

Original Paper

# Effectiveness of Mental Health Apps for Distress During COVID-19 in US Unemployed and Essential Workers: Remote Pragmatic Randomized Clinical Trial

Katherine Anne Comtois<sup>1</sup>, PhD; Felicia Mata-Greve<sup>1</sup>, PhD; Morgan Johnson<sup>1</sup>, MSc; Michael D Pullmann<sup>1</sup>, PhD; Brittany Mosser<sup>1,2</sup>, MSW; Patricia Arean<sup>1,2</sup>, PhD

<sup>1</sup>Advanced Laboratories for Accelerating the Reach and Impact of Treatments for Youth and Adults with Mental Illness, Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, United States

<sup>2</sup>Conducting Research to Enhance Assessment and Treatment Through Innovation in Mental Health Lab, Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, United States

**Corresponding Author:**

Katherine Anne Comtois, PhD

Advanced Laboratories for Accelerating the Reach and Impact of Treatments for Youth and Adults with Mental Illness

Department of Psychiatry and Behavioral Sciences

University of Washington

1959 Pacific Stree NE

Seattle, WA, 98195

United States

Phone: 1 14152727222

Email: [parean@uw.edu](mailto:parean@uw.edu)

## Abstract

**Background:** During the COVID-19 pandemic, the general public was concerned about the mental health impacts of unemployment due to COVID-19 and the stress essential workers experienced during this time. Several reports indicated that people in distress were turning to digital technology, but there was little evidence about the impact of these tools on mitigating distress.

**Objective:** This study seeks to determine the acceptability, feasibility, usability, and effectiveness of mobile mental health apps for decreasing mental health symptoms in essential workers and unemployed individuals with suicide risk.

**Methods:** We recruited participants who indicated that they were unemployed because of COVID-19 or were COVID-19–designated essential workers. Participants were randomized to 1 of 4 free commercial mobile apps for managing distress that were (1) highly rated by PsyberGuide and (