Thus, the time 2 increase in perceived social support is hypothesized to have been related to the early pandemic surge in social connections as families, peers, and communities reached out to individuals to ensure safety and well-being. This early pandemic increase in social support is similar to findings from a community sample of adults, where an increase in social support was reported at the start of the pandemic [71]. However, although interesting patterns emerged in terms of students' well-being during this time, the findings of this study are tentative, as several factors may have affected the results, and should be interpreted with caution.

## Limitations and Future Directions

This study is not without limitations. Considering that the program was disseminated during a time of change because of the onset of the COVID-19 pandemic, the results may have been different if the web-based outreach program had been provided to manage regular day-to-day stress. Thus, these results may not be generalizable to nonpandemic times, and future research would benefit from evaluating program effectiveness in a different context. Furthermore, given the importance of intervention dosage [63], future studies may want to examine specifically how much time students engage with strategy practice. Although this study asked students to retrospectively rate on a Likert scale from *never* to *frequently* how much they used the program strategies, it may be helpful to have students provide a daily report of the time spent engaging with each strategy to be able to better evaluate whether the dosage had an impact on the effectiveness of the intervention. A further limitation of this study is the use of a nonstandardized, researcher-designed measure to assess program satisfaction owing to the lack of relevant standardized measures in this area. Finally, there is limited generalizability of the findings owing to a sample where the participants predominantly identified as White women. Further research is needed to explore student responses to web-based mental health outreach programs among groups that may be underrepresented in our sample.

## Conclusions

The elevated levels of mental health distress reported by university students and the difficulties associated with the developmental period of emerging adulthood highlight the need to provide university students with appropriate mental health support. Although this study did not demonstrate program effectiveness, students reported high acceptability of both mental health service provider–presented and peer-presented programs. The findings highlight that the content presented (strategies for skill-building and psychoeducation) may play a more important role in students' acceptability than who is delivering the program. Hence, future initiatives may want to consider the involvement of peers in delivering similar web-based programs as an effective approach to address barriers to program dissemination, such as limited resources and increasing acceptability of students with a preference for peer approaches.

---

## Acknowledgments

This study was funded by the Canadian Institute of Health Research.

---

## Authors' Contributions

LB made substantial contributions to the study methodological design, data collection, data analysis, interpretation of the results, manuscript writing, and editing of the final manuscript. BNB, SZ, JM, and NLH made substantial contributions to the study conceptualization, design, data collection, interpretation of the results, and critical revision of manuscript drafts and approved the final version. LDG, VR, SPL, RW, and SNI made substantial contributions to the study conceptualization and design and critically revised manuscript drafts.

## Conflicts of Interest

None declared.

### Editorial Notice

This randomized study was only retrospectively registered, as the authors believed that the randomization approach used in this study would not fall under the umbrella of a randomized controlled trial. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials, because the risk of bias appears low and the authors had not intended to conduct a randomized controlled trial. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 9091 KB - mental_v9i7e34168_app1.pdf ]

## References

1. American College Health Association-National College Health Assessment II: Canadian Consortium Executive Summary Spring 2019. American College Health Association. Silver Spring, MD, USA: American College Health Association; 2019. URL: https://www.cacuss.ca/files/Research/NCHA-II%20SPRING%202019%20CANADIAN%20REFERENCE%20GROUP%20EXECUTIVE%20SUMMARY.pdf [accessed 2020-03-11]
2. Matthews M. A Review of Mental Health Services Offered by Canada's English language Universities. Western University. 2017. URL: http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=1&sid=08d8428e-c11a-49b0-b24e-906cfdc1f7e4%40sessionmgr4007 [accessed 2021-06-21]
3. Making the Case for Investing in Mental Health in Canada. Mental Health Commission of Canada. 2016. URL: https://www.mentalhealthcommission.ca/sites/default/files/2016-06/Investing_in_Mental_Health_FINAL_Version_ENG.pdf [accessed 2020-03-11]
4. Watkins DC, Hunt JB, Eisenberg D. Increased demand for mental health services on college campuses: perspectives from administrators. Qual Soc Work 2011 Aug 09;11(3):319-337. [doi: 10.1177/1473325011401468]
5. Ali K, Farrer L, Gulliver A, Griffiths KM. Online peer-to-peer support for young people with mental health problems: a systematic review. JMIR Ment Health 2015 May 19;2(2):e19 [FREE Full text] [doi: 10.2196/mental.4418] [Medline: 26543923]
6. Byrom N. An evaluation of a peer support intervention for student mental health. J Ment Health 2018 Jun;27(3):240-246. [doi: 10.1080/09638237.2018.1437605] [Medline: 29451411]
7. Cyr C, Mckee H, O'Hagan M, Priest R. Making the case for peer support. Mental Health Commission of Canada. 2016 Jul. URL: https://www.mentalhealthcommission.ca/wp-content/uploads/drupal/2016-07/MHCC_Making_the_Case_for_Peer_Support_2016_Eng.pdf [accessed 2020-04-21]
8. Fortuna KL, Naslund JA, LaCroix JM, Bianco CL, Brooks JM, Zisman-Ilani Y, et al. Digital peer support mental health interventions for people with a lived experience of a serious mental illness: systematic review. JMIR Ment Health 2020 Apr 03;7(4):e16460 [FREE Full text] [doi: 10.2196/16460] [Medline: 32243256]
9. Arnett JJ. Emerging adulthood. A theory of development from the late teens through the twenties. Am Psychol 2000 May;55(5):469-480. [Medline: 10842426]
10. Sussman S, Arnett JJ. Emerging adulthood: developmental period facilitative of the addictions. Eval Health Prof 2014 Jun;37(2):147-155. [doi: 10.1177/0163278714521812] [Medline: 24492245]
11. Arnett JJ. Emerging Adulthood: The Winding Road from the Late Teens Through the Twenties. Oxford, UK: Oxford University Press; 2014.
12. Arnett JJ, Žukauskienė R, Sugimura K. The new life stage of emerging adulthood at ages 18-29 years: implications for mental health. Lancet Psychiatry 2014 Dec;1(7):569-576. [doi: 10.1016/S2215-0366(14)00080-7] [Medline: 26361316]
13. Schiller M, Hammen CC, Shahar G. Links among the self, stress, and psychological distress during emerging adulthood: comparing three theoretical models. Self Identity 2016 Feb 23;15(3):302-326. [doi: 10.1080/15298868.2015.1131736]
14. Strengthening resilience: a priority shared by health 2020 and the and the sustainable development goals. World Health Organization. 2017. URL: https://www.euro.who.int/__data/assets/pdf_file/0005/351284/resilience-report-20171004-h1635.pdf [accessed 2022-05-27]
15. Regehr C, Glancy D, Pitts A. Interventions to reduce stress in university students: a review and meta-analysis. J Affect Disord 2013 May 15;148(1):1-11. [doi: 10.1016/j.jad.2012.11.026] [Medline: 23246209]

16. Biasi V, Patrizi N, Mosca M, De Vincenzo C. The effectiveness of university counselling for improving academic outcomes and well-being. Br J Guid Couns 2016 Dec 03;45(3):248-257. [doi: 10.1080/03069885.2016.1263826]
17. Vescovelli F, Melani P, Ruini C, Ricci Bitti PE, Monti F. University counseling service for improving students' mental health. Psychol Serv 2017 Nov;14(4):470-480. [doi: 10.1037/ser0000166] [Medline: 29120205]
18. Salganik MJ. Bit by Bit: Social Research in the Digital Age. Princeton, NJ, USA: Princeton University Press; 2019.
19. Chan JK, Farrer LM, Gulliver A, Bennett K, Griffiths KM. University students' views on the perceived benefits and drawbacks of seeking help for mental health problems on the internet: a qualitative study. JMIR Hum Factors 2016 Jan 19;3(1):e3 [FREE Full text] [doi: 10.2196/humanfactors.4765] [Medline: 27026140]
20. Irish M, Kuso S, Simek M, Zeiler M, Potterton R, Musiat P, et al. Online prevention programmes for university students: stakeholder perspectives from six European countries. Eur J Public Health 2021 Jul 07;31(31 Suppl 1):i64-i70 [FREE Full text] [doi: 10.1093/eurpub/ckab040] [Medline: 34240152]
21. Clarke AM, Kuosmanen T, Barry MM. A systematic review of online youth mental health promotion and prevention interventions. J Youth Adolesc 2015 Jan;44(1):90-113. [doi: 10.1007/s10964-014-0165-0] [Medline: 25115460]
22. Harrer M, Adam SH, Fleischmann RJ, Baumeister H, Auerbach R, Bruffaerts R, et al. Effectiveness of an Internet- and app-based intervention for college students with elevated stress: randomized controlled trial. J Med Internet Res 2018 Apr 23;20(4):e136 [FREE Full text] [doi: 10.2196/jmir.9293] [Medline: 29685870]
23. Garrido S, Millington C, Cheers D, Boydell K, Schubert E, Meade T, et al. What works and what doesn't work? A systematic review of digital mental health interventions for depression and anxiety in young people. Front Psychiatry 2019 Nov 13;10:759 [FREE Full text] [doi: 10.3389/fpsyt.2019.00759] [Medline: 31798468]
24. Lattie EG, Adkins EC, Winquist N, Stiles-Shields C, Wafford QE, Graham AK. Digital mental health interventions for depression, anxiety, and enhancement of psychological well-being among college students: systematic review. J Med Internet Res 2019 Jul 22;21(7):e12869 [FREE Full text] [doi: 10.2196/12869] [Medline: 31333198]
25. Repper J, Carter T. A review of the literature on peer support in mental health services. J Ment Health 2011 Aug;20(4):392-411. [doi: 10.3109/09638237.2011.583947] [Medline: 21770786]
26. Davidson L, Bellamy C, Guy K, Miller R. Peer support among persons with severe mental illnesses: a review of evidence and experience. World Psychiatry 2012 Jun;11(2):123-128 [FREE Full text] [doi: 10.1016/j.wpsyc.2012.05.009] [Medline: 22654945]
27. Goodwin J, Behan L, Kelly P, McCarthy K, Horgan A. Help-seeking behaviors and mental well-being of first year undergraduate university students. Psychiatry Res 2016 Dec 30;246:129-135. [doi: 10.1016/j.psychres.2016.09.015] [Medline: 27693865]
28. Lannin DG, Barrowclough M, Vogel DL. An examination of help-seeking preferences via best-worst scaling. J Clin Psychol 2020 Sep;76(9):1677-1695. [doi: 10.1002/jclp.22945] [Medline: 32077504]
29. Montagni I, Tzourio C, Cousin T, Sagara JA, Bada-Alonzi J, Horgan A. Mental health-related digital use by university students: a systematic review. Telemed J E Health 2020 Feb;26(2):131-146. [doi: 10.1089/tmj.2018.0316] [Medline: 30888256]
30. Nicholas J, Boydell K, Christensen H. mHealth in psychiatry: time for methodological change. Evid Based Ment Health 2016 May;19(2):33-34. [doi: 10.1136/eb-2015-102278] [Medline: 27044849]
31. Baum F, MacDougall C, Smith D. Participatory action research. J Epidemiol Community Health 2006 Oct;60(10):854-857 [FREE Full text] [doi: 10.1136/jech.2004.028662] [Medline: 16973531]
32. Macaulay AC, Commanda LE, Freeman WL, Gibson N, McCabe ML, Robbins CM, et al. Participatory research maximises community and lay involvement. North American Primary Care Research Group. BMJ 1999 Sep 18;319(7212):774-778 [FREE Full text] [doi: 10.1136/bmj.319.7212.774] [Medline: 10488012]
33. Bajaj B, Pande N. Mediating role of resilience in the impact of mindfulness on life satisfaction and affect as indices of subjective well-being. Pers Individ Dif 2016 Apr;93:63-67. [doi: 10.1016/j.paid.2015.09.005]
34. Cosco TD, Kaushal A, Richards M, Kuh D, Stafford M. Resilience measurement in later life: a systematic review and psychometric analysis. Health Qual Life Outcomes 2016 Jan 28;14:16 [FREE Full text] [doi: 10.1186/s12955-016-0418-6] [Medline: 26821587]
35. Polk LV. Toward a middle-range theory of resilience. ANS Adv Nurs Sci 1997 Mar;19(3):1-13. [doi: 10.1097/00012272-199703000-00002] [Medline: 9055026]
36. Sanderson B, Brewer M. What do we know about student resilience in health professional education? A scoping review of the literature. Nurse Educ Today 2017 Nov;58:65-71. [doi: 10.1016/j.nedt.2017.07.018] [Medline: 28898767]
37. Thompson G, McBride RB, Hosford CC, Halaas G. Resilience among medical students: the role of coping style and social support. Teach Learn Med 2016;28(2):174-182. [doi: 10.1080/10401334.2016.1146611] [Medline: 27064719]
38. Wilson CJ, Deane FP, Ciarrochi JV, Rickwood D. Measuring help seeking intentions: Properties of the General Help Seeking Questionnaire. University of Wollongong. 2005. URL: https://ro.uow.edu.au/cgi/viewcontent.cgi?article=2580&context=hbspapers [accessed 2020-03-11]
39. Altman DG, Bland JM. Treatment allocation by minimisation. BMJ 2005 Apr 09;330(7495):843 [FREE Full text] [doi: 10.1136/bmj.330.7495.843] [Medline: 15817555]

40. Scott NW, McPherson GC, Ramsay CR, Campbell MK. The method of minimization for allocation to clinical trials. A review. Control Clin Trials 2002 Dec;23(6):662-674. [doi: 10.1016/s0197-2456(02)00242-8] [Medline: 12505244]
41. Son C, Hegde S, Smith A, Wang X, Sasangohar F. Effects of COVID-19 on college students' mental health in the United States: interview survey study. J Med Internet Res 2020 Sep 03;22(9):e21279 [FREE Full text] [doi: 10.2196/21279] [Medline: 32805704]
42. Kirkpatrick JD, Kirkpatrick WK. Kirkpatrick's Four Levels of Training Evaluation. Alexandria, VA, USA: Association for Talent Development; 2016.
43. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
44. Cohen S. Perceived stress in a probability sample of the United States. In: Spacapan S, Oskamp S, editors. The Social Psychology of Health: The Claremont Symposium on Applied Social Psychology. New York, NY, USA: Sage Publications; 1988:31-67.
45. Roberti JW, Harrington LN, Storch EA. Further psychometric support for the 10-item version of the Perceived Stress Scale. J Coll Couns 2006;9(2):135-147. [doi: 10.1002/j.2161-1882.2006.tb00100.x]
46. Chesney AM, Neilands TB, Chambers DB, Taylor JM, Folkman S. A validity and reliability study of the coping self-efficacy scale. Br J Health Psychol 2006 Sep;11(Pt 3):421-437 [FREE Full text] [doi: 10.1348/135910705X53155] [Medline: 16870053]
47. Zimet GD, Dahlem NW, Zimet SG, Farley GK. The multidimensional scale of perceived social support. J Pers Assess 1988 Mar;52(1):30-41. [doi: 10.1207/s15327752jpa5201_2]
48. Zimet GD, Powell SS, Farley GK, Werkman S, Berkoff KA. Psychometric characteristics of the Multidimensional Scale of Perceived Social Support. J Pers Assess 1990;55(3-4):610-617. [doi: 10.1080/00223891.1990.9674095] [Medline: 2280326]
49. Lee RM, Draper M, Lee S. Social connectedness, dysfunctional interpersonal behaviors, and psychological distress: testing a mediator model. J Couns Psychol 2001;48(3):310-318. [doi: 10.1037/0022-0167.48.3.310]
50. Brown KW, Ryan RM. The benefits of being present: mindfulness and its role in psychological well-being. J Pers Soc Psychol 2003 Apr;84(4):822-848. [doi: 10.1037/0022-3514.84.4.822] [Medline: 12703651]
51. MacKillop J, Anderson EJ. Further psychometric validation of the Mindful Attention Awareness Scale (MAAS). J Psychopathol Behav Assess 2007 May 22;29(4):289-293. [doi: 10.1007/s10862-007-9045-1]
52. The WHOQOL Group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. The WHOQOL Group. Psychol Med 1998 May;28(3):551-558. [doi: 10.1017/s0033291798006667] [Medline: 9626712]
53. Vahedi S. World Health Organization Quality-of-Life Scale (WHOQOL-BREF): analyses of their item response theory properties based on the graded responses model. Iran J Psychiatry 2010;5(4):140-153 [FREE Full text] [Medline: 22952508]
54. Conley CS, Durlak JA, Kirsch AC. A meta-analysis of universal mental health prevention programs for higher education students. Prev Sci 2015 May;16(4):487-507. [doi: 10.1007/s11121-015-0543-1] [Medline: 25744536]
55. Rith-Najarian LR, Boustani MM, Chorpita BF. A systematic review of prevention programs targeting depression, anxiety, and stress in university students. J Affect Disord 2019 Oct 01;257:568-584. [doi: 10.1016/j.jad.2019.06.035] [Medline: 31326690]
56. Spijkerman MP, Pots WT, Bohlmeijer ET. Effectiveness of online mindfulness-based interventions in improving mental health: a review and meta-analysis of randomised controlled trials. Clin Psychol Rev 2016 Apr;45:102-114 [FREE Full text] [doi: 10.1016/j.cpr.2016.03.009] [Medline: 27111302]
57. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
58. Fall 2020 Admissions*. McGill University — Enrolment Services. 2020. URL: https://www.mcgill.ca/es/admissions-profile/fall-2020 [accessed 2022-05-27]
59. Tabachnick BG, Fidell LS. Using Multivariate Statistics. 6th edition. Boston, MA, USA: Pearson Education; 2007.
60. Rickwood D, Wallace A, Kennedy V, O'Sullivan S, Telford N, Leicester S. Young people's satisfaction with the online mental health service eheadspace: development and implementation of a service satisfaction measure. JMIR Ment Health 2019 Apr 17;6(4):e12169 [FREE Full text] [doi: 10.2196/12169] [Medline: 30994470]
61. Charles NE, Strong SJ, Burns LC, Bullerjahn MR, Serafine KM. Increased mood disorder symptoms, perceived stress, and alcohol use among college students during the COVID-19 pandemic. Psychiatry Res 2021 Feb;296:113706 [FREE Full text] [doi: 10.1016/j.psychres.2021.113706] [Medline: 33482422]
62. Amanvermez Y, Rahmadiana M, Karyotaki E, de Wit L, Ebert DD, Kessler RC, et al. Stress management interventions for college students: a systematic review and meta-analysis. Clin Psychol Sci Pract (forthcoming) 2020 May 22:e12342. [doi: 10.1111/cpsp.12342]
63. Conley CS, Durlak JA, Shapiro JB, Kirsch AC, Zahniser E. A meta-analysis of the impact of universal and indicated preventive technology-delivered interventions for higher education students. Prev Sci 2016 Aug;17(6):659-678. [doi: 10.1007/s11121-016-0662-3] [Medline: 27225631]

64. Hamza CA, Ewing L, Heath NL, Goldstein AL. When social isolation is nothing new: a longitudinal study on psychological distress during COVID-19 among university students with and without preexisting mental health concerns. Can Psychol 2021 Feb;62(1):20-30. [doi: 10.1037/cap0000255]
65. Antonova E, Schlosser K, Pandey R, Kumari V. Coping with COVID-19: mindfulness-based approaches for mitigating mental health crisis. Front Psychiatry 2021 Mar 23;12:563417 [FREE Full text] [doi: 10.3389/fpsyt.2021.563417] [Medline: 33833695]
66. Weis R, Ray SD, Cohen TA. Mindfulness as a way to cope with COVID‐19‐related stress and anxiety. Couns Psychother Res 2021 Mar;21(1):8-18. [doi: 10.1002/capr.12375]
67. Huckins JF, daSilva AW, Wang W, Hedlund E, Rogers C, Nepal SK, et al. Mental health and behavior of college students during the early phases of the COVID-19 pandemic: longitudinal smartphone and ecological momentary assessment study. J Med Internet Res 2020 Jun 17;22(6):e20185 [FREE Full text] [doi: 10.2196/20185] [Medline: 32519963]
68. Kaparounaki CK, Patsali ME, Mousa DP, Papadopoulou EV, Papadopoulou KK, Fountoulakis KN. University students' mental health amidst the COVID-19 quarantine in Greece. Psychiatry Res 2020 Aug;290:113111 [FREE Full text] [doi: 10.1016/j.psychres.2020.113111] [Medline: 32450416]
69. Karlin NJ, Marrow S, Weil J, Baum S, Spencer TS. Social support, mood, and resiliency following a Peruvian natural disaster. J Loss Trauma 2012 Sep;17(5):470-488. [doi: 10.1080/15325024.2012.665019]
70. Madsen W, O'Mullan C. Perceptions of community resilience after natural disaster in a rural Australian town. J Community Psychol 2016 Mar 09;44(3):277-292. [doi: 10.1002/jcop.21764]
71. Tull MT, Edmonds KA, Scamaldo KM, Richmond JR, Rose JP, Gratz KL. Psychological outcomes associated with stay-at-home orders and the perceived impact of COVID-19 on daily life. Psychiatry Res 2020 Jul;289:113098 [FREE Full text] [doi: 10.1016/j.psychres.2020.113098] [Medline: 32434092]

## Abbreviations

**CSE:** Coping Self-Efficacy Scale
**GHSQ:** General Help-Seeking Questionnaire
**MAAS:** Mindful Attention Awareness Scale
**MSPSS:** Multidimensional Scale of Perceived Social Support
**PSS:** Perceived Stress Scale
**QoL:** quality of life
**SCS-R:** Social Connectedness Scale-Revised
**WHOQOL-BREF:** World Health Organization Quality of Life Brief questionnaire

*Edited by J Torous; submitted 11.10.21; peer-reviewed by C Bedard, J Torous; comments to author 20.03.22; accepted 01.06.22; published 22.07.22.*

*Please cite as:*
*Bastien L, Boke BN, Mettler J, Zito S, Di Genova L, Romano V, Lewis SP, Whitley R, Iyer SN, Heath NL*
*Peer-Presented Versus Mental Health Service Provider–Presented Mental Health Outreach Programs for University Students: Randomized Controlled Trial*
*JMIR Ment Health 2022;9(7):e34168*
*URL: https://mental.jmir.org/2022/7/e34168*
*doi:10.2196/34168*
*PMID:35762935*

Original Paper

# Telehealth Autism Diagnostic Assessments With Children, Young People, and Adults: Qualitative Interview Study With England-Wide Multidisciplinary Health Professionals

Debbie Spain[1,2*], PhD; Gavin R Stewart[1*], PhD; David Mason[1*], MSc; Victoria Milner[1,2*], MSc; Bryony Fairhurst[2,3], DClinPsy; Janine Robinson[4], DClinPsy; Nicola Gillan[5], DClinPsy; Ian Ensum[5], DClinPsy; Eloise Stark[6], DPhil; Francesca Happe[1], PhD

[1]Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

[2]The National Psychology Clinic, London, United Kingdom

[3]Berrywood Hospital, Northampton Healthcare National Health Service Foundation Trust, Northampton, United Kingdom

[4]Cambridge Lifespan Autism Spectrum Service, Cambridgeshire and Peterborough National Health Service Foundation Trust, Cambridge, United Kingdom

[5]Bristol Autism Spectrum Service, Avon and Wiltshire Mental Health Partnership National Health Service Trust, Bristol, United Kingdom

[6]University of Oxford, Oxford, United Kingdom

*these authors contributed equally

**Corresponding Author:**
Debbie Spain, PhD
Institute of Psychiatry, Psychology & Neuroscience
King's College London
16 De Crespigny Park
Denmark Hill
London, SE5 8AF
United Kingdom
Phone: 44 207 848 5388
Email: debbie.spain@kcl.ac.uk

## Abstract

**Background:** Autism spectrum disorder (hereafter, autism) is a common neurodevelopmental condition. Core traits can range from subtle to severe and fluctuate depending on context. Individuals can present for diagnostic assessments during childhood or adulthood. However, waiting times for assessment are typically lengthy, and many individuals wait months or even years to be seen. Traditionally, there has been a lack of standardization between services regarding how many and which multidisciplinary health professionals are involved in the assessment and the methods (diagnostic tools) that are used. The COVID-19 pandemic has affected routine service provision because of stay-at-home mandates and social distancing guidelines. Autism diagnostic services have had to adapt, such as by switching from conducting assessments in person to doing these fully via telehealth (defined as the use of remote technologies for the provision of health care) or using blended in-person or telehealth methods.

**Objective:** This study explored health professionals' experiences of and perspectives about conducting telehealth autism diagnostic assessments, including barriers and facilitators to this, during the COVID-19 pandemic; potential telehealth training and supervision needs of health professionals; how the quality and effectiveness of telehealth autism diagnostic services can be enhanced; and experiences of delivering postdiagnostic support remotely.

**Methods:** A total of 45 health professionals, working in varied settings across England, participated in one-off, in-depth semistructured qualitative interviews. These were conducted via videoconferencing or telephone. Altogether, participants represented 7 professional disciplines (psychiatry, medicine, psychology, speech and language therapy, occupational therapy, nursing, and social work). The data were then analyzed thematically.

**Results:** Thematic analysis indicated the following 7 themes: practicalities of telehealth, telehealth autism diagnostic assessments, diagnostic conclusions, clinical considerations, postdiagnostic support, future ways of working, and health professionals' experiences and needs. Overall, telehealth autism diagnostic assessments were deemed by many participants to be convenient, flexible, and efficient for some patients, families, and health professionals. However, not all patients could be assessed in this way, for example, because of digital poverty, complex clinical presentation, or concerns about risk and safeguarding. Working remotely encouraged innovation, including the development of novel assessment measures. However, some participants expressed significant concerns about the validity and reliability of remotely assessing social communication conditions.

**Conclusions:** A shift to telehealth meant that autism diagnostic services remained operational during the COVID-19 pandemic. However, this method of working has potentially affected the parity of service, with people presenting with clinical complexity having to potentially wait longer to be seen or given a diagnostic opinion. There is also a lack of standardization in the provision of services. Further research should identify evidence-based ways of enhancing the timeliness, accessibility, and robustness of the autism diagnostic pathway, as well as the validity and reliability of telehealth methods.

***(JMIR Ment Health 2022;9(7):e37901)*** doi:10.2196/37901

**KEYWORDS**

autism; COVID-19 pandemic; autism diagnostic assessment; telehealth; health professionals; clinical supervision; training; COVID-19

# *Introduction*

## Background

Autism spectrum disorder (henceforth, autism) is a lifelong neurodevelopmental condition affecting 1% to 2% of the population [1]. Core autism traits include social communication differences (impairments), difficulties tolerating change and uncertainty, sensory sensitivities, and restricted or repetitive interests and behaviors [2]. Autism is a substantially heterogeneous condition. Traits may be subtle or severe, affecting functioning to varying degrees [2]. Some individuals are diagnosed early in life, for example, when parents or teachers notice difficulties. Conversely, many individuals are only seen for diagnostic assessment in adulthood, commonly but not exclusively at the point they are required to become more independent and autonomous [3]. There is also growing recognition that many autistic individuals are undiagnosed or remain misdiagnosed [4].

Autism diagnostic assessments have traditionally lacked standardization between services and settings. For example, data on clinical practice in the United States, Canada, New Zealand, and the United Kingdom indicate that there is variation in how many and which health professionals are involved in the diagnostic process, the semistructured or structured diagnostic tools that are used, from whom information is obtained apart from the person (eg, family and educators), and the topics that the person is asked about (eg, a sole focus on autism or wider themes that include mental health) [5-7]. This is important as the assessment process can influence outcomes (ie, what diagnostic conclusions are reached) [8] and, in turn, the service provision that patients and their families can access.

In England, health professionals are expected to follow the National Institute for Health and Care Excellence guidelines pertaining to autism diagnostic assessment [9,10]. Traditionally, assessments have been conducted in person, with very limited use of telehealth (defined as the use of remote technologies, including videoconferencing and the telephone, for the provision of health care). Irrespective of age, the National Institute for Health and Care Excellence recommends that the assessment comprises a minimum of three components: (1) a clinical interview or assessment with the person, (2) behavioral observation, and (3) a review of developmental history. Although the guidelines provide an indication of the types of semistructured or structured tools that may be useful, they do not mandate the use of one over the other, resulting in differences in practice [9,10].

The emergence of the COVID-19 pandemic in early 2020 substantially affected the provision of emergency, routine, and specialist clinical services. Stay-at-home mandates and social distancing measures meant that nonemergency services needed to adapt the standard ways of working [11,12]. Some autism diagnostic services temporarily shut down waiting lists and suspended direct clinical work for several months, exacerbating already lengthy waiting times [13,14]. However, overall, many services started conducting partial or complete autism diagnostic assessments using telehealth [13-16].

There is a precedent for conducting autism diagnostic assessments remotely [17,18]. For example, 10 studies conducted in the United States before 2020 examined the feasibility and acceptability of conducting assessments via telehealth rather than in person, with preliminary evidence of effectiveness and good interrater reliability (when comparing both methods) [17]. However, the pandemic context, including stay-at-home mandates and social distancing measures, has introduced additional complexities and considerations for clinical practice, such as the need to rapidly develop new systems and processes to facilitate telehealth appointments, the expectation that professionals will adopt new ways of working without formal training, and the need to make clinical decisions about eligibility or contraindications for telehealth in a clinical rather than a research sample. A recently published systematic review of studies investigating telehealth methods of autism assessment and interventions for autistic individuals, conducted before and during the pandemic, also found that this is feasible, effective, and reliable [18].

Since the start of the COVID-19 pandemic, a handful of studies, primarily conducted in the United States and Canada, have examined the feasibility and acceptability of telehealth autism diagnostic assessments or the perspectives of patients or health professionals [14-16,19-21]. Preliminary findings indicate that some professionals and services can find telehealth to be convenient, flexible, and satisfactory when working with individuals across the life span. However, consistent concerns have also been raised by some professionals, including difficulties with engaging patients and families, assessing subtleties in nonverbal social communication and performing risk assessment, limited confidence in reaching diagnostic conclusions, and wider challenges such as technological problems (digital poverty and poor internet connection) and environmental considerations (eg, lack of privacy during appointments). In addition, professionals have identified a paucity of diagnostic tools validated for use via telehealth [22] and more general uncertainty about the validity and reliability of remotely assessing a condition underpinned by social communication differences. Some have also noted the potential for a widening gulf in health care disparities, as factors such as digital poverty, clinical complexity, risk, and the need for interpreters may mean that services want to meet patients and their families in person, resulting in a longer waiting time for assessment. Taken together, the evidence to date suggests that telehealth has merit as an approach for assessing autism [17,18]. However, professionals also face challenges in practice as a direct result of this approach, which may directly affect confidence and clarity in reaching diagnostic conclusions and the resultant service provision available for patients and their families. Further investigation of professionals' experiences of conducting telehealth autism diagnostic assessments in other settings and contexts is warranted to better understand how service provision can be more suitably tailored during and beyond the COVID-19 pandemic. Moreover, understanding more about barriers and facilitators to telehealth autism diagnostic assessment may help inform future iterations of service provision, ideally incorporating input from the range of involved stakeholders (including patients and their families).

### Study Aims

This study aimed to investigate England-wide multidisciplinary team (MDT) health professionals' experiences of and perspectives about (1) conducting telehealth autism diagnostic assessments, including barriers and facilitators; (2) potential training and supervision needs of health professionals using telehealth; (3) how quality and effectiveness of telehealth autism diagnostic services can be enhanced; and (4) experiences of delivering postdiagnostic support.

## *Methods*

We report the methods and findings based on the Consolidated Criteria for Reporting Qualitative Research guidelines [23] (Multimedia Appendix 1 [23]).

### Research Team

The team included autistic and nonautistic researchers working clinically (primarily in adult autism services or mental health settings) or in autism research departments. All members of the team were invited to comment on the study design and methods, as well as interpret the findings and contribute to the write-up. Members of the autism community (autistic teenagers, adults, and a parent carer) were asked to comment on the study materials and findings and offer their perspectives on the implications arising from the research.

### Study Design

This study was informed by phenomenological principles and used a qualitative study design. MDT health professionals attended one-off semistructured interviews between March and June 2021.

### Ethics Approval

Ethics approval was obtained from King's College London (Research Ethics Committee reference MRA-20/21-22168). The participants provided informed consent, including for the dissemination of anonymized quotes.

### Participants

Participant inclusion criteria were MDT health professionals (eg, representing psychiatry and psychology) with experience in conducting autism diagnostic assessments or providing postdiagnostic support for children, adolescents, or adults in any setting and in England.

We used convenience and snowball recruitment methods via the authors' existing England-wide collaborations and networks, gatekeepers at health organizations and universities, word of mouth, and social media. Recruitment ceased once (1) the breadth of health professional disciplines involved in autism diagnostic assessments in England was represented and (2) data saturation (defined as no new themes emerging) was reached.

A total of 45 MDT health professionals from across England participated (Figure 1). This comprised most of the total number of potential participants who initially expressed interest in the study. A total of 7 professional disciplines were represented. Most participants worked at least part-time in the National Health Service (NHS). Their expertise in working with autistic individuals ranged from 6 months to 30 years (mean 12 years; Table 1).

**Figure 1.** Location of services within which participants were based.

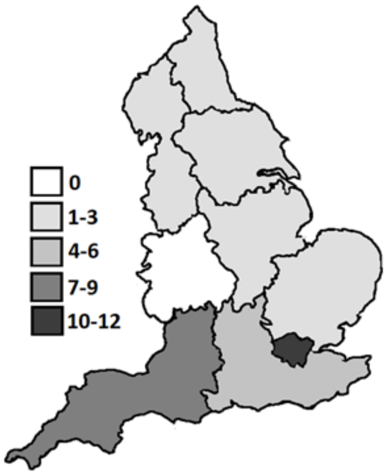


**Table 1.** Participants' professional demographic characteristics (N=45).

| Characteristics | Participants |
|---|---|
| **Profession[a], n (%)** | |
| Clinical psychologist | 13 (29) |
| Speech and language therapist | 6 (13) |
| Occupational therapist | 6 (13) |
| Psychiatrist | 5 (11) |
| Neurodevelopmental worker[b] | 5 (11) |
| Social worker | 3 (6) |
| Pediatrician | 2 (5) |
| Nurse | 2 (4) |
| Medical physician | 2 (4) |
| Counseling psychologist | 1 (2) |
| **Experience (years), mean (SD); range** | |
| Since core or primary professional training (n=29) | 13.32 (7.11); 1-23 |
| Working with autistic individuals | 12.14 (8.53); 0.5-30 |
| **Service context, n (%)** | |
| NHS[c] | 32 (70) |
| Private | 13 (30) |
| Accepts referrals for people with a learning disability | 8 (17) |
| Digital health provider | 5 (11) |
| **Age of patient group (n=35), n (%)** | |
| Child (<18 years) | 13 (37) |
| Adult (≥18 years) | 17 (49) |
| Life span | 5 (14) |

[a] Participants could endorse >1 professional discipline.

[b] Unqualified practitioner specializing in administering semistructured diagnostic tools as part of the multidisciplinary team assessment.

[c] NHS: National Health Service.

## Materials

The topic guide informed the interviews. This was developed in collaboration with experts with experience and health professionals. Briefly, the topic guide included (1) demographic questions, (2) contextual questions about participants' service context and experience of using telehealth or hybrid assessment, (3) prompts about views on telehealth, (4) perceived telehealth training and clinical supervision needs, and (5) thoughts about improving service provision during and beyond the COVID-19 pandemic. Multimedia Appendix 1 provides more information regarding the topic guide.

## Procedure

Interviews were conducted by 3 female researchers (DS, BF, and VM)—1 postdoctoral nurse, 1 clinical psychologist, and 1 doctoral student—all of whom had experience in conducting qualitative research and autism diagnostic assessments in varied settings (including inpatient and community settings, the criminal justice system, and research studies). Interviews were conducted via videoconferencing, as well as, infrequently, by telephone, at the time of participants' choice. Of the 45 participants, 6 (13%) had prior working relationships with their interviewers. This was acknowledged but not considered to impede participants' responses, as in all but one instance, individuals were not routinely working together at the time of study participation. The participants were aware that the study focused on the clinical practice and research interests of the team.

During the interviews, participants were asked questions based on topic guide prompts, allowing them to lead the conversation. The mean duration of the interviews was 46 (range 20-73) minutes. Interviews were recorded (audio and video), excluding 2 instances because of technical issues. The participants were not asked to specify their location at the time of the interview, although most appeared to be at work or home. We did not ask whether there were others in the vicinity during study participation; however, there were no obvious interruptions. Interviewers took hand notes and met intermittently during the study setup and recruitment for peer reflection on the interview process and content and to ensure that they each met participants with different backgrounds (ie, working with young people or adults, in the NHS or independently, and from varied professional disciplines).

## Data Analysis

Data pertaining to participants' professional demographic characteristics and descriptions of service-related factors were summarized descriptively. Qualitative data were analyzed thematically, involving (1) becoming familiar with the data, (2) generating initial codes, (3) searching for themes, (4) reviewing tentative themes, (5) labeling themes, and (6) summarizing the data [24].

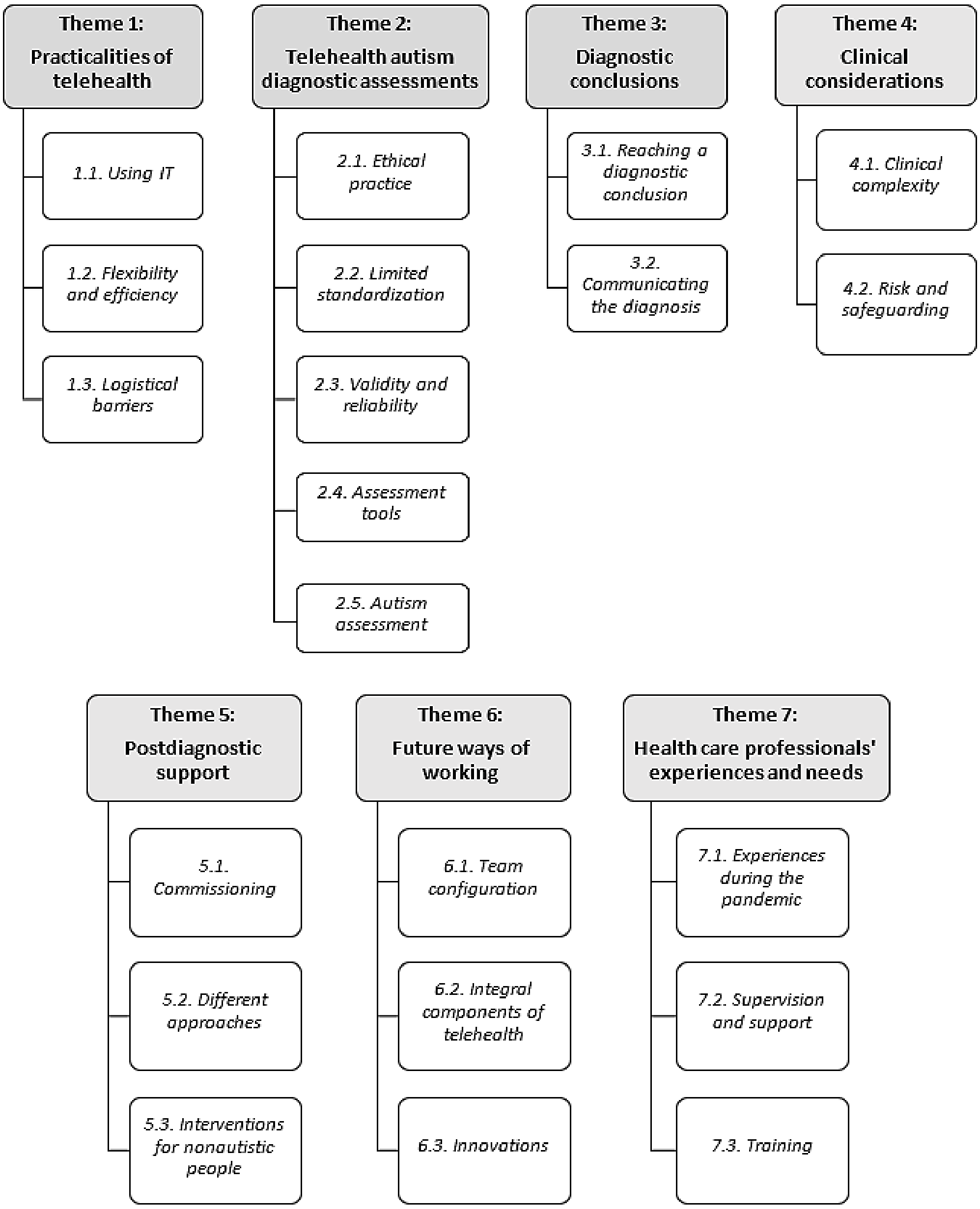
Reflexivity was embedded in the research process and during the thematic analysis [25]. This included members of the research team reflecting on perceptions of how autism can and should be assessed; minimum standards for autism assessment; views about the utility, validity, and reliability of telehealth; and experiences of receiving or providing autism and nonautism diagnoses in person and remotely.

The interviewers transcribed their own interviews. Transcripts were not sent to the participants for comments or checking. These were collated into a master document organized according to the question topic. One of the researchers (DS) created initial codes that were subsequently refined through categorization, with labels assigned to tentative themes and subthemes as they were identified. To ensure consistency within the coding, 2 researchers (VM and BF) randomly selected 10% of the interview responses at random and coded them. The codes were then compared among all 3 interviewers, with a high degree of comparability. Tentative themes were then finalized and presented to the wider research team for their comments.

# *Results*

Thematic analysis of the data indicated there were seven themes (1) practicalities of telehealth, (2) telehealth autism diagnostic assessments, (3) diagnostic conclusions, (4) clinical considerations, (5) postdiagnostic support, (6) future ways of working, and (7) health professionals' experiences and needs. Figure 2 presents the themes and subthemes (Multimedia Appendices 2-8).

**Figure 2.** Overview of themes and subthemes.



## Theme 1: Practicalities of Telehealth

The first theme pertained to participants' views on the practicalities of using telehealth, with three subthemes: (1) using IT, (2) flexibility and efficiency, and (3) logistical barriers.

### *Using IT*

*Digital poverty* was an issue for some participants. One of the participants said the following:

> *...privileged people can access a lot better and get a much more robust kind of assessment, because it's not constantly losing connection all the time. So that's a real concern to me.*

Professionals working in teams were said to have differing *computer literacy levels* and access to technology or devices. Some patients and families were described as being accustomed to IT, whereas others could *struggle with technology*. The lack of familiarity with this could be *an independent source of anxiety*. One participant remarked they "have one laptop to read, write and call from...the IT is not enough...we need bigger [wider], and more screens," specifically, one to use for clinical interactions and a second for reading and writing notes.

The use of videoconferencing platforms also differs. For example, Zoom videoconference was permitted only in some NHS Health Trusts. One of the digital health services used a custom-built platform. Some participants noted complexities associated with not being able to blur the screen background:

> *[this] means that people might know more about your personal life than any of us might share*

The visible contrast between some participants' and patients' home environments could be stark:

> *...there might be things in the background that are a bit distracting for somebody*

These factors may have influenced engagement and rapport building between patients and professionals, as well as patients' attention during an assessment.

### *Flexibility and Efficiency*

The consensus was that telehealth "gives us flexibility and choices." Together, participants said that using telehealth could result in (1) fluid appointment times, (2) more accessible appointments, (3) fewer no-shows, (4) options for swiftly filling last-minute cancellations, (5) the possibility for audio and video recording of assessments, (6) less travel and minimal expenditure (eg, on travel or parking), (7) fewer room bookings, (8) environmental benefits or lower carbon footprint, and (9) capacity for recruiting staff living outside of the area. Consequently, many participants felt that flexibility in telehealth could benefit all stakeholders.

### *Logistical Barriers*

Participants described a range of logistical issues related to patient circumstances that could potentially influence the viability, practical implementation, and success of a telehealth autism assessment (Textbox 1).

**Textbox 1.** Logistical issues affecting telehealth autism diagnostic assessments.

- Environmental factors
    - Location not optimal for an assessment (eg, nowhere suitable to sit at home, dialing into the appointment from work, and walking or riding a bicycle while doing the assessment)
    - Domestic situations not optimal for an assessment (eg, lack of privacy and caring for young children during the assessment)
    - Poor lighting or curtains closed
    - Poor sound or much background noise
    - Distracted by extraneous cues, finding it hard to sustain attention, and experiencing difficulty in sitting still
- IT-related factors
    - Issues relating to the camera (eg, height and position of the camera, proximity to the patient, and declining to turn the camera on or camera turned off unexpectedly)
    - Battery of device running out of charge unexpectedly
    - Only possible to see what is in front of the camera and not behind or at the periphery
- Engagement-related factors
    - Can feel intrusive to speak to someone while at home
    - Patient may choose not to join the assessment or wander off part of the way through
    - Displaying behavior that seems inappropriate for the context (eg, patient or family members not fully clothed, disappearing to make a sandwich or go for a walk, and answering the telephone)
- Risk-related factors
    - Domestic abuse
    - Safeguarding issues

## Theme 2: Telehealth Autism Diagnostic Assessments

The second theme pertained to the views of the participants when using telehealth for autism diagnostic assessments, with five subthemes related to (1) ethical practice, (2) limited standardization, (3) validity and reliability, (4) assessment tools, and (5) autism assessment.

### *Ethical Considerations*

Several participants stated that it was unethical for patients to wait longer than necessary. Thus, telehealth was a reasonable option, given the COVID-19 pandemic. Conversely, another participant highlighted that some patients cannot be seen using telehealth (eg, as they do not have IT or their clinical presentation precludes this [see the *Theme 4: Clinical Considerations* section]), and thus, "there's a bit of an ethical dilemma there, because obviously they've lost their place on our waiting list."

It was also reported that a purely remote assessment may contravene *ethical professional standards*. As autism is a social communication condition, not meeting a patient in person may mean that naturalistic interactions cannot be adequately assessed.

### *Limited Standardization*

Autism assessments lacked standardization, as highlighted by one of the participants, who said that "everybody's making their best guess at what might work." Services differed in terms of the (1) number of health professionals involved, (2) depth of information obtained, (3) range of sources from which information was gleaned, (4) types of behavioral observation assessments used, (5) setup of in-person appointments when offered, (6) total number of appointments offered (including feedback), and (7) overall duration of the assessment.

### Validity and Reliability

Participants' views differed regarding whether this method of diagnostic assessment was valid and reliable. Reflecting the views of many, as well as a change in usual practice since March 2020, one of the participants said the following:

> *I've been really surprised as how useful it is...[before the pandemic], I thought it would be a really bad idea and it wouldn't be valid, and it would be very limited, not reliable...now I've really shifted.*

However, this was commonly caveated with *curiosity* and, more specifically, a worry, about "how valid and reliable it is," especially the behavioral observation components of the assessment (see the *Assessment Tools* subsection). Conversely, others said the following:

> *...feel so strongly about it that it's not valid...if the full assessment is done remotely, it's not clinically valid*
>
> *I couldn't in all conscience assign a diagnosis [about] something as profound as how you interact socially with another human being having never sat in a room with them.*

Some participants thought telehealth assessments "do work very well, but there are always going to be [patients] when they're not going to be sufficient."

The reliability of telehealth assessments could be dependent on the age of the patient, such as being less appropriate for younger children. Others have suggested that this is less reliable for people with "definite speech and language difficulties, with intellectual [learning] disability, learning difficulties such as dyslexia, dyspraxia," or parents with a learning disability, who may find this a more overwhelming, ambiguous, or confusing meeting.

### Assessment Tools

Obtaining a developmental history, such as with the Autism Diagnostic Interview–Revised [26], was considered easy via telehealth, and indeed, this commonly occurred before the COVID-19 pandemic. Formal behavioral observation assessments, such as the Autism Diagnostic Observation Schedule (ADOS)–2 [27], translated less well to web-based forums. Some services decided to complete the ADOS-2 when social distancing measures were no longer in place, resulting in patients being placed on an internal waiting list.

Many services demonstrated innovation and developed an *ADOS-informed assessment*, comprising play-based and conversational tasks. Participants found this beneficial for structuring an appraisal of behaviors suggestive of autism. However, it was noted that these assessments had not been empirically tested, and thus, their psychometric properties (eg, test-retest reliability and interrater reliability) were unknown. On reflection, one of the participants felt that the ADOS-informed assessment they were using "is slightly limiting. We've done the best we can." Another said they had adopted "a really low threshold for review when we weren't certain...we feel that we may be missing things." Others described their newly developed assessment as "really successful, and I think it's been amazing that there's certain things you can pick up doing it."

### Autism Assessment

Opinions on conducting telehealth autism assessments varied. Some participants said that, with practice and experience, this did not differ substantially from assessing someone in person:

> *...we've certainly adjusted to it and for a significant majority of people, doing online assessment has been absolutely fine, and I think the diagnostic conclusion we've made has been the same as to whether we'd seen them in a room or not*

Another suggested the following:

> *...a difficult case is a difficult case, and a straightforward case is a straightforward case...I'm not sure that meeting somebody in person would have made a big difference*

Some characteristics prototypically associated with autism could be challenging to observe via telehealth, summarized as "you lose a lot of the subtleties...lose out on the interaction." These included (1) nonverbal behavior (eg, eye contact and quality, flexibility, range, congruence, and integration of facial expressions and gestures); (2) fluidity, responsivity, and reciprocity of social interaction with familiar and unfamiliar others; (3) hypo- and hypersensory sensitivities (eg, to light or noise); (4) repetitive movements and mannerisms, especially those outside the camera shot; and (5) gait and posture.

In addition, it could be difficult to assess coping strategies patients use in their day-to-day lives to manage difficulties or traits:

> *...you might not see that the curtains are drawn, or you might not see that there's particular lighting that they need*

The medium of telehealth could affect judgment about why a trait or behavior was observed. Echoing others' comments, one of the participants said the following:

> *...how much of that [social interaction difficulties] is a deficit on their part, and how much of it is just because there might be a slight delay in the internet? Or there might have been a break in the connection. So, it can be complicated to figure out whether their difficulty with reciprocity is because of that, or whether it's a typical issue.*

With younger children, there was a specific concern because of the following:

> *...they've not really had much social interaction over the last year [2020], and then you're trying to discern whether that's a COVID thing, or whether that's related to how they prefer things to be anyway*

Seeing patients' home environments could help with finding out about their preferences and difficulties:

> *I like that the person can show you things in their home. So, if you're asking somebody about collections, they can then show you that collection, or if they if you're asking about organisation, you know they can show you things that they've organised and so you get that sort of evidence and insight that you wouldn't get by bringing somebody to a clinic.*"

It was also easier to "see family dynamics" and "parent child interactions...like mum putting a hand on the child's shoulder...little things that actually show you a little bit about what their relationship is like." In contrast, patients could access their "favorite toys" at home; more easily rely on a "scripted, rehearsed kind of story"; and thereby, manage some interactions and ADOS-informed tasks more adeptly.

The remote assessment of domains other than autism could pose challenges. Several participants noted that there is "no possibility of doing a physical examination, not even just blood pressure and pulse, or you know if you thought someone would benefit from blood tests." Alongside this, many participants described difficulties in assessing the nonspecific elements of social interaction informing diagnostic conclusions:

> *...do they hold the door open for the informant, then let it go in their face*
>
> *...who sits next to who*
>
> *...how do they greet me...how do they sit*
>
> *...how they cope by coming to a clinic*
>
> *...if someone's trousers are stained, or if someone smells, like you're not getting that information about self-care and things that they might be struggling with and they don't always have the insight to be able to give you that information verbally*

Some participants said it was more difficult to develop a rapport on the web:

> *[it's] nice and tangible sitting in a room and there's some natural toys, and let's do this task together and let's work on it together*

A few participants wondered the following:

> *...people find it easier to sort of spin you a mistruth, when people want a diagnosis...I've got a lady at the moment that I don't know if she has autism or not, but she's giving me a lot of conflicting information*

## Theme 3: Diagnostic Conclusions

The third theme pertained to formulating and sharing diagnostic conclusions with two subthemes: (1) reaching a diagnostic conclusion and (2) communicating the diagnosis.

### *Reaching a Diagnostic Conclusion*

As for reaching a diagnostic conclusion, one of the participants said the following:

> *I think a lot of it's to do with the experience. The more you see people with different types of ASD and different presentations, and it takes a long time, but you find patterns in things and in people's behaviours, so you know you can read [about] it as much as you want or go on as much training courses that you can, but you never quite get it until you've been working with individuals for a long period of time*

Others felt that telehealth assessments introduced greater "uncertainty...we spend longer discussing cases," and the following:

> *I think it [telehealth] makes it much harder as a clinician to be sure of the diagnosis...you can't rely on your feeling and your responses because you're just listening to what they're saying*

There was a sense of *complexity* and difficult diagnostic decisions, and more difficulties for *newly qualified* health professionals. Several participants said they would not "confirm a diagnosis with anyone that I have not seen in person."

A few services had adopted an open-door policy, with one participant describing the following:

> *...assessment is limited by the set up [telehealth]...we generally say we would be happy to review in two years if problems persist. So, if we haven't given a diagnosis, we're leaving the door open that we may have missed it*

Moreover, their service "made recommendations based on the young person's current needs and situations, so we might give autism-related recommendations even without the diagnosis."

### *Communicating the Diagnosis*

In some services, feedback was provided in person, resulting in a lag between assessment and diagnosis. Some participants felt the following:

> *...quite callous and not particularly warm and friendly to be doing it over [the internet], like you're giving someone a life changing diagnosis and you can't even offer them a cup of tea while you're doing it or something. You know there's nothing to kind of soften the blow*

Another said the following:

> *...it's difficult if they are very emotional—you can say warm empathic things, but you can't hand them the tissue box...you feel a bit inadequate*

Some noted that giving a diagnosis jointly with colleagues seemed easier than giving a diagnosis just as the sole health professional.

The patients' experiences of receiving a diagnosis were important. It was difficult to know "whether it feels better for them to be in their own space and try to process that, or whether it's better to be in a clinic room." Some patients were said to "underestimate what the impact of a diagnosis might be like for them." One participant highlighted the following:

> *...there's a bit [of a] difference about you sitting in your bedroom and somebody giving you some news and then hanging up and you're still kind of sitting in your bedroom, versus coming to a room, somebody telling you something, you've been given kind of that time in the room, and then leaving the place where*

> *you'll be given the outcome to travel somewhere different*

Others said receiving a diagnosis could be a relief for patients; however, this was communicated, such as "it explains my past. You know, I've got a different narrative now." Overall, ensuring that patients have "the right emotional support around them" was deemed crucial.

Not receiving an autism diagnosis could incur *frustration and sadness*. Participants reported that relaying this in person or via telehealth could be difficult. One of the participants said their service goes "that extra mile" if a patient does not receive a diagnosis, as "you've got to do that in a way that doesn't [seem] over rejecting...like a huge disappointment." A few participants dealt with formal complaints whereby parents had said that a diagnosis of autism was not made as the assessment had been conducted via telehealth.

## Theme 4: Clinical Considerations

The fourth theme pertained to clinical considerations associated with the feasibility of using telehealth in their service, with two subthemes: (1) clinical complexity and (2) risk and safeguarding.

### *Clinical Complexity*

Participants said that referrals were increasingly being received for "more and more complex cases." Examples of complex cases may include patients presenting with limited verbal communication or selective mutism, mental health conditions, enduring personality traits or personality disorders, attachment-based problems, complex trauma, *looked-after* child status, fetal alcohol syndrome, sensory processing disorders, multimorbidity, or a forensic history. Some participants felt that the COVID-19 pandemic resulted in *a 2-tiered system*, with patients with more straightforward presentations being seen via telehealth versus patients with more complexity possibly waiting for longer.

Clinical complexity typically meant that the assessment was "more of a challenge":

> *...because we need to have more discussion and the MDT process becomes more lengthy, because you have got more to consider*

Participants found this could make it "really hard to tell whether they're [patients] autistic or not autistic, and you go away kind of thinking, well after 10 years, I should be able to know whether someone's autistic or not. It's very rare that I can't reach a conclusion [in person], but it seems to be far more complicated [via telehealth]."

### *Risk and Safeguarding*

In some services, moderate to high risk to self or others, recent suicidality, substance use, *high mental health needs*, impaired capacity, and known safeguarding concerns precluded the offer of a telehealth assessment.

Participants identified a range of risks inherent in clinical work, including to self, to others, and from others. However, the current pandemic context potentially increased the risk for some people, such as "from the fact that you're doing [the] assessment remotely." For instance, the following was more crucial:

> *...know where somebody is when you're speaking to them [as] they may not be at home...if there were kind of risk issues that came up, it will be important to know where they were.*

Another participant said that *risky topics* could arise when someone "doesn't want to disclose the ASD assessment to their family or partner or their children." Providing feedback for a diagnostic conclusion that patients are not happy with could also feel risky, especially in the absence of good rapport developed in person. Several participants highlighted that there may be an increased "risk of getting it [the diagnosis] wrong" and incurring "false positives and false negatives."

Several participants expressed uncertainty about whether the risk can be accurately gauged remotely, with some feeling "it can be quite difficult to hold that risk remotely." This was deemed especially tricky, as "there isn't anyone else that's going to come and pick up and monitor that risk." Another said the following:

> *...assessing high risk patients...[such as]...someone who's very psychotic... creates a bit more anxiety rather than being with the person in the same room and kind of getting a sense of the situation."*

Alongside this, it was noted that risk assessment and management is core work for some professional disciplines (eg, psychiatry, clinical psychology, and nursing); however, there may be less emphasis on this in the training of other disciplines:

> *...there's extra training to try and bring everybody up to that standard, which is really good, but then sometimes there are still gaps in people's knowledge and experience*

It was also apparent that a few services declined referrals for patients deemed to present any risk, again highlighting the potential disparities.

Seeing patients in their homes, via the web, could raise unexpected safeguarding concerns. One of the participants said the following regarding the period of a break:

> *...parents forgot to turn off the camera and volume, and they [the professional] heard inappropriate things where they shouted at their children...it made them feel uncomfortable and they filed a safeguarding concern*

Another participant identified that talking about safeguarding could potentially increase the risk of further safeguarding issues; for example, when assessing someone in an "abusive...coercive relationship" seen in the company of the abuser.

Participants talked about the complexity of dealing with safeguarding issues from their own homes:

> *...there's just something about being in a clinic environment where you know you almost kind of have your like safeguarding hat on more. I think because you're kind of in a role, whereas when you're at home, sometimes say you know you hear something or you even see something in the background, and I get a moment where I think gosh, this is actually really, you know important...sometimes that's difficult*

> *and not having just that constant kind of liaison with your colleagues is really hard*

## Theme 5: Postdiagnostic Support

The fifth theme pertained to the participants' views on postdiagnostic support and how their service currently handles this, with three subthemes: (1) commissioning, (2) different approaches, and (3) interventions for nonautistic individuals.

### Commissioning

Many services were "not commissioned to provide any postdiagnostic support," although this was described as follows:

> *...crucial, because we've got lots and lots and lots of children and adults who are being diagnosed with autism. But then, [they ask], what now? Where do I go with this? How can I make this useful?*

Another participant emphasized the following:

> *...they [commissioners, managers] sometimes lack the understanding that it is much more than a diagnosis or not. It's about being able to come away knowing that you feel that you've got a pretty good understanding of that child to not only feel comfortable making the diagnostic decision that you made, but also that you've been able to do something helpful for families*

The consensus was that services should be better resourced to provide input after the assessment.

### Different Approaches

The nature of postdiagnostic interventions differed between services, ranging from no intervention; signposting; resource leaflets; in-depth assessments of functioning; psychoeducation workshops and groups for patients, families, or friends; regular drop-in sessions; and, infrequently, individual sessions.

Some services had moved groups to the web, with varying degrees of success. One of the participants described their group now "feels much more like a teaching session...most of the clients don't want the camera on… so you can feel you could be speaking into the empty [void]." Others considered the move to web-based groups to have "been more successful than I thought it would be"—a valuable asset for patients who may have opted out of or been unable to attend in person. Attending the group on the web also meant that "you don't have to talk, but you can listen," reducing potential pressure on patients.

### Interventions for Nonautistic Individuals

The lack of a postdiagnostic intervention for people who do not receive a diagnosis of autism was mentioned:

> *...if you don't have a diagnosis of autism...this is a big issue. Too many autism services just dump them*

In one of the services, importance was placed on parity of understanding irrespective of diagnosis:

> *...you still get all of that same process. You still get the formulation. You still get told you will still get a differential diagnosis and opinion and we will still make recommendations for you. So, no matter where you are, autistic or not, you come up with the full assessment and what's deemed to be your diagnosis, but also what's deemed to be a formulation, so that if you do have to go into other services, you can take that with you, not have to answer the same questions again*

## Theme 6: Future Ways of Working

The sixth theme pertained to participants' thoughts about optimal service provision, with three subthemes: (1) team configuration, (2) integral components of telehealth, and (3) innovations.

### Team Configuration

There was wide variation in workforce configurations. Few participants worked as sole practitioners. Most teams had between 2 and ≥6 professional disciplines represented or available to participate in assessments ad hoc.

Echoing many participants' sentiments, one of the participants said the following:

> *I don't think they [health professionals] need to be from a particular professional background. What's more important is that they have adequate experience and training and confidence in differential diagnosis across a range of mental conditions and a range of neurodevelopmental conditions and that they know the [care] pathways, whether that's in the private sector or the NHS, you know, to refer people on for follow-up assessments and follow up treatment*

### Integral Components of Telehealth

Of the 45 participants, 5 (11% of the sample) worked for a digital health service. Of the remaining participants, most perceived services will continue to use telehealth beyond the COVID-19 pandemic. Whether this would be augmented with at least one in-person appointment depended on factors such as (1) organizational policy, (2) patient choice, (3) clinical complexity, (4) potential risk and safeguarding issues, (5) health professionals' preferences, and (6) environmental considerations (eg, whether there is somewhere quiet and confidential that patients and health professionals can use in their own homes or work areas). Reflecting many others' perspectives, one of the participants noted they are "happy to advocate a hybrid model, as long as the hybrid model is being hybrid to increase capacity without losing quality."

Participants outlined the fundamental elements they considered necessary to ensure good quality of telehealth autism assessments. This included suggestions for what the assessment comprises, how it is offered, who conducts it, health professionals' proficiency and ongoing supervision and training needs, and robust processes for service delivery (Textbox 2).

**Textbox 2.** Fundamental elements of telehealth autism diagnostic assessments.

- Professional resources
    - Remote assessment tools that are evidence based, standardized, and validated
    - A computer that has a reliable internet connection
    - Robust IT systems, prompt support with IT problems, and clear IT policies
    - Excellent admin support and tight admin processes
- Service design and processes
    - Collaboration and coproduction of service design and delivery with patients and families
    - Blended or hybrid model of service delivery, incorporating remote and in-person options based on needs and preferences
    - Differentiated pathways and options for straightforward and complex assessments
    - Allocation of patients to a professional within the team to offer continuity from referral to discharge
    - Clear procedures for assessing and managing risk or safeguarding concerning and a mechanism for obtaining urgent clinical advice
    - Options to conduct a neurodevelopmental assessment
- Team working and supervision
    - Input from a range of multidisciplinary team professionals
    - Options for joint working with colleagues if clinically indicated
    - High-quality clinical supervision
    - Opportunities to obtain peer support and build consensus about good practice in telehealth with colleagues at wider services
- Training
    - Adequately trained professionals with expertise in autism and mental health
    - Guidance on what a good practice telehealth assessment should incorporate and minimum standards for this
- Patient-friendly
    - Culturally aware service provision
    - Resources for patients (eg, visual information about the assessment process, overview of telehealth etiquette, and computer and internet access)
    - Accessible clinical reports and options for patients to comment on a draft report and discuss the final report
    - Postdiagnostic support via varied means (eg, written resources, in-depth assessment of functioning, individual sessions, and group support)

### *Innovations*

One of the participants highlighted the following:

> *...how fortuitous it is that COVID's come along at exactly the time when we've got the technological ability to do this stuff*

The key to this was the development of new autism assessment tools, taking into account "the cultural differences, and the social cultural context that people are living in." There was a keenness for "something that does what the ADOS [does], but works in an online environment," with established validity and reliability.

Potential identified innovations included allowing the patient or parents to forward videos of behavior and functioning in everyday situations, using 2 cameras to observe behavior from different angles in the clinic or at home, developing more eye-tracking or neuropsychological tasks for remote use, and having more sophisticated *screen-sharing* options.

## Theme 7: Health Professionals' Experiences and Needs

The seventh and final theme pertained to the participants' experiences and needs as health professionals, with three subthemes: (1) experiences during the COVID-19 pandemic, (2) supervision and support, and (3) training.

### *Experiences During the COVID-19 Pandemic*

The convenience, flexibility, and efficiency of working from home were favored by the participants. However, this was not without its limitations. One of the participants said the following:

> *...all of us are females in our team and [the] majority of us [have] got children as well, so it's been a bit of a balance really, having time to home school and time to do the assessments*

Many participants reported that they "like going to an office and seeing people and being around people" and "prefer sitting in a room with somebody...just to maintain human connection."

Time spent, in person, with colleagues was "absolutely critical...[for] things like humor, team building."

Several participants had experienced a sense of *isolation*, with one remarking the following:

> *I've never met my team. I've never met my supervisor. I've never met my patients in person...also it felt very isolated with the team and definitely didn't help with some team dynamics...sometimes it's nice to knock on somebody's door and asking the question, or at least meet the people we work with*

Another said the following:

> *...it's difficult working with silence...my mental health is not so good, I think, since I'm always on my own*

Some general health implications of working at home are highlighted. This could be "more tiring" and "physically intense...I've been having eyestrain and more headaches." Back problems because of "sitting so much" were more common. Another participant said, "the longer [you] spend on a screen, the more burnt out you feel." Overall, it was suggested that "actually getting up and out of your seat, and not working from a computer all the time, is actually physically more healthy."

### *Supervision and Support*

In keeping with several participants' viewpoints, one of the participants noted the following:

> *I think we've been making things up as we go along and there hasn't been very much guidance from anywhere about what we* [should do]

Supervision was deemed "more important now than ever, but it's more avoided. I think because people are just so tired with it all [the pandemic]." Some participants expressly wanted "safeguarding supervision."

Some pandemic-specific reasons for supervision were described, including the following:

> *...thinking about the impact of us not having our own routines or home life balance being so blurred, and helping people to find ways to separate work and home when they're in the same environment...the impact of the pandemic on everybody and how it changed everybody's life...emotional demands [of the] clinical job...we're all kind of going through you know extreme stress in our lives*

Several participants felt that current ways of working raised *ethical considerations* for discussion in supervision:

> *...holding [the] tension between what do I clinically feel is the right thing to do...what do families want...what is driving the decision-making process?*

One of the supervisors reflected they are "a bit more careful when [they're] supervising remotely and they've [the supervisee] assessed remotely," to ensure the diagnosis reached is accurate. Peer supervision was also described as "really important...[there is] a real power in hearing from other people."

Forums bringing together health professionals working across services were considered useful, with one participant saying the following:

> *...it [would] be fantastic to you know, see what other people [health professionals] have done and how people have changed things and what they feel, or even if it's just to confirm that what we're doing is as good as we can do*

### *Training*

The following was highlighted:

> *...none of us were trained to do electronic-based assessments as part of our background core clinical trainings. We've been forced into it. Some people have flourished staff wise, others haven't*

Few patients had received any telehealth-specific training.

Participants identified five telehealth-specific training areas for health professionals: (1) IT skills (eg, general computer literacy, using video conferencing platforms, touch typing, and digital security), (2) clinical skills (eg, knowledge of mental health and differential diagnoses and how to assess them through telehealth, conducting virtual risk assessments and management, and addressing safeguarding concerns remotely), (3) therapeutic skills (eg, deportment on the web, how to enhance virtual engagement, and rapport building), (4) autism-specific skills (eg, how to assess core symptoms and strengths on the web and training in using new [validated] diagnostic tools), and (5) reliability meetings (ie, checking consistency for clinical assessments and standardized tool use).

# Discussion

## Principal Findings

This study gathered the perspectives of professionals working across services in England and with people across the life span regarding their thoughts about and experiences of conducting telehealth autism diagnostic assessments since the start of the COVID-19 pandemic. The participants represented 7 professional disciplines and had varied experiences with autism services.

A thematic analysis of participants' responses indicated that there are several advantages associated with telehealth, particularly in relation to convenience, flexibility, and efficiency for patients, their families, and professionals and giving rise to opportunities for innovation. However, participants also reported that telehealth incurs a range of challenges, including increasing potential health care disparities; affecting confidence in assessing, formulating, and sharing diagnostic conclusions; and contributing to clinical, environmental, and practical complexities.

## Comparison With Prior Work

The findings reported here are broadly consistent with those outlined in a handful of recent studies examining professionals' experiences of providing telehealth autism diagnostic assessments in the United States and Australia [13,16,19]. Studies have reported that professionals appreciated the convenience, flexibility, efficiency, and cost and space savings of telehealth. Moreover, many professionals felt this was an acceptable and satisfactory approach during the COVID-19 pandemic, even if they would not have traditionally opted to work in this way—a finding that echoes the broader literature on telehealth [28]. This is encouraging, although further studies are needed to establish why some professionals are more in favor of telehealth methods for autism diagnostic assessments than others; for example, whether contributory factors for this include the amount of autism- and mental health–relevant experience or expertise professionals have, the type of setting they work in, their age (eg, familiarity with IT), and the degree of training and clinical supervision or support provided within services.

Similar to the findings of this study, professionals elsewhere have raised concerns about the validity and reliability of telehealth autism diagnostic assessments and difficulties in assessing core autism traits remotely [13,15-19]. Concerns have likely been amplified by the fact that professionals are unable to use mainstay diagnostic tools, notably the ADOS-2 [27], and they may not yet be trained in alternatives with preliminary validation (eg, the Brief Observation of Symptoms of Autism) [22].

A consistent theme in the emerging literature is that the subtleties of social communication (eg, modulation of eye contact and use of descriptive or emphatic gestures) can be more challenging to assess via videoconferencing; for example, given the relatively small screen and that nonverbal gestures may not be oriented toward the camera, even if directed to the screen. Similarly, repetitive behaviors (eg, mannerisms) may manifest during an assessment but outside of the camera view. In addition, stay-at-home mandates and social distancing measures have meant that many individuals have had less social contact outside their immediate family in the past 2 years than before 2020 [27]. Indirectly, this may have altered the frame of reference for social situations or social norms for young children or individuals who have been more isolated [29]. Conceivably, some individuals may experience heightened social anxiety. Therefore, in some instances, social difficulties may be evident at assessment; however, causal mechanisms (eg, autism, anxiety, and lack of exposure to social situations) may be uncertain. The implication is that professionals may need to spend more time with patients; for example, conducting an assessment over several appointments so that the individual becomes more familiar with the professional and process, speaking to others who know the person well, or clarifying differences in social styles before and during the COVID-19 pandemic. Conducting assessments jointly with a colleague or developing checklists or prompts for quantifying subtle and overt traits associated with autism may prove useful. Although there is tentative evidence of the reliability of telehealth autism diagnostic assessments [17,18], most studies were conducted before the COVID-19 pandemic. Further research is needed to establish the psychometric properties of the newly developed diagnostic tools and ways through which the validity of telehealth can be enhanced.

Importantly, digital poverty was highlighted as a potential contributory factor increasing health care disparities in this study (ie, not all patients waiting for an autism diagnostic assessment could be seen as they did not have the requisite IT equipment or reliable internet access). This reflects findings reported in studies of telehealth autism services [16,19,29] and telehealth health services more generally [30,31]. Digital poverty is not uncommon [31-33]. For example, approximately 1 in 10 United Kingdom households does not have access to a PC or mobile device with interactive access [31]. For individuals who do have access to an internet-enabled device, it may be that practicalities or cost render internet access difficult, or it may be that they lack the skills or confidence to use this adeptly [32]. Although internet use has broadly increased over the past decade (from 79.7% to 90% of the United Kingdom adult population), it is a cause for concern that individuals with longer-term health issues or from lower socioeconomic status backgrounds may be excluded from telehealth opportunities, or lack access to skill-based training or support to use this. In addition, poorer than required computer literacy of patients, their families, and professionals was highlighted as an important consideration in this study, mirroring findings elsewhere [13,16]. The unexpected onset of the COVID-19 pandemic is likely to have meant that some services were unable to swiftly assess digital poverty and competencies of patients, families, and professionals and accommodate needs accordingly. However, going forward, it is imperative that these factors are addressed with future implementation of telehealth policies, ensuring that patients and families, including individuals with neurodevelopmental or intellectual disabilities [34], are supported to access and use telehealth with ease and that professionals have the correct training and tools to conduct high-quality assessments [35]. Moreover, it would be ideal for all stakeholders to input into co-designing telehealth methods and platforms [34]. This may also include identifying which methods of telehealth are deemed more satisfactory by patients, their families, and professionals; why this is the case; and how this can be used to further iterate the services provided.

Participants identified several fundamental aspects that they considered pivotal for enhancing telehealth autism diagnostic assessments during and after the COVID-19 pandemic. These related to iterating service provision in collaboration with patients and families, offering blended models of care (ie, in person and telehealth), streamlining administrative and IT processes, ensuring patients and families have access to resources, providing professionals with the necessary equipment and training, improving team cohesion, and providing professionals with adequate clinical supervision and sources of support locally and nationally.

Recent studies [16,36,37] have similarly reported that systemic changes to service provision may enhance telehealth in autism services, as well as the acceptability and satisfaction of patients, their families, and professionals. At the onset of the COVID-19 pandemic, many services were able to maintain routine care. There was likely limited time to stop and think, broadly and systemically, about what processes might be best and why. In addition, it was not clear how long the service provision would be disrupted. Now that there is more clarity and possibly more stability in light of vaccination programs, it would be useful for clinical services to evaluate and audit provision and practice during the past 2 years.

Participants in this study identified telehealth-related areas for continuing professional development, including general clinical (eg, engagement), autism-specific, and practical skills. Studies conducted with professionals using telehealth for autism diagnostic assessment [13,16,19] or interventions for autistic individuals [26] have similarly highlighted additional training needs arising in this context. Although it is understandable that services may not have been geared up to offer specialized training at the onset of the COVID-19 pandemic, 2 years in, it seems crucial that core professional training and postqualification training incorporate skills-based sessions to support professionals in developing their competence and confidence in using telehealth. Future research could examine the impact of training on clinical work and whether the mode of delivery (eg, in person vs lectures on the web vs simulation methods) is a moderating mechanism. Ultimately, professionals are likely to require skills that enable them to relatively adeptly use blended in-person or telehealth methods.

## Limitations

This study had several limitations. We recruited participants from a wide sampling frame but were unable to assess the reach of the study information (ie, the number of potential participants who saw the study information vs the number who contacted the research team to express interest in participating). We also did not clarify the motivations for study participation (eg, strong views in favor of or against telehealth). A wide range of health professional disciplines involved in autism assessments was represented; however, there were comparatively fewer medically trained participants. Together, participants worked across different settings and types of services; however, we did not purposively recruit participants based on each service that may conduct autism diagnostic assessments (eg, the criminal justice system). All participants were based in England, which may have affected the generalizability of the findings to other countries.

## Conclusions

This is one of the first studies to explore, in-depth, health professionals' views on conducting autism diagnostic assessments via telehealth in England since the onset of the COVID-19 pandemic. The study participants represented 7 clinical disciplines and conducted diagnostic assessments with children, adolescents, and adults across most regions of the country. Together, participants were enthusiastic about many ways in which telehealth can be efficient, flexible, and limit costs, with clear examples of innovation. However, it was also evident that some patients may wait for a disproportionately long time for assessment as telehealth is not deemed appropriate, given their clinical presentation, risk issues, or digital poverty. Views differed regarding the degree to which solely using telehealth is a sufficiently valid and reliable way of assessing autism and sharing diagnostic conclusions. Further studies are needed to establish what best practice telehealth autism diagnostic assessments should comprise, alongside research that focuses on reducing health care disparities and enhancing professionals' skills and confidence in working in this way. In addition, the development of telehealth service provision should ideally incorporate stakeholder engagement and collaboration.

## Acknowledgments

The authors would like to thank the health professionals who participated in the study.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary information.
[DOCX File , 33 KB - mental_v9i7e37901_app1.docx ]

Multimedia Appendix 2
Theme 1 and indicative participant points.
[PDF File (Adobe PDF File), 78 KB - mental_v9i7e37901_app2.pdf ]

Multimedia Appendix 3
Theme 2 and indicative participant points.
[PDF File (Adobe PDF File), 94 KB - mental_v9i7e37901_app3.pdf ]

Multimedia Appendix 4
Theme 3 and indicative participant points.
[PDF File (Adobe PDF File), 75 KB - mental_v9i7e37901_app4.pdf ]

Multimedia Appendix 5
Theme 4 and indicative participant points.

[PDF File (Adobe PDF File), 76 KB - mental_v9i7e37901_app5.pdf ]

---

Multimedia Appendix 6
Theme 5 and indicative participant points.
[PDF File (Adobe PDF File), 79 KB - mental_v9i7e37901_app6.pdf ]

---

Multimedia Appendix 7
Theme 6 and indicative participant points.
[PDF File (Adobe PDF File), 79 KB - mental_v9i7e37901_app7.pdf ]

---

Multimedia Appendix 8
Theme 7 and indicative participant points.
[PDF File (Adobe PDF File), 80 KB - mental_v9i7e37901_app8.pdf ]

---

## References

1. Roman-Urrestarazu A, van Kessel R, Allison C, Matthews FE, Brayne C, Baron-Cohen S. Association of race/ethnicity and social disadvantage with autism prevalence in 7 million school children in England. JAMA Pediatr 2021 Jun 01;175(6):e210054 [FREE Full text] [doi: 10.1001/jamapediatrics.2021.0054] [Medline: 33779707]
2. American Psychiatric Association (APA). The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Washington, DC: American Psychiatric Association; 2013.
3. Happé FG, Mansour H, Barrett P, Brown T, Abbott P, Charlton RA. Demographic and cognitive profile of individuals seeking a diagnosis of autism spectrum disorder in adulthood. J Autism Dev Disord 2016 Nov;46(11):3469-3480. [doi: 10.1007/s10803-016-2886-2] [Medline: 27549589]
4. Lai M, Baron-Cohen S. Identifying the lost generation of adults with autism spectrum conditions. Lancet Psychiatry 2015 Nov;2(11):1013-1027. [doi: 10.1016/S2215-0366(15)00277-1] [Medline: 26544750]
5. Hayes J, Ford T, Rafeeque H, Russell G. Clinical practice guidelines for diagnosis of autism spectrum disorder in adults and children in the UK: a narrative review. BMC Psychiatry 2018 Jul 13;18(1):222 [FREE Full text] [doi: 10.1186/s12888-018-1800-1] [Medline: 30005612]
6. Taylor LJ, Eggleston MJ, Thabrew H, van der Meer L, Waddington H, Whitehouse AJ, et al. An investigation of adherence to best practice guidelines for autism diagnosis in New Zealand. Autism 2021 Oct 17;25(7):2087-2100. [doi: 10.1177/13623613211015757] [Medline: 34000858]
7. Penner M, Anagnostou E, Andoni LY, Ungar WJ. Systematic review of clinical guidance documents for autism spectrum disorder diagnostic assessment in select regions. Autism 2018 Jul;22(5):517-527 [FREE Full text] [doi: 10.1177/1362361316685879] [Medline: 28548543]
8. Charman T, Gotham K. Measurement issues: screening and diagnostic instruments for autism spectrum disorders - lessons from research and practise. Child Adolesc Ment Health 2013 Feb 04;18(1):52-63 [FREE Full text] [doi: 10.1111/j.1475-3588.2012.00664.x] [Medline: 23539140]
9. National Institute for Health and Care Excellence (NICE). Autism Spectrum Disorder in Under 19s: Support and Management. London: National Institute for Health and Care Excellence: Clinical Guidelines; 2021.
10. Autism: Recognition, Referral, Diagnosis and Management of Adults on the Autism Spectrum. Leicester (UK): British Psychological Society; 2012.
11. Monaghesh E, Hajizadeh A. The role of telehealth during COVID-19 outbreak: a systematic review based on current evidence. BMC Public Health 2020 Aug 01;20(1):1193 [FREE Full text] [doi: 10.1186/s12889-020-09301-4] [Medline: 32738884]
12. Baweja R, Brown SL, Edwards EM, Murray MJ. COVID-19 pandemic and impact on patients with autism spectrum disorder. J Autism Dev Disord 2022 Jan 10;52(1):473-482 [FREE Full text] [doi: 10.1007/s10803-021-04950-9] [Medline: 33689088]
13. Gibbs V, Cai RY, Aldridge F, Wong M. Autism assessment via telehealth during the Covid 19 pandemic: experiences and perspectives of autistic adults, parents/carers and clinicians. Res Autism Spectr Disord 2021 Oct;88:101859 [FREE Full text] [doi: 10.1016/j.rasd.2021.101859] [Medline: 34512797]
14. Spain D, Stewart GR, Mason D, Robinson J, Capp SJ, Gillan N, et al. Autism diagnostic assessments with children, adolescents, and adults prior to and during the COVID-19 pandemic: a cross-sectional survey of professionals. Front Psychiatry 2022;13:789449 [FREE Full text] [doi: 10.3389/fpsyt.2022.789449] [Medline: 35573336]
15. Adamou M, Jones SL, Fullen T, Galab N, Abbott K, Yasmeen S. Remote assessment in adults with Autism or ADHD: a service user satisfaction survey. PLoS One 2021 Mar 25;16(3):e0249237 [FREE Full text] [doi: 10.1371/journal.pone.0249237] [Medline: 33765076]
16. Kryszak EM, Albright CM, Fell LA, Butter EM, Kuhlthau KA. Clinician perspectives on telehealth assessment of autism spectrum disorder during the COVID-19 pandemic. J Autism Dev Disord 2022 Feb 01:1-16 (forthcoming) [FREE Full text] [doi: 10.1007/s10803-022-05435-z] [Medline: 35103899]

17. Alfuraydan M, Croxall J, Hurt L, Kerr M, Brophy S. Use of telehealth for facilitating the diagnostic assessment of Autism Spectrum Disorder (ASD): a scoping review. PLoS One 2020;15(7):e0236415 [FREE Full text] [doi: 10.1371/journal.pone.0236415] [Medline: 32702017]
18. Ellison KS, Guidry J, Picou P, Adenuga P, Davis TE. Telehealth and autism prior to and in the age of COVID-19: a systematic and critical review of the last decade. Clin Child Fam Psychol Rev 2021 Sep;24(3):599-630 [FREE Full text] [doi: 10.1007/s10567-021-00358-0] [Medline: 34114135]
19. Matthews NL, Skepnek E, Mammen MA, James JS, Malligo A, Lyon A, et al. Feasibility and acceptability of a telehealth model for autism diagnostic evaluations in children, adolescents, and adults. Autism Res 2021 Dec 11;14(12):2564-2579. [doi: 10.1002/aur.2591] [Medline: 34378858]
20. Wagner L, Weitlauf AS, Hine J, Corona LL, Berman AF, Nicholson A, et al. Transitioning to telemedicine during COVID-19: impact on perceptions and use of telemedicine procedures for the diagnosis of autism in toddlers. J Autism Dev Disord 2022 May 04;52(5):2247-2257 [FREE Full text] [doi: 10.1007/s10803-021-05112-7] [Medline: 34085153]
21. Jang J, White S, Esler A, Kim SH, Klaiman C, Megerian JT, et al. Diagnostic evaluations of Autism Spectrum Disorder during the COVID-19 pandemic. J Autism Dev Disord 2022 Feb;52(2):962-973 [FREE Full text] [doi: 10.1007/s10803-021-04960-7] [Medline: 33730320]
22. Dow D, Holbrook A, Toolan C, McDonald N, Sterrett K, Rosen N, et al. The brief observation of symptoms of autism (BOSA): development of a new adapted assessment measure for remote telehealth administration through COVID-19 and beyond. J Autism Dev Disord 2021 Dec 16:1-12 (forthcoming) [FREE Full text] [doi: 10.1007/s10803-021-05395-w] [Medline: 34914016]
23. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
24. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
25. Berger R. Now I see it, now I don't: researcher's position and reflexivity in qualitative research. Qual Res 2013 Jan 03;15(2):219-234. [doi: 10.1177/1468794112468475]
26. Lord C, Rutter M, Le Couteur A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J Autism Dev Disord 1994 Oct;24(5):659-685. [doi: 10.1007/BF02172145] [Medline: 7814313]
27. Lord C, Rutter M, DiLavore PC, Risi S, Gotham K, Bishop S. Autism Diagnostic Observation Schedule, Second Edition. Torrance, CA: Western Psychological Services; 2012.
28. Johnsson G, Bulkeley K. Practitioner and service user perspectives on the rapid shift to teletherapy for individuals on the autism spectrum as a result of COVID-19. Int J Environ Res Public Health 2021 Nov 11;18(22):11812 [FREE Full text] [doi: 10.3390/ijerph182211812] [Medline: 34831567]
29. Pietrabissa G, Simpson SG. Psychological consequences of social isolation during COVID-19 outbreak. Front Psychol 2020;11:2201 [FREE Full text] [doi: 10.3389/fpsyg.2020.02201] [Medline: 33013572]
30. Wijaya R, Bunga B, Kiling I. Socio-emotional struggles of young children during COVID-19 pandemic: social isolation and increased use of technologies. J Early Childhood Res 2021 Oct 24;20(1):113-127 [FREE Full text] [doi: 10.1177/1476718x211052789]
31. Aisbitt GM, Nolte T, Fonagy P. Editorial Perspective: the digital divide - inequalities in remote therapy for children and adolescents. Child Adolesc Ment Health 2022 Feb 18 (forthcoming). [doi: 10.1111/camh.12545] [Medline: 35179284]
32. van Kessel R, Hrzic R, O'Nuallain E, Weir E, Wong BL, Anderson M, et al. Digital health paradox: international policy perspectives to address increased health inequalities for people living with disabilities. J Med Internet Res 2022 Feb 22;24(2):e33819 [FREE Full text] [doi: 10.2196/33819] [Medline: 35191848]
33. Honeyman M, Maguire D, Evans H, Davies A. Digital technology and health inequalities: a scoping review. Public Health Wales NHS Trust. 2020. URL: https://phw.nhs.wales/publications/publications1/digital-technology-and-health-inequalities-a-scoping-review/ [accessed 2022-04-26]
34. Samuels-Kalow M, Jaffe T, Zachrison K. Digital disparities: designing telemedicine systems with a health equity aim. Emerg Med J 2021 Jun 04;38(6):474-476. [doi: 10.1136/emermed-2020-210896] [Medline: 33674277]
35. Exploring the UK's digital divide. Office for National Statistics. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/articles/exploringtheuksdigitaldivide/2019-03-04 [accessed 2022-03-02]
36. Chiauzzi E, Clayton A, Huh-Yoo J. Videoconferencing-based telemental health: important questions for the COVID-19 era from clinical and patient-centered perspectives. JMIR Ment Health 2020 Dec 08;7(12):e24021 [FREE Full text] [doi: 10.2196/24021] [Medline: 33180739]
37. Spain D, Mason D, J Capp S, Stoppelbein L, W White S, Happé F. "This may be a really good opportunity to make the world a more autism friendly place": professionals' perspectives on the effects of COVID-19 on autistic individuals. Res Autism Spectrum Disorders 2021 May;83:101747 [FREE Full text] [doi: 10.1016/j.rasd.2021.101747]

## Abbreviations

**ADOS:** Autism Diagnostic Observation Schedule
**MDT:** multidisciplinary team
**NHS:** National Health Service

---

*Edited by J Torous; submitted 10.03.22; peer-reviewed by R van Kessel, T Ewais; comments to author 31.03.22; revised version received 27.04.22; accepted 27.04.22; published 20.07.22.*

*Please cite as:*
*Spain D, Stewart GR, Mason D, Milner V, Fairhurst B, Robinson J, Gillan N, Ensum I, Stark E, Happe F*
*Telehealth Autism Diagnostic Assessments With Children, Young People, and Adults: Qualitative Interview Study With England-Wide Multidisciplinary Health Professionals*
*JMIR Ment Health 2022;9(7):e37901*
*URL: https://mental.jmir.org/2022/7/e37901*
*doi:10.2196/37901*
*PMID:35857358*

Original Paper

# Acoustic and Linguistic Features of Impromptu Speech and Their Association With Anxiety: Validation Study

Bazen Gashaw Teferra[1], BSc, MSc; Sophie Borwein[2], BA, MPP, PhD; Danielle D DeSouza[3], BSc, MSc, PhD; William Simpson[3,4], BSc, PhD; Ludovic Rheault[5], PhD; Jonathan Rose[1], BASc, MASc, PhD

[1]The Edward S Rogers Sr Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada
[2]School of Public Policy, Simon Fraser University, Vancouver, BC, Canada
[3]Winterlight Labs, Toronto, ON, Canada
[4]Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, ON, Canada
[5]Department of Political Science, Munk School of Global Affairs and Public Policy, University of Toronto, Toronto, ON, Canada

**Corresponding Author:**
Bazen Gashaw Teferra, BSc, MSc
The Edward S Rogers Sr Department of Electrical and Computer Engineering
University of Toronto
10 King’s College Road
Toronto, ON, M5S 3G4
Canada
Phone: 1 416 978 6992
Email: bazen.teferra@mail.utoronto.ca

## Abstract

**Background:** The measurement and monitoring of generalized anxiety disorder requires frequent interaction with psychiatrists or psychologists. Access to mental health professionals is often difficult because of high costs or insufficient availability. The ability to assess generalized anxiety disorder passively and at frequent intervals could be a useful complement to conventional treatment and help with relapse monitoring. Prior work suggests that higher anxiety levels are associated with features of human speech. As such, monitoring speech using personal smartphones or other wearable devices may be a means to achieve passive anxiety monitoring.

**Objective:** This study aims to validate the association of previously suggested acoustic and linguistic features of speech with anxiety severity.

**Methods:** A large number of participants (n=2000) were recruited and participated in a single web-based study session. Participants completed the Generalized Anxiety Disorder 7-item scale assessment and provided an impromptu speech sample in response to a modified version of the Trier Social Stress Test. Acoustic and linguistic speech features were a priori selected based on the existing speech and anxiety literature, along with related features. Associations between speech features and anxiety levels were assessed using age and personal income as covariates.

**Results:** Word count and speaking duration were negatively correlated with anxiety scores ($r$=–0.12; $P$<.001), indicating that participants with higher anxiety scores spoke less. Several acoustic features were also significantly ($P$<.05) associated with anxiety, including the mel-frequency cepstral coefficients, linear prediction cepstral coefficients, shimmer, fundamental frequency, and first formant. In contrast to previous literature, second and third formant, jitter, and zero crossing rate for the $z$ score of the power spectral density acoustic features were not significantly associated with anxiety. Linguistic features, including negative-emotion words, were also associated with anxiety ($r$=0.10; $P$<.001). In addition, some linguistic relationships were sex dependent. For example, the count of words related to power was positively associated with anxiety in women ($r$=0.07; $P$=.03), whereas it was negatively associated with anxiety in men ($r$=–0.09; $P$=.01).

**Conclusions:** Both acoustic and linguistic speech measures are associated with anxiety scores. The amount of speech, acoustic quality of speech, and gender-specific linguistic characteristics of speech may be useful as part of a system to screen for anxiety, detect relapse, or monitor treatment.

***(JMIR Ment Health 2022;9(7):e36828)*** doi:10.2196/36828

**KEYWORDS**

mental health; generalized anxiety disorder; impromptu speech; acoustic features; linguistic features; mobile phone

## *Introduction*

### Background

Anxiety disorders are among the most common mental health issues, with an incidence of approximately 10% in the Canadian population [1]. Many Canadians are unable to access psychological and psychiatric resources to help those affected [2], in part, because of the cost of professional help [3]. It may be possible to address some of this deficit using methods that automate the measurement and diagnosis of anxiety disorders. The first step in this direction is to explore methods for the automatic detection of mental health issues that could be used to trigger early intervention, monitor treatment response, or detect relapse. In addition, frequent monitoring together with other time-series information could be used to help understand the mechanisms of generalized anxiety disorder (GAD) itself. An avenue of such automation is recording an individual's speech and looking for signals of anxiety within the recordings.

In this work, we focused specifically on GAD [4]. A reason that GAD may be detectable in speech is that those with anxiety disorders exhibit higher activation of the sympathetic nervous system under stress than those without anxiety [5], which in turn influences the production of speech [6]. The goal of this work was to collect a large set of samples of audio speech, each with a self-reported measure of anxiety scale, and explore whether acoustic and linguistic signals correlated with measured anxiety. We built on previous studies by collecting approximately 10 times greater number of human participants than previous research on the detection of anxiety in speech. Many of the signals that we explored have been previously reported as significantly correlated with anxiety in the literature, and our goal was to leverage our larger sample size to examine which signals could be most useful in identifying anxiety in speech. We also explored linguistic indicators of anxiety that have not been considered before.

This paper is organized as follows: the next section summarizes related work in anxiety detection. The *Methods* section describes the speech sample collection methods and the set of features considered for correlation with anxiety. The *Results* section reports on the demographics of participants and feature correlations, whereas the *Discussion* section discusses the results and their implications for future research on anxiety detection. A final section provides our conclusions.

### Related Work

Although it is important to note that some scholarship is skeptical that biomarkers correlate with emotions [7], here we review existing work exploring associations between both acoustic and linguistic speech features and anxiety severity in healthy and clinical cohorts. It should be noted that these studies explore broader classes of anxiety disorders, including internalizing disorders, social phobia or social anxiety disorder (SAD), panic disorder, and agoraphobia, as well as GAD.

McGinnis et al [8] identified several acoustic characteristics of speech that can be used to detect anxiety disorders in children. Studying 71 participants aged 3 to 8 years, the researchers were able to detect internalizing disorders—a collective term for anxiety and depression—from speech. The authors extracted and selected several acoustic features from the speech produced in a 3-minute task based on the Trier Social Stress Test (TSST) for children [9]. These features included zero crossing rate, mel-frequency cepstral coefficients (MFCCs) [10], zero crossing rate for the *z* score of the power spectral density (ZCR-zPSD), dominant frequency, mean frequency, perceptual spectral centroid, spectral flatness, and the skew and kurtosis of the power spectral density. Using the Davies-Bouldin index–based feature selection [11], the MFCC features and ZCR-zPSD had the highest Davies-Bouldin score. Several models were built to predict which children had an internalizing disorder (n=43 out of 71) or were healthy. Both logistic regression and support vector machine [12] analysis achieved a classification accuracy of 80%.

Özseven et al [13] conducted a study of the speech of 43 adults aged 17 to 55 years. Of these 43 adults, 21 were clinically diagnosed with GAD, 2 were diagnosed with panic disorder, and 20 were healthy controls. The study explored 122 acoustic features derived from the participants' speech to determine the correlation between these features and anxiety. Their results showed that 42 of the features (including MFCCs, linear prediction cepstral coefficients [LPCCs], fundamental frequency [F0], first formant [F1], second formant [F2], third formant [F3], jitter, and shimmer) showed a significant change between a neutral state and an anxious state in the participants with anxiety.

Weeks et al [14] found a relationship between anxiety and alterations in voice. Specifically, their study showed a link between vocal pitch (characterized by F0) and SAD. They collected impromptu speech samples from 46 undergraduate students, 25 with a diagnosis of SAD and 21 healthy controls. Participants also completed the Beck Anxiety Scale as a measure of self-reported anxiety severity [15]. Their results indicated that mean F0 was positively correlated ($r$=0.72; $P$=.002) with anxiety severity across all male participants. However, the correlation for female participants was weaker ($r$=0.02; $P$=.92), indicating possible sex differences in the relationship between anxiety severity and vocal pitch.

Laukka et al [16] explored the relationship between anxiety and the acoustic features of speech. They collected speech data from 71 patients with social phobia delivering public speeches and extracted 4 types of speech features: pitch (F0 mean, F0 SD, and F0 maximum), loudness (intensity mean), voice quality (HF 500, relative proportion of high-frequency spectral energy above vs below 500), and temporal aspects of speech (articulation rate and percentage of silence). The researchers observed a significant change from before treatment to after treatment (a pharmacological anxiolytic treatment for social anxiety) in F0 mean, F0 maximum, HF 500, and percentage of silence. They also calculated the Pearson correlation coefficient between state anxiety measured by the Spielberger State-Trait Anxiety Inventory [17] and the speech features. Those with a significant correlation were F0 SD (*r*=–0.24; *P*<.05) and percentage of silence (*r*=0.36; *P*<.01).

Albuquerque et al [18] investigated the relationship between acoustic speech features and anxiety. They recruited 112 adult Portuguese speakers who performed 2 tasks: reading vowels in disyllabic words and picture description. The authors extracted 18 acoustic features, including F0, F1, F2, speech duration, number of pauses, and articulation rate. They measured the percentage change between participants who were nonanxious (Hospital Anxiety and Depression Scale, Anxiety subscale [19] score ≤7) and those who were anxious (Hospital Anxiety and Depression Scale, Anxiety subscale score >7) and observed a change of >10% in speech duration.

Wörtwein et al [20] assessed the behaviors of participants experiencing anxiety caused by public speaking through audiovisual features. A total of 45 participants were recruited from Craigslist. These participants were asked to complete the Personal Report of Confidence as a Speaker scale [21], which estimates public speaking anxiety levels. Several audio features were extracted from the audio and their results showed significant relationships between the Personal Report of Confidence as a Speaker scale and SD of the 0th coefficient of the MFCC [10] (*r*=–0.36; *P*<.05), SD of F1 (*r*=–0.41; *P*<.01), and the total pause duration (*r*=0.35; *P*<.05).

Hagenaars and van Minnen [22] explored whether the activation of fear was manifested in the speech of 25 female patients diagnosed with panic disorder. Their results showed that patients with panic disorder have a significantly higher pitch (*P*<.001) during autobiographical fear memory. Respondents also spoke significantly slower (*P*<.001) during autobiographical talking than during script talking.

Di Matteo et al [23] explored the relationship between linguistic features of speech and anxiety. Their work used *passively* collected intermittent samples of audio data from participants' smartphones, collected over a 2-week period, as input. The study had 84 nonclinical participants recruited from a web-based recruitment platform. The audio was converted to text, and the authors used the Linguistic Inquiry and Word Count (LIWC) approach [24] to classify the words into 67 different categories. They calculated correlations with 4 self-report measures: SAD, GAD, depression, and functional impairment. They observed a significant correlation between words related to perceptual processes (eg, *see* in the LIWC) with SAD (*r*=0.31; *P*=.003) and words related to rewards with GAD (*r*=–0.29; *P*=.007).

In a similar study that used LIWC features, Anderson et al [25] recruited 42 participants diagnosed with SAD and 27 healthy controls to explore the differences in the words used between these 2 groups. The participants were asked to write a distinct autobiographical and socially painful passage. They used the LIWC to extract the word count in each of the LIWC categories, such as first-person singular pronouns, anxiety-related words, and fear-related words. Their results showed that patients with SAD used more first-person singular pronouns (I, me, and mine), anxiety-related words, sensory and perceptual words, and words denoting physical touch, as well as fewer references to other people.

Overall, previous work identifies several audio features that are correlated with anxiety. However, the results are mixed because of differences in participants recruited, speech measures assessed, statistical methods used, and amount of mood induction. In addition, the largest sample size among these studies was 112, which limits the potential for generalizability to the larger population, a necessary step before considering the deployment of technologies for passive anxiety monitoring. In this study, we recruited a substantially larger cohort (n=2000) to explore features of speech from previous findings at a greater scale.

## Methods

### Data Collection

Participants from a nonclinical population were recruited for a 10- to 15-minute task implemented through a custom website. Self-report measures of anxiety were collected once at the beginning of the study and at the end of each of 2 specific tasks. In the following subsections, we describe the recruitment of participants, the data collection procedure, and the assessment of anxiety and speech measures.

### Ethics Approval

The study was approved by the University of Toronto Research Ethics Board (37584).

### Recruitment and Demographics

A total of 2000 participants were recruited using the Prolific [26] web-based human participant recruitment platform. Prolific maintains a list of registered participants and, for each participant, many characteristics, including age, income, sex, primary language spoken, country of birth, and residence. The inclusion criteria for this study were as follows: age range 18 to 65 years; fluency in English; English as a first language; and at least 10 previous studies completed on Prolific, with 95% of these previous Prolific tasks completed satisfactorily, as labeled by the study author. The data set was also balanced for sex (1000/2000, 50% female, and 1000/2000, 50% male). The Prolific platform provides us with some relevant demographics of the participants, including their age and income.

Participants who completed the study were paid £2 (US $2.74). They were able to complete the entire study remotely, using their PCs.

### Study Procedure

Participants were presented with the opportunity to participate in this study on Prolific if they met the aforementioned inclusion criteria. Those who wished to participate clicked on the study link, which brought them to a consent form that described the procedure and goals of the study and also provided information on data privacy. After they gave consent, a hyperlink brought participants to an external web application (a screenshot of which is presented in Multimedia Appendix 1) that implemented the tasks described in the following sections.

Participants were first asked to fill out the standard Generalized Anxiety Disorder 7-item scale (GAD-7) questionnaire [27], which is described in more detail in the *Anxiety Measures* section. Next, they were asked to complete 2 speech tasks, which were recorded using their computer's internal microphone. It should be noted that our protocol also involved recording a video of the participants' faces during both speech tasks. Although that video is not used in the work reported here, the fact that the video was requested may have influenced the set of participants willing to continue participation, as discussed later in this paper.

For the first speech task (task 1), participants were asked to read aloud a specific passage titled *My Grandfather*, which is a public domain passage that contains nearly all the phonemes of American English [28]. The full script of this passage is presented in Multimedia Appendix 2. This passage is not intended to induce stress or anxiety but to provide a baseline speech sample for each participant. It was used in this work to test the quality of the speech-to-text (STT) transcription.

For the second speech task (task 2), the participant followed a modified version of the widely used TSST [29] for the purpose of inducing a moderate amount of stress. We chose to base our anxiety stimulus on the TSST as previous studies [30,31] have shown a higher activation in participants with relatively higher anxiety after exposure to moderate stress induced by the TSST.

In this modified version of the TSST, participants were told to imagine that they were a job applicant for a job that they really want (their *dream* job) and they were invited for an interview with a hiring manager. They were given a few minutes to prepare—to decide what their *dream* job is—and how they would convince an interviewer that they are the right person for the position. Participants were also told that the recorded video would be viewed by researchers studying their behavior and language. Participants were then asked to speak for 5 minutes, making the case for themselves to be hired for that dream job.

It should be noted that in the original TSST [29], participants would normally deliver their speech in front of a live panel of judges. If a participant finished their delivery in <5 minutes, the judges in the original TSST design would encourage the participant to keep speaking for the full 5 minutes. An example statement of encouragement is as follows: "What are your personal strengths?" In our modified TSST, we implemented a similar method to encourage participants to speak for the full 5 minutes. When our software detected silence (the absence of speech for >6 seconds), it displayed several different prompts, which are reproduced in Multimedia Appendix 3, inviting participants to keep speaking on different topics related to the task. Finally, it should be noted that the modified TSST only made use of the first part of the original TSST and not the second task involving mental arithmetic.

### Anxiety Measures

Our goal was to examine possible correlations between features of speech and GAD, based largely on previously suggested features. To measure the severity of GAD, we used the GAD-7 [27], which is a 7-item questionnaire that asks participants how often they were bothered by anxiety-related problems during the previous 2 weeks. Although the 2-week period suggests that the GAD-7 measures a temporary condition, this seems to be in contradiction with the fact that a GAD diagnosis requires a 6-month duration of symptoms [32,33]. However, the GAD-7 has been validated as a diagnostic tool for GAD (using a value of 10 as the cutoff threshold) with a sensitivity of 89% and a specificity of 82% [27]. Thus, we chose to use the GAD-7 to obtain a binary label of GAD (using the same threshold of 10) as our main indicator of anxiety.

Each of the 7 questions on the GAD-7 has 4 options for the participant to select from, indicating how often they have been bothered by the 7 problems on the scale. These options and their numerical ratings are as follows: 0=not at all, 1=several days, 2=more than half the days, and 3=nearly every day. The final GAD-7 score is a summation of the values for each question, giving a severity measure for GAD in the range of 0 (no anxiety symptoms) to 21 (severe anxiety symptoms).

We also used a second, informal anxiety measure in this study to serve as an internal check to measure how much, on average, the modified TSST (task 2) induced stress and anxiety compared with task 1 (the reading or speaking of the *MyGrandfather* passage). Here, we used a single question to measure self-reported levels of anxiety on a 4-point scale. We asked participants how anxious they felt during the task and to choose from the following numerical rating: 0=not anxious at all, 1=somewhat anxious, 2=very anxious, and 3=extremely anxious. This question was deployed immediately after the first and second tasks had been completed.

### Selection of Acoustic and Linguistic Features

#### Overview

Prior work suggested that information about the mental state of a person may be acquired from the signals within speech acoustics [34] and the language used [35]. We refer to each kind of this extracted information as a *feature* using the terminology used in the field of machine learning.

In this work, we considered both acoustic and linguistic features, which are described in the following sections. These features were extracted from each of the 5-minute speech samples in which the participant responded to the modified TSST task. It should be noted that all the participants were prompted to speak for the full 5 minutes, as described in the *Study Procedure* section, although the total speech duration of each participant may vary.

#### Acoustic Features

##### Overview

Previous research has identified several acoustic features that are correlated with anxiety, as described in the *Related Work* section. Using these previous findings as a reference point, we selected the acoustic features described in the following sections for our empirical analysis. The features were extracted using the following software packages: My-Voice Analysis [36], Surfboard [37], and Librosa [38].

#### MFCC Features

These are coefficients derived from a mel-scale cepstral representation of an audio signal. We included 13 MFCCs, a common set of acoustic signals designed to reflect changes in perceivable pitch. The MFCC features were shown to be related to anxiety in 3 studies [8,13,20]. Descriptive statistics (mean and SD) of the 13 MFCC features were used in this study. It should be noted that not all MFCC features included in this study were determined to be significant in prior work; however, these 13 are most commonly assessed together, and thus, we included them all as features of interest. The parameters we used when extracting these 13 MFCC features were as follows: window length=2048 samples, length of fast Fourier transform window=2048 samples, samples advance between successive frames=512 samples, window type=Hanning, and number of mel bands=128.

#### LPCC Features

These are coefficients derived from a linear prediction cepstral representation of an audio signal. The first 13 cepstrum coefficients were used here. The LPCC features were shown to be related with anxiety in the study by Özseven et al [13]. Descriptive statistics (mean and SD) of the 13 LPCCs were used in our study.

#### ZCR-zPSD Features

In the study by McGinnis et al [8], ZCR-zPSD was one of the top features selected using Davies-Bouldin index–based feature selection [11] for an anxiety-prediction task.

#### Amount of Speech

This refers to the amount of speech and related metrics such as the percentage of silence. These features have been shown to be related to anxiety in 3 studies [16,18,20]. Our specific feature was the amount of time, in seconds, that speech was present. We also counted the total number of words present in an STT transcript as a separate measure of the amount of speech.

#### Articulation Rate

This indicates how fast the participant spoke. The study by Hagenaars and van Minnen [22] suggested that patients with panic disorder spoke significantly slower ($P<.001$) during autobiographical talking than when reading a script.

#### F0 Feature

This is the frequency at which the glottis vibrates, also known as the *pitch* of the voice. Multiple studies have shown F0 to be one of the acoustic features affected by anxiety [13,14,16,22]. F0 varies throughout a person's speech; therefore, both the mean and SD of F0 were used as features.

#### F1, F2, and F3 Features

These are the F1, F2, and F3 [39]. The study by Özseven et al [13] showed a significant relation with anxiety. The mean and SD of each formant were used as features.

#### Jitter

This refers to the cycle-to-cycle F0 variation of the sound wave. *Jitter* has been shown to be an indicator of anxiety [13,40,41].

#### Shimmer

This refers to the cycle-to-cycle amplitude variation of the sound wave. *Shimmer* has been shown to be related to anxiety severity [13].

#### Intensity

The squared mean of the amplitude of the sound wave within a given frame, also known as *intensity*, has been shown to be related to anxiety [16]. As the amplitude of a sound wave varies during speech, the mean and SD were used as features.

### *Linguistic Features*

Using Amazon's AWS STT [42] program, a transcript was produced from the audio recordings. From the transcripts, linguistic features were extracted using the LIWC software (Pennebaker Conglomerates, Inc) [24], which places words into dictionaries based on semantic categories. For example, 1 category is called *negemo* and contains words that relate to negative emotions, such as *hurt*, *ugly*, and *nasty*. Another category is called *health* and contains words such as *clinic*, *flu*, and *pill*. There is also a category called *anxiety*, which includes words such as *anxiety* and *fearful*. Some categories are contained within others; for example, *anxiety* is contained within *negemo*.

To apply the LIWC dictionaries, one simply counts the number of words that belong to each category, and each count becomes a feature. There are 93 categories in the LIWC, although not all are relevant for an STT transcript. We removed those features that were not relevant; for example, informal language words such as *lol* and *btw*. Other excluded categories included those related to some punctuation marks (eg, colons, quotation marks, and parentheses). After removing these, 80 linguistic features remained. Prior work [23,25], which was discussed in the *Related Work* section, has shown that LIWC categories related to perceptual processes (see, hear, and feel), words related to rewards, the use of the first-person singular pronoun, and anxiety-related words were associated with anxiety.

## Separation of Data for Analysis

The overarching objective of this study was to gain an understanding of which features of speech—both acoustic and linguistic—are correlated with the GAD-7. However, it is known that certain demographic attributes are directly indicative of anxiety. For example, sex is known to influence the prevalence of anxiety [43]. In addition, both age [44] and income [45] influence anxiety, which suggests the need to control for these demographics. An additional reason to control for the demographics is that both age and income have been shown to be related to speech features [46,47]. towing to the strong effect of sex on the GAD-7 score, we created separate data sets for analysis of female and male samples, in addition to the combined data set. We chose to do this, rather than correcting for sex computationally, because it leaves the data intact.

## Statistical Analysis

The partial Pearson correlation coefficient [48] was computed between each of the features and the GAD-7 (controlling for the effect of age and personal income). Correlations were examined for 3 versions of the data set: the entire sample data set and separately by sex for male and female participants. We considered a result statistically significant at a significance level of *P*=.05. The *P* values were not corrected to account for the large number of tests as we attempted to use features that were determined to be significant in previous works.

# Results

## Overview

This section reports the main empirical results. We begin by discussing the recruitment yield, the demographic characteristics of the participants, and the relationship between demographic attributes and the reported GAD-7 score. Next, we report correlations for the features described in the *Selection of Acoustic and Linguistic Features* section.

## Recruitment and Data Inclusion

A total of 4542 participants accepted the offer from the Prolific recruitment platform to participate in the study, of whom 2212 (48.7%) completed the study, giving a recruitment yield of approximately 49%.

Of the 2212 participants who completed the study, 2000 (90.42%) provided acceptable submissions (and thus received payment), giving a submission-to-approval yield of approximately 90%. To be clear, the recruitment continued until 2000 acceptable submissions were received. The reasons for which submissions were deemed unacceptable included the following: a missing video, a missing or grossly imperfect audio, or failure to complete one or both tasks. These acceptability criteria were distinct from those used in the subsequent review of audio quality that is described in the following paragraphs. The period of recruitment ranged from November 23, 2020, to May 28, 2021. Of note, the recruitment took place during the global COVID-19 pandemic.

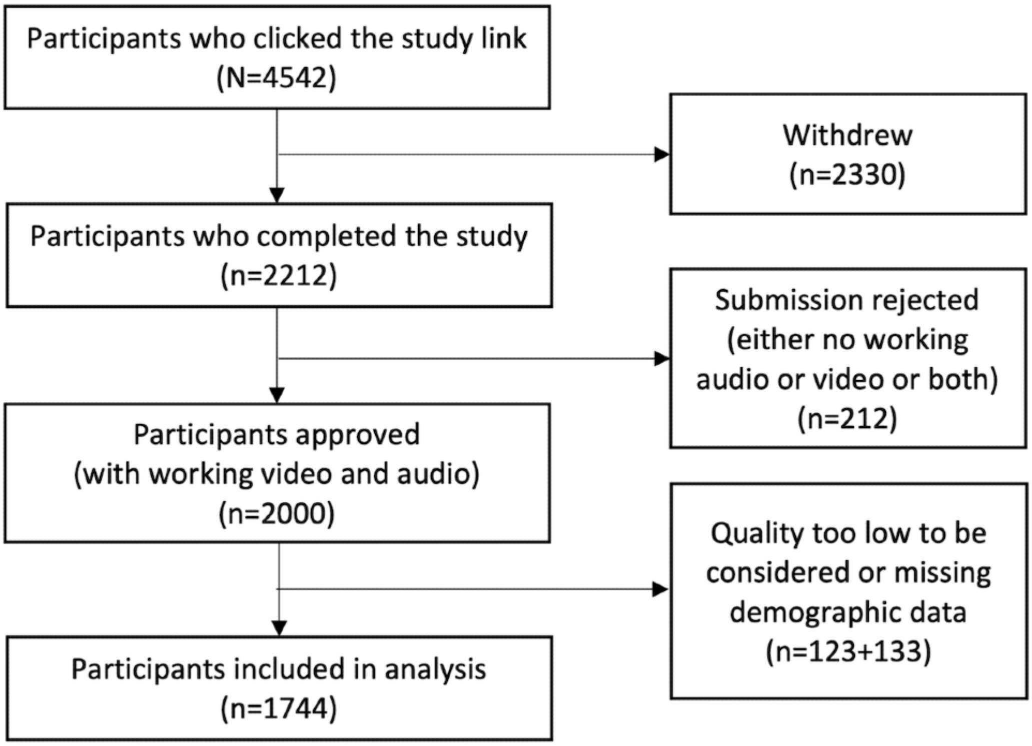
In addition to the aforementioned submission approval criteria, we reviewed the input data and audio for acceptability using the following procedure. To begin, we computed all acoustic and linguistic features described in the *Selection of Acoustic and Linguistic Features* section. Recordings with poor quality were filtered out for manual review based on the following criteria:

1. A task 2 word count of <125
2. A speaking duration for task 2 of <60 seconds (compared with the full 5 minutes)
3. Any other feature value being beyond 3 SDs from the mean in either direction (outliers)

Of the 2000 participant recordings, 193 (9.65%) were flagged based on these criteria. For each of these, a researcher (BGT) listened to the task 2 audio recordings. The researcher discarded any samples that were deemed, subjectively, to be of insufficient audio quality or those whose response to task 2 was not responsive to the task itself. Of the 193 flagged participants, 123 (63.7%) were rejected through this manual review, meaning that of the 2000 samples, 1877 (93.85%) remained.

Finally, the 1877 samples were checked for missing data, and 133 (7.09%) participants had missing demographic information; consequently, the final number of participants included in our analysis was 1744 (92.91%). The flow chart of the study recruitment and quality control is presented in Figure 1. We also explored correlations of the excluded data with the GAD-7, often called missingness analysis, and this is presented in Multimedia Appendix 4.

**Figure 1.** Study recruitment flow chart.



## Data Overview and Demographics of Participants

Of the 1744 participants, 540 (30.96%) were above the GAD-7 screening threshold of 10 and 1204 (69.04%) were below the GAD-7 screening threshold of 10. Hereon, we will refer to those participants with a GAD-7 score ≥10 as the group *with anxiety* and those with a GAD-7 score <10 as the *nonanxious* group.

Table 1 shows participants' demographics, obtained from the Prolific recruitment platform. Columns 1 and 2 of the table show the name of demographic attributes and each category, whereas columns 3 and 4 give the number (and percentage) of participants with that attribute in the group with anxiety and the nonanxious group, respectively. Column 5 gives the *P* value for a chi-square test of the null of independence to determine whether there is a significant difference between the group with anxiety and the nonanxious group for each categorical factor.

**Table 1.** Demographic characteristics of participants in the group with anxiety and the nonanxious group (N=1744).

| Demographic factors | | Group with anxiety (n=540), n (%) | Nonanxious group (n=1204), n (%) | *P* value from chi-square test |
|---|---|---|---|---|
| **Sex** | | | | <.001 |
| | Male | 229 (42.41) | 653 (54.24) | |
| | Female | 311 (57.59) | 551 (45.76) | |
| **Self-reported ongoing mental health illness or condition** | | | | <.001 |
| | Yes | 297 (55) | 311 (25.83) | |
| | No | 243 (45) | 893 (74.17) | |
| **Personal income, pounds sterling (£1=US $1.37)** | | | | <.001 |
| | <10,000 | 181 (33.52) | 281 (23.34) | |
| | 10,000 to 19,999 | 112 (20.74) | 208 (17.28) | |
| | 20,000 to 29,999 | 92 (17.04) | 259 (21.51) | |
| | 30,000 to 39,999 | 60 (11.11) | 184 (15.28) | |
| | 40,000 to 49,999 | 36 (6.67) | 109 (9.05) | |
| | 50,000 to 59,999 | 20 (3.7) | 74 (6.15) | |
| | ≥60,000 | 39 (7.22) | 89 (7.39) | |
| **Age (years)** | | | | <.001 |
| | 18 to 19 | 27 (5) | 44 (3.65) | |
| | 20 to 29 | 239 (44.26) | 379 (31.48) | |
| | 30 to 39 | 162 (30) | 334 (27.74) | |
| | 40 to 49 | 67 (12.41) | 219 (18.19) | |
| | 50 to 59 | 39 (7.22) | 132 (10.96) | |
| | ≥60 | 6 (1.11) | 96 (7.97) | |

## Posttask Self-report Anxiety Measure

As described in the *Anxiety Measures* section, participants were asked to rate their state of anxiety after each task on a scale of 0 to 3, where 3 was the highest level of anxiety. A paired 2-tailed *t* test was conducted to assess the difference between the 2 measurements. The test validates that the modified TSST task successfully induced some anxiety in participants, with the average score on the self-reported state anxiety measure increasing from 0.5 (SD 0.6) to 1.5 (SD 0.9; *P*<.001) before and after completing task 2, respectively.

## Feature Correlations

### Overview

The *Selection of Acoustic and Linguistic Features* section describes the set of acoustic and linguistic features that were selected. These were features that were reported as significant in prior work on anxiety and speech, as well as closely associated features. These features were computed on the speech samples of participants performing task 2—the modified TSST. The following subsections summarize the main empirical results. The correlation between demographics and the acoustic and linguistic features is presented in Multimedia Appendix 5, and the intercorrelation among the significant features is presented in Multimedia Appendix 6, Multimedia Appendix 7, and Multimedia Appendix 8 for the all-sample, female-sample, and male-sample data sets, respectively.

### Amount of Speech

The features with one of the highest correlations for both the male-sample and female-sample data sets were those related to the amount the participant spoke during task 2. The 2 specific features used to estimate speech length were speaking duration (the number of seconds of speech present within the 5-minute speech task) and the word count derived from an STT transcript. Table 2 presents the correlation for the all-sample data set (controlling for sex, age, and income) and for separated female-sample and male-sample data sets (controlling for age and income). Figure 2 presents a scatter plot of speaking duration versus the GAD-7, as well as the distribution of both variables, for all 3 data sets. The scatter plot is colored to give a better sense of the density of data points. Figure 3 provides the same kind of scatter plots and distributions for the word count metric of task 2.

**Table 2.** Correlation of amount of speech features with the Generalized Anxiety Disorder 7-item scale.

| Sample and feature | *r* | *P* value |
|---|---|---|
| **All samples (N=1744)** | | |
| Speaking duration | –0.12 | <.001 |
| Word count | –0.12 | <.001 |
| **Female samples (n=862)** | | |
| Word count | –0.13 | <.001 |
| Speaking duration | –0.11 | <.001 |
| **Male samples (n=882)** | | |
| Speaking duration | –0.13 | <.001 |
| Word count | –0.12 | <.001 |

**Figure 2.** Speaking duration versus Generalized Anxiety Disorder 7-item scale (GAD-7) scatter plot and distributions.

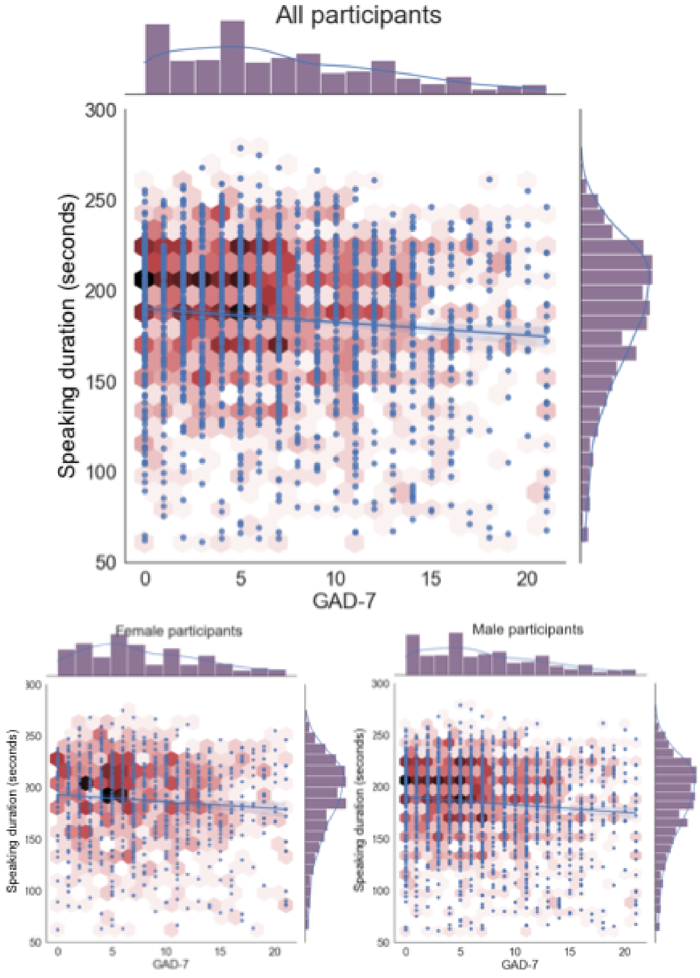
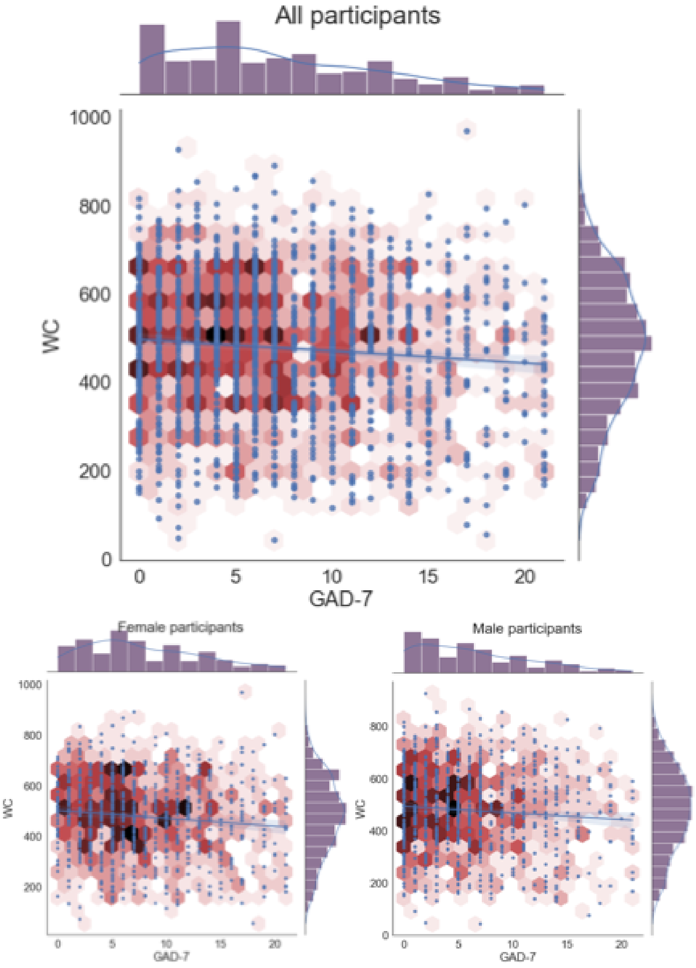


**Figure 3.** Word count (WC) versus Generalized Anxiety Disorder 7-item scale (GAD-7) scatter plot and distributions.



### *Acoustic Feature Correlation With the GAD-7*

Table 3 presents the correlation and *P* values for all the acoustic features (presented in the *Acoustic Features* section) that had *P* values above the 95% CI for the 3 data sets: all participants, female-only participants, and male-only participants. Again, it should be noted that all correlations were computed after controlling for age and personal income, whereas the calculations involving all participants also controlled for sex.

Table 4 reports results for features that previous work found to be statistically significant but for which we found no correlation in our sample. In our results, these features were not significantly associated with anxiety in any of the 3 data sets: all participants, female-only participants, and male-only participants.

Table 5 makes a direct comparison between previous work on the specific features (and their relation to anxiety) and the results from this study.

**Table 3.** Correlation of significant acoustic features with the Generalized Anxiety Disorder 7-item scale.

| Sample and feature | *r* | *P* value |
|---|---|---|
| **All samples (N=1744)** | | |
| Shimmer | 0.08 | <.001 |
| mfcc_std_2 | –0.08 | .002 |
| mfcc_std_3 | –0.07 | .002 |
| mfcc_mean_2 | –0.07 | .004 |
| f0_std | 0.06 | .01 |
| mfcc_std_5 | –0.06 | .01 |
| mfcc_std_4 | –0.05 | .03 |
| **Female samples (n=862)** | | |
| mfcc_std_3 | –0.10 | .002 |
| Shimmer | 0.10 | .004 |
| lpcc_std_6 | –0.09 | .008 |
| lpcc_std_4 | –0.09 | .008 |
| mfcc_mean_2 | –0.09 | .01 |
| Intensity_mean | –0.09 | .01 |
| mfcc_mean_1 | –0.09 | .01 |
| lpcc_std_10 | –0.07 | .03 |
| intensity_std | –0.07 | .03 |
| lpcc_std_12 | –0.07 | .04 |
| mfcc_mean_8 | 0.07 | .04 |
| lpcc_mean_4 | 0.07 | .049 |
| **Male samples (n=882)** | | |
| mfcc_std_2 | –0.09 | .005 |
| mfcc_std_5 | –0.09 | .01 |
| mfcc_mean_5 | –0.08 | .01 |
| f0_std | 0.07 | .03 |
| mfcc_std_4 | –0.07 | .04 |
| Shimmer | 0.07 | .04 |
| mfcc_std_11 | –0.07 | .046 |
| f1_mean | 0.07 | .047 |

**Table 4.** Correlation of acoustic features not found to be significant.

| Feature | Previous works | This study | | | | | |
|---|---|---|---|---|---|---|---|
| | | All samples | | Female samples | | Male samples | |
| | | *r* | *P* value | *r* | *P* value | *r* | *P* value |
| Jitter | Showed a significant increase from a neutral state to an anxious state [13] | 0.03 | .18 | –0.01 | .76 | 0.06 | .06 |
| ZCR-zPSD[a] | ZCR-zPSD was one of the top selected features using the Davies-Bouldin index–based feature selection [8] | 0.01 | .67 | –0.04 | .29 | 0.05 | .14 |
| Articulation rate | Patients with panic disorder spoke significantly slower (*P*<.001) during autobiographical talking than during script talking [22] | –0.01 | .64 | –0.05 | .12 | 0.02 | .55 |
| F1[b] SD | Showed a significant change between neutral state and anxious state [13] | –0.03 | .18 | –0.02 | .53 | –0.04 | .25 |
| F2[c] mean | Showed a significant change between neutral state and anxious state [13] | 0.004 | .85 | 0.04 | .26 | –0.04 | .22 |
| F2 SD | Showed a significant change between neutral state and anxious state [13] | 0.01 | .59 | 0.03 | .38 | –0.02 | .60 |
| F3[d] mean | Showed a significant change between neutral state and anxious state [13] | 0.02 | .49 | 0.04 | .21 | –0.01 | .72 |

[a] ZCR-zPSD: zero crossing rate for the *z* score of the power spectral density.

[b] F1: first formant.

[c] F2: second formant.

[d] F3: third formant.

**Table 5.** Comparison of previous works’ correlations with those of this study.

| Feature | Previous work | | This study | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | All samples | | Female samples | | Male samples | |
| | *r* | *P* value | *r* | *P* value | *r* | *P* value | *r* | *P* value |
| Speaking duration | –0.36 | <.01 | –0.12 | <.001 | –0.11 | <.001 | –0.13 | <.001 |
| MFCC[a]_std_1 | –0.36 | <.05 | 0.01 | .54 | 0.02 | .61 | 0.02 | .52 |
| F0[b]_mean | Female: 0.02; male: 0.72 | Female: 0.92; male: 0.002 | 0.02 | .37 | –0.03 | .33 | 0.06 | .06 |
| F0_SD | –0.24 | <.05 | 0.06 | .01 | 0.03 | .30 | 0.07 | .03 |
| Intensity mean | –0.2 | —[c] | –0.04 | .13 | –0.09 | .01 | 0.01 | .72 |

[a] MFCC: mel-frequency cepstral coefficient.

[b] F0: fundamental frequency.

[c] Not available.

### *Linguistic Feature Correlation With the GAD-7*

The quality of the transcript produced using Amazon’s AWS STT program [42] was analyzed by comparing the transcript produced from the task 1 audio with the actual *My Grandfather* passage. The word error rate was calculated, and the STT transcript had an average word error rate of 7% (SD 4.6%).

Table 6 presents the set of linguistic features (described in the *Linguistic Features* section) that had *P* values <.05 for the same 3 data sets: all participants, male-only participants, and female-only participants. Each section in the table is sorted in decreasing order of absolute value of correlation. As described previously, the partial correlations account for age and personal income across all data sets, and we also controlled for sex in the full data set.

**Table 6.** Correlation of significant Linguistic Inquiry and Word Count linguistic features with the Generalized Anxiety Disorder 7-item scale.

| Sample and feature | *r* | *P* value |
|---|---|---|
| **All samples (N=1744)** | | |
| AllPunc | 0.13 | <.001 |
| Period | 0.12 | <.001 |
| assent | 0.10 | <.001 |
| negemo | 0.10 | <.001 |
| relativ | –0.09 | <.001 |
| motion | –0.08 | <.001 |
| swear | 0.08 | <.001 |
| anger | 0.08 | <.001 |
| focusfuture | –0.07 | .003 |
| adverb | –0.07 | .004 |
| time | –0.07 | .004 |
| function | –0.07 | .005 |
| negate | 0.07 | .006 |
| prep | –0.06 | .007 |
| WPS[a] | –0.06 | .007 |
| anx | 0.06 | .008 |
| hear | 0.06 | .01 |
| death | 0.06 | .01 |
| ipron | –0.06 | .01 |
| see | –0.06 | .01 |
| affect | 0.06 | .02 |
| i | 0.05 | .02 |
| family | 0.05 | .02 |
| sad | 0.05 | .03 |
| ppron | 0.05 | .03 |
| space | –0.05 | .04 |
| article | –0.05 | .04 |
| leisure | 0.05 | .04 |
| friend | 0.05 | .047 |
| **Female samples (n=862)** | | |
| Period | 0.16 | <.001 |
| AllPunc | 0.14 | <.001 |
| adverb | –0.11 | <.001 |
| negemo | 0.11 | <.001 |
| anger | 0.11 | .002 |
| motion | –0.10 | .003 |
| assent | 0.10 | .004 |
| see | –0.09 | .006 |
| relativ | –0.09 | .006 |
| sad | 0.08 | .01 |
| Dic | –0.08 | .02 |

| Sample and feature | *r* | *P* value |
|---|---|---|
| power | 0.07 | .03 |
| WPS | –0.07 | .03 |
| death | 0.07 | .04 |
| percept | –0.07 | .046 |
| **Male samples (n=882)** | | |
| AllPunc | 0.13 | <.001 |
| assent | 0.11 | .001 |
| relativ | –0.10 | .002 |
| leisure | 0.10 | .002 |
| hear | 0.10 | .003 |
| swear | 0.10 | .004 |
| time | –0.10 | .004 |
| Apostro | 0.09 | .005 |
| power | –0.09 | .01 |
| ppron | 0.09 | .01 |
| Sixltr | –0.09 | .01 |
| anx | 0.08 | .01 |
| negate | 0.08 | .01 |
| negemo | 0.08 | .01 |
| article | –0.08 | .01 |
| Period | 0.08 | .02 |
| prep | –0.08 | .02 |
| focusfuture | –0.08 | .02 |
| family | 0.08 | .02 |
| ipron | –0.07 | .04 |
| affect | 0.07 | .04 |
| motion | –0.07 | .048 |

[a] WPS: words per sentence.

# Discussion

## Principal Findings

### *Overview*

Our central objective was to test specific acoustic and linguistic features of impromptu speech for their association with anxiety and to do so with a larger number of participants. In this section, we discuss the implications of the findings presented in the previous section, as well as the limitations of the study.

The results presented in the *Results* section quantified the relationship between features computed from recorded speech and the self-reported GAD-7 score using Pearson correlation coefficients, controlling for age and income. The results show several significant correlations between features extracted from speech and anxiety, which can help to inform future efforts in the automatic monitoring of anxiety. We discuss these in the following sections.

### *Recruitment and Data Inclusion*

Figure 1, the study recruitment flow chart, shows that the recruitment yield was 48.7% (2212/4542). Regarding the 51.3% (2330/4542) of participants who dropped out after accepting the study, we can only speculate as to why. Some may have been unwilling to have their words audio recorded or their full video recorded, and although the consent form makes this task clear, it may be that the participants who dropped out only really understood this when they saw their video on the screen.

We also conducted a missingness analysis on the 4.98% (256/4542) of samples excluded from the study (Multimedia Appendix 4). The results show that in the excluded data, the mention of words related to anxiety and those related to home had a significant positive correlation with anxiety and the count of longer words (>6 letters) was negatively correlated with anxiety. We found similar positive and negative correlations of these features in the 38.4% (1744/4542) samples included in our analysis. This indicates that excluding the 256 samples did not affect the correlation results.

### Demographics of Participants

The proportion of participants in the group with anxiety (those above the GAD-7 screening threshold of 10) was 30.96% (540/1744), which is much higher than the general population rate of approximately 10% [1]. This result, indicating that English speakers recruited from Prolific have elevated rates of anxiety and depression, is consistent with our prior studies using recruits from Prolific and suggests that this population exhibits a higher incidence of anxiety [23,49-51]. Table 1 sheds some light on this difference: it shows that a similar high fraction of participants self-reported on their Prolific profile that they have an ongoing mental health condition.

The demographic data listed in Table 1 provide several interesting insights into the recruited cohort with respect to the presence or absence of above-threshold GAD-7 scores. First, there was a significantly larger proportion of women in the group with anxiety than men. This is consistent with previous findings suggesting that anxiety is more prevalent in women than in men [43]. We feel that this confirms that it is useful to consider separate female-only and male-only data sets to avoid the bias introduced by sex when exploring features that may correlate with the GAD-7. For example, pitch (F0) would typically be higher for women, and as a result, sex effects could easily confound the association between pitch and anxiety.

The rows in Table 1 that show the proportions of participants classified as anxious and nonanxious by income suggest that there is a relationship between income and anxiety: the 2 very lowest categories of income show a disproportionately higher amount of anxiety. There is a downward trend in anxiety with income until the very last category, which is ≥£60,000 (US $82,200). It is interesting that above a certain income level, anxiety seems to increase, although this is consistent with prior studies on anxiety and income [45].

Similarly, with respect to age, younger participants were more likely to be in the group with anxiety, which is consistent with previous work [44].

### Posttask Self-report Anxiety Measure

As described in the *Anxiety Measures* section, we used the posttask self-reported anxiety measure as an internal check to see whether task 2 (the modified TSST task) induced more self-reported anxiety than task 1. A paired *t* test conducted on the 2 informal ratings of anxiety of the 2 tasks had a *P* value of <.001, indicating a significant difference and implying that task 2 induced greater anxiety. Recall that most of the prior work discussed in the *Related Work* section also used mood induction tasks.

### Amount of Speech

The results suggest that features related to the amount of speech that the participants delivered in response to task 2 had one of the highest correlations with their GAD-7 response across all the features explored in this work. In particular, 2 features captured this aspect: *speaking duration* and *word count*, as shown in Table 2 (their intercorrelation with each other is presented in Multimedia Appendix 6). In all cases, the negative direction of the correlation suggests that participants who spoke more tended to have lower GAD-7 scores. This result is consistent with previous work, as shown in the first data row of Table 5; however, our study gives a much lower Pearson correlation than prior work (*r*=0.12 in this study vs *r*=0.36 in the study by Laukka et al [16]). We speculate that the more anxious a person is, the less confidence they would have about their speech; therefore, perhaps, they speak less.

### Acoustic Features

The main purpose of this work was to explore how acoustic features relate to anxiety. We wanted to determine whether associations found in previous studies still hold with the larger sample size. Table 3 lists the features that have significant correlations, with *P*<.05, across all 3 data sets. The features with the strongest correlation in this set were *shimmer* on the all-sample data set and the SDs of the second and third MFCCs for the male-sample and female-sample data sets, respectively. We note that there are multiple parameters used in the extraction of MFCC features; therefore, a direct comparison of the specific MFCC features of our study with specific features of previous work is not possible as the prior work does not provide the exact parameters used to compute the MFCCs. The parameters used in this study are provided in the *Acoustic Features* section under the *Methods* section. That being said, in previous research, the fourth MFCC was the most significant among the 13 MFCC features in the study by Özseven et al [13] and the SD of the first MFCC in the study by Wörtwein et al [20] had a significant correlation (*r*=–0.36; *P*<.05) with an anxiety scale. These results, from both our study and previous work, suggest that signals of anxiety are present in the MFCC features.

The following features, listed as relevant in prior work, did not show significant correlations with the GAD-7: F2 and F3, jitter, ZCR-zPSD, and the articulation rate. Table 4 presents prior works' associations with anxiety regarding these features and the correlation values obtained in our study. It is important to note that in previous research, these features were noted as significant or relevant; however, no correlations with an indicator of anxiety were provided. This makes it difficult to compare directly with the correlations obtained in our study.

### Linguistic Features

Correlations between linguistic features extracted using the LIWC dictionaries [24] and the GAD-7 have been presented in the *Results* section. These had a higher correlation than the acoustic features, as presented in Table 6. The top LIWC category with the highest correlation in all the data sets is the count of punctuations. This includes the count of periods, which would indicate the number of separate sentences. The count of periods together with a negative correlation of words per sentence indicates that the use of shorter sentences is positively associated with anxiety.

Other LIWC categories with high correlation in the all-sample data set were negative emotion (*negemo*; eg, hurt, ugly, and nasty), anger (*anger*; eg, hate, kill, and annoyed), anxiety (*anx*; eg, worried and fearful), and sad (*sad*; eg, crying, grief, and sad). The anger, anxiety, and sad categories were constituent subsets of the negative emotion (*negemo*) category; that is, words counted under one of the anger, anxiety, or sad categories were also counted for the *negemo* category. The high intercorrelation with each other is shown in Multimedia Appendix 6. The *negemo* count had a higher correlation than these individual subcategories, suggesting that words related to anger, anxiety, and sad captured different dimensions of self-reported anxiety.

An LIWC category with a significant correlation that is present in the male-sample data set but not in the female-sample data set is the use of apostrophes (*apostro*), indicating that words with contractions (such as *I'll*) were positively associated with the GAD-7. In addition, only for men, function words, including personal pronouns (*ppron*), had a significant positive correlation with anxiety. We speculate that male individuals with anxiety might use personal pronouns (which include I, me, and mine) to divert their attention from the anxiety-inducing event and focus on themselves. More generally, the increased use of personal pronouns has been shown to occur in individuals with depression [52], a highly comorbid mental health illness with GAD (but not only for men).

Another differentiation between men and women occurs in the LIWC feature for words related to *power* (eg, superior and bully). The *power* count had a positive correlation with the GAD-7 for women and a negative correlation for men. We speculate that the negative correlation is somehow related to the stereotypical dominance behavior associated with men.

In prior work studying associations between LIWC scores and anxiety, words related to anxiety and first-person singular pronouns were shown to be significantly associated with social anxiety [25], similar to our results. The same work has also shown that perceptual process words (see, hear, and feel) are significantly associated with anxiety, which does not align with our results. For example, the LIWC category for *see* has a negative correlation in both the all-sample and the female-sample data sets (as shown in Table 6). However, in the study by Di Matteo et al [23], the category *see* had a positive correlation ($r$=0.31; $P$=.02) with a social anxiety measure. We speculate that the use of perceptual process words (eg, *see*) might be a differentiating factor between social anxiety and GAD as it was positively correlated in the former and negatively correlated in the latter. By contrast, the LIWC category for the perceptual process word *hear* had a positive correlation in both the all-sample and the male-sample data set (also shown in Table 6). Notice that both *see* and *hear* are perceptual processes; however, the category for *see* is significant for women, whereas the category for *hear* is significant for men.

Furthermore, in prior work, death-related words were shown to have a positive correlation with anxiety [23]. Our results (as shown in Table 6) show a similar trend where death-related words had a significant positive correlation in the male-sample and all-sample data sets. However, a significant correlation was not observed in the female-sample data set.

The fact that there are several single-word categories that have significant correlations suggests that techniques that are able to look at multiple word meanings may have greater potential in making predictions.

## Limitations

A limitation of this study is the use of self-report measures to assess GAD. Self-report measures, by nature, are subjective opinions that individuals have about themselves while filling out the questionnaires and may not completely capture clinical symptoms. In this study, we took these self-report questionnaires as the true label of the audio samples. However, we believe that this is a good first step that gave us encouraging preliminary results. A psychiatric diagnosis would be an improved label but is clearly much more expensive to acquire.

A further limitation of this study is the selection bias that might be introduced during the recruitment of the participants. As presented in Figure 1, only 48.7% (2212/4542) of the participants who initially accepted the offer from Prolific to participate finished the study. We were not able to collect the GAD-7 scores of the participants who did not complete the study; therefore, we do not know their levels of anxiety. It is possible that these participants had higher levels of anxiety, which caused them to drop out of the study.

Another limitation concerns the differences in the recording devices and recording locations of the participants performing each task. Ideally, we would want every sample to be recorded using the same microphone in the same location with the same acoustics. This would reduce the potential bias introduced by different factors such as recording quality or background noise. At the same time, in a real-life scenario where an application to detect anxiety might be deployed, the recording equipment and the location will likely differ for everyone. Hence, this limitation could be unavoidable, and it might even be essential to take these types of differences into consideration.

## Conclusions

We present results from a large-N study examining the relationship between speech and GAD. Our data collection relied on participants using home recording devices, hence capturing variations in acoustic environments, which will need to be factored in when deploying tools for the detection of mental health disorders in the wild. Our goal was to provide a useful benchmark for future research by assessing the extent to which results from previous research are generalizable to our data collection approach and larger data set. We tested the most common acoustic and linguistic features associated with anxiety in previous studies and provided detailed correlation tables broken down by demographics.

Our findings are decidedly mixed. On the one hand, with our larger data set, we found modest correlations between anxiety and several features of speech, including speaking duration and acoustic features such as MFCCs, LPCCS, shimmer, F0, and F1. However, other features shown to correlate with anxiety elsewhere—including F2 and F3, jitter, and ZCR-zPSD—were not significantly associated with anxiety in our study. Although these null findings do not entirely rule out the potential of more sophisticated learning models for this task, we believe that researchers should be wary of inherent difficulties. Readers should also note that our data collection already sidestepped additional challenges that we expected to influence the detection of anxiety disorders from speech, such as variations in accents, dialects, and spoken language. On the other hand, we found statistically significant correlations for a subset of speech features from previous research. This suggests that there may be a fundamental pathway between anxiety and the production of speech, one that is robust enough to be generalized to the population.

Future investigations could explore whether features of speech from task 1 (simple reading of a passage) exhibit correlations with the GAD-7 or whether these features could be used as a control for the features of task 2 (the modified TSST task). It may also be informative to separate out different age groups (eg, younger and older) to see whether there is a specific impact of speech features on the GAD-7.

## Acknowledgments

This research was funded by a University of Toronto XSeed Grant, Natural Sciences and Engineering Research Council of Canada Discovery Grant (RGPIN-2019-04395), and Social Sciences and Humanities Research Council Partnership Engage Grant (892-2019-0011).

## Conflicts of Interest

WS is an employee of Winterlight Labs and hold equity within the company, and DDD is a former employee of Winterlight Labs.

Multimedia Appendix 1
Web application screenshot.
[PDF File (Adobe PDF File), 444 KB - mental_v9i7e36828_app1.pdf ]

Multimedia Appendix 2
My Grandfather passage.
[PDF File (Adobe PDF File), 28 KB - mental_v9i7e36828_app2.pdf ]

Multimedia Appendix 3
Speech encouragement statements.
[PDF File (Adobe PDF File), 34 KB - mental_v9i7e36828_app3.pdf ]

Multimedia Appendix 4
Excluded data analysis.
[PDF File (Adobe PDF File), 43 KB - mental_v9i7e36828_app4.pdf ]

Multimedia Appendix 5
Correlation between demographics and acoustic and linguistic features.
[PDF File (Adobe PDF File), 103 KB - mental_v9i7e36828_app5.pdf ]

Multimedia Appendix 6
Significant feature intercorrelations of the all-sample data set.
[XLSX File (Microsoft Excel File), 21 KB - mental_v9i7e36828_app6.xlsx ]

Multimedia Appendix 7
Significant feature intercorrelations of the female-sample data set.
[XLSX File (Microsoft Excel File), 17 KB - mental_v9i7e36828_app7.xlsx ]

Multimedia Appendix 8
Significant feature intercorrelations of the male-sample data set.
[XLSX File (Microsoft Excel File), 18 KB - mental_v9i7e36828_app8.xlsx ]

## References

1. Mental Health - Anxiety Disorders. Public Health Canada. 2009. URL: https://www.canada.ca/en/health-canada/services/healthy-living/your-health/diseases/mental-health-anxiety-disorders.html [accessed 2022-01-24]
2. Roberge P, Fournier L, Duhoux A, Nguyen CT, Smolders M. Mental health service use and treatment adequacy for anxiety disorders in Canada. Soc Psychiatry Psychiatr Epidemiol 2011 Apr;46(4):321-330. [doi: 10.1007/s00127-010-0186-2] [Medline: 20217041]

3. Koerner N, Dugas MJ, Savard P, Gaudet A, Turcotte J, Marchand A. The economic burden of anxiety disorders in Canada. Can Psychol 2004;45(3):191-201. [doi: 10.1037/h0088236]
4. Hidalgo RB, Sheehan DV. Generalized anxiety disorder. Handb Clin Neurol 2012;106:343-362. [doi: 10.1016/B978-0-444-52002-9.00019-X] [Medline: 22608630]
5. Hoehn-Saric R, McLeod DR. The peripheral sympathetic nervous system. Its role in normal and pathologic anxiety. Psychiatr Clin North Am 1988 Jun;11(2):375-386. [Medline: 3047706]
6. Thompson AR. Pharmacological agents with effects on voice. Am J Otolaryngol 1995;16(1):12-18. [doi: 10.1016/0196-0709(95)90003-9] [Medline: 7717466]
7. Barrett LF. How Emotions Are Made: The Secret Life of the Brain. Boston, MA, USA: Mariner Books; 2018.
8. McGinnis EW, Anderau SP, Hruschak J, Gurchiek RD, Lopez-Duran NL, Fitzgerald K, et al. Giving voice to vulnerable children: machine learning analysis of speech detects anxiety and depression in early childhood. IEEE J Biomed Health Inform 2019 Nov;23(6):2294-2301 [FREE Full text] [doi: 10.1109/JBHI.2019.2913590] [Medline: 31034426]
9. Buske-Kirschbaum A, Jobst S, Wustmans A, Kirschbaum C, Rauh W, Hellhammer D. Attenuated free cortisol response to psychosocial stress in children with atopic dermatitis. Psychosom Med 1997;59(4):419-426. [doi: 10.1097/00006842-199707000-00012] [Medline: 9251162]
10. Ali S, Tanweer S, Khalid S, Rao N. Mel frequency cepstral coefficient: a review. In: Proceedings of the 2nd International Conference on ICT for Digital, Smart, and Sustainable Development. 2020 Presented at: ICIDSSD '20; February 27-28, 2020; New Delhi, India. [doi: 10.4108/eai.27-2-2020.2303173]
11. Davies DL, Bouldin DW. A cluster separation measure. IEEE Trans Pattern Anal Mach Intell 1979 Apr;PAMI-1(2):224-227. [doi: 10.1109/tpami.1979.4766909]
12. Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995 Sep;20(3):273-297. [doi: 10.1007/BF00994018]
13. Özseven T, Dügenci M, Doruk A, Kahraman H. Voice traces of anxiety: acoustic parameters affected by anxiety disorder. Arch Acoust 2018;43(4):625-636. [doi: 10.24425/AOA.2018.125156]
14. Weeks JW, Lee CY, Reilly AR, Howell AN, France C, Kowalsky JM, et al. "The Sound of Fear": assessing vocal fundamental frequency as a physiological indicator of social anxiety disorder. J Anxiety Disord 2012 Dec;26(8):811-822. [doi: 10.1016/j.janxdis.2012.07.005] [Medline: 23070030]
15. Julian LJ. Measures of anxiety: State-Trait Anxiety Inventory (STAI), Beck Anxiety Inventory (BAI), and Hospital Anxiety and Depression Scale-Anxiety (HADS-A). Arthritis Care Res (Hoboken) 2011 Nov;63 Suppl 11:S467-S472 [FREE Full text] [doi: 10.1002/acr.20561] [Medline: 22588767]
16. Laukka P, Linnman C, Åhs F, Pissiota A, Frans Ö, Faria V, et al. In a nervous voice: acoustic analysis and perception of anxiety in social phobics' speech. J Nonverbal Behav 2008 Jul 18;32(4):195-214. [doi: 10.1007/s10919-008-0055-9]
17. Spielberger CD. State-trait anxiety inventory. In: Weiner IB, Craighead WE, editors. The Corsini Encyclopedia of Psychology. Hoboken, NJ, USA: John Wiley & Sons; 2010.
18. Albuquerque L, Valente AR, Teixeira A, Figueiredo D, Sa-Couto P, Oliveira C. Association between acoustic speech features and non-severe levels of anxiety and depression symptoms across lifespan. PLoS One 2021 Apr 8;16(4):e0248842 [FREE Full text] [doi: 10.1371/journal.pone.0248842] [Medline: 33831018]
19. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [doi: 10.1111/j.1600-0447.1983.tb09716.x] [Medline: 6880820]
20. Wortwein T, Morency L, Scherer S. Automatic assessment and analysis of public speaking anxiety: a virtual audience case study. In: Proceedings of the 2015 International Conference on Affective Computing and Intelligent Interaction. 2015 Presented at: ACII '15; September 21-24, 2015; Xi'an, China p. 187-193. [doi: 10.1109/acii.2015.7344570]
21. Gilkinson H. Social fears as reported by students in college speech classes*. Speech Monogr 1942 Jan;9(1):141-160. [doi: 10.1080/03637754209390068]
22. Hagenaars MA, van Minnen A. The effect of fear on paralinguistic aspects of speech in patients with panic disorder with agoraphobia. J Anxiety Disord 2005;19(5):521-537. [doi: 10.1016/j.janxdis.2004.04.008] [Medline: 15749571]
23. Di Matteo D, Wang W, Fotinos K, Lokuge S, Yu J, Sternat T, et al. Smartphone-detected ambient speech and self-reported measures of anxiety and depression: exploratory observational study. JMIR Form Res 2021 Jan 29;5(1):e22723 [FREE Full text] [doi: 10.2196/22723] [Medline: 33512325]
24. Pennebaker JW, Boyd RL, Jordan K, Blackburn K. The development and psychometric properties of LIWC2015. The University of Texas at Austin. 2015. URL: https://repositories.lib.utexas.edu/bitstream/handle/2152/31333/LIWC2015_LanguageManual.pdf [accessed 2022-01-24]
25. Anderson B, Goldin PR, Kurita K, Gross JJ. Self-representation in social anxiety disorder: linguistic analysis of autobiographical narratives. Behav Res Ther 2008 Oct;46(10):1119-1125 [FREE Full text] [doi: 10.1016/j.brat.2008.07.001] [Medline: 18722589]
26. Palan S, Schitter C. Prolific.ac—a subject pool for online experiments. J Behav Exp Finance 2018 Mar;17:22-27. [doi: 10.1016/j.jbef.2017.12.004]
27. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]

28. Reilly J, Fisher JL. Sherlock Holmes and the strange case of the missing attribution: a historical note on "The Grandfather Passage". J Speech Lang Hear Res 2012 Feb;55(1):84-88. [doi: 10.1044/1092-4388(2011/11-0158)] [Medline: 22354714]
29. Kirschbaum C, Pirke KM, Hellhammer DH. The 'Trier Social Stress Test'-a tool for investigating psychobiological stress responses in a laboratory setting. Neuropsychobiology 1993;28(1-2):76-81. [doi: 10.1159/000119004] [Medline: 8255414]
30. Gerra G, Zaimovic A, Zambelli U, Timpano M, Reali N, Bernasconi S, et al. Neuroendocrine responses to psychological stress in adolescents with anxiety disorder. Neuropsychobiology 2000;42(2):82-92. [doi: 10.1159/000026677] [Medline: 10940763]
31. Jezova D, Makatsori A, Duncko R, Moncek F, Jakubek M. High trait anxiety in healthy subjects is associated with low neuroendocrine activity during psychosocial stress. Prog Neuropsychopharmacol Biol Psychiatry 2004 Dec;28(8):1331-1336. [doi: 10.1016/j.pnpbp.2004.08.005] [Medline: 15588760]
32. Endler NS, Kocovski NL. State and trait anxiety revisited. J Anxiety Disord 2001;15(3):231-245. [doi: 10.1016/s0887-6185(01)00060-3] [Medline: 11442141]
33. First MB. Structured Clinical Interview for DSM-IV Axis I Disorders SCID-I: Clinician Version. Washington, DC, USA: American Psychiatric Press; 1997.
34. Kliper R, Portuguese S, Weinshall D. Prosodic analysis of speech and the underlying mental state. In: Proceedings of the 5th International Conference on the Pervasive Computing Paradigms for Mental Health. 2015 Presented at: MindCare '15; September 24-25, 2015; Milan, Italy p. 52-62. [doi: 10.1007/978-3-319-32270-4_6]
35. Pennebaker JW, Mehl MR, Niederhoffer KG. Psychological aspects of natural language use: our words, our selves. Annu Rev Psychol 2003;54:547-577. [doi: 10.1146/annurev.psych.54.101601.145041] [Medline: 12185209]
36. Sabahi S. my-voice-analysis. GitHub. 2021 Sep 1. URL: https://github.com/Shahabks/my-voice-analysis [accessed 2022-01-24]
37. Lenain R, Weston J, Shivkumar A, Fristed E. Surfboard: audio feature extraction for modern machine learning. In: Proceedings of the 2020 Inter Speech. 2020 Presented at: IS '20; October 25-29, 2020; Shanghai, China p. 2917-2921. [doi: 10.21437/interspeech.2020-2879]
38. McFee B, Raffel C, Liang D, Ellis D, McVicar M, Battenberg E, et al. librosa: audio and music signal analysis in Python. In: Proceedings of the 14th Python in Science Conference. 2015 Presented at: SciPy '15; July 6-12, 2015; Austin, TX, USA p. 18-24. [doi: 10.25080/majora-7b98e3ed-003]
39. Aalto D, Malinen J, Vainio M. Formants. In: Aronoff M, editor. Oxford Research Encyclopedia of Linguistics. Oxford, UK: Oxford University Press; 2018.
40. Silber-Varod V, Kreiner H, Lovett R, Levi-Belz Y, Amir N. Do social anxiety individuals hesitate more? The prosodic profile of hesitation disfluencies in Social Anxiety Disorder individuals. In: Proceedings of the 8th Speech Prosody. 2016 Presented at: SpeechProsody '16; May 31-June 3, 2016; Boston, MA, USA p. 1211-1215. [doi: 10.21437/speechprosody.2016-249]
41. Fuller BF, Horii Y, Conner DA. Validity and reliability of nonverbal voice measures as indicators of stressor-provoked anxiety. Res Nurs Health 1992 Oct;15(5):379-389. [doi: 10.1002/nur.4770150507] [Medline: 1529122]
42. Hashemipour S, Ali M. Amazon Web Services (AWS) – an overview of the on-demand cloud computing platform. In: Proceedings of the 3rd EAI International Conference on Emerging Technologies in Computing. 2020 Presented at: iCETiC '20; August 19-20, 2020; London, UK p. 40-47. [doi: 10.1007/978-3-030-60036-5_3]
43. Cameron OG, Hill EM. Women and anxiety. Psychiatr Clin North Am 1989 Mar;12(1):175-186. [doi: 10.1016/s0193-953x(18)30459-3]
44. Krasucki C, Howard R, Mann A. The relationship between anxiety disorders and age. Int J Geriatr Psychiatry 1998 Feb;13(2):79-99. [doi: 10.1002/(sici)1099-1166(199802)13:2<79::aid-gps739>3.0.co;2-g] [Medline: 9526178]
45. Dijkstra-Kersten SM, Biesheuvel-Leliefeld KE, van der Wouden JC, Penninx BW, van Marwijk HW. Associations of financial strain and income with depressive and anxiety disorders. J Epidemiol Community Health 2015 Jul;69(7):660-665. [doi: 10.1136/jech-2014-205088] [Medline: 25636322]
46. Pellegrini T, Hämäläinen A, de Mareüil PB, Tjalve M, Trancoso I, Candeias S, et al. A corpus-based study of elderly and young speakers of European Portuguese: acoustic correlates and their impact on speech recognition performance. In: Proceedings of the 2013 Inter Speech. 2013 Presented at: InterSpeech '13; August 25-29, 2013; Lyon, France p. 853-856. [doi: 10.21437/interspeech.2013-241]
47. Farrow K, Grolleau G, Mzoughi N. What in the word! The scope for the effect of word choice on economic behavior. Kyklos 2018 Oct 10;71(4):557-580. [doi: 10.1111/kykl.12186]
48. Baba K, Shibata R, Sibuya M. Partial correlation and conditional correlation as measures of conditional independence. Aust NZ J Stat 2004 Dec;46(4):657-664. [doi: 10.1111/j.1467-842x.2004.00360.x]
49. Di Matteo D, Fotinos K, Lokuge S, Yu J, Sternat T, Katzman MA, et al. The relationship between smartphone-recorded environmental audio and symptomatology of anxiety and depression: exploratory study. JMIR Form Res 2020 Aug 13;4(8):e18751 [FREE Full text] [doi: 10.2196/18751] [Medline: 32788153]
50. Di Matteo D, Fotinos K, Lokuge S, Mason G, Sternat T, Katzman MA, et al. Automated screening for social anxiety, generalized anxiety, and depression from objective smartphone-collected data: cross-sectional study. J Med Internet Res 2021 Aug 13;23(8):e28918 [FREE Full text] [doi: 10.2196/28918] [Medline: 34397386]

51. Di Matteo D. Inference of anxiety and depression from smartphone-collected data. University of Toronto. 2021. URL: https://tspace.library.utoronto.ca/bitstream/1807/108894/1/Di_Matteo_Daniel_202111_PhD_thesis.pdf [accessed 2022-01-24]
52. Eichstaedt JC, Smith RJ, Merchant RM, Ungar LH, Crutchley P, Preoţiuc-Pietro D, et al. Facebook language predicts depression in medical records. Proc Natl Acad Sci U S A 2018 Oct 30;115(44):11203-11208 [FREE Full text] [doi: 10.1073/pnas.1802331115] [Medline: 30322910]

## Abbreviations

**F0:** fundamental frequency
**F1:** first formant
**F2:** second formant
**F3:** third formant
**GAD:** generalized anxiety disorder
**GAD-7:** Generalized Anxiety Disorder 7-item scale
**LIWC:** Linguistic Inquiry and Word Count
**LPCC:** linear prediction cepstral coefficient
**MFCC:** mel-frequency cepstral coefficient
**SAD:** social anxiety disorder
**STT:** speech-to-text
**TSST:** Trier Social Stress Test
**ZCR-zPSD:** zero crossing rate for the z score of the power spectral density

*Edited by J Torous; submitted 28.01.22; peer-reviewed by E McGinnis, R McGinnis, V Yadav; comments to author 25.03.22; revised version received 27.04.22; accepted 23.05.22; published 08.07.22.*

*Please cite as:*
*Teferra BG, Borwein S, DeSouza DD, Simpson W, Rheault L, Rose J*
*Acoustic and Linguistic Features of Impromptu Speech and Their Association With Anxiety: Validation Study*
*JMIR Ment Health 2022;9(7):e36828*
*URL: https://mental.jmir.org/2022/7/e36828*
*doi:10.2196/36828*
*PMID:35802401*

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org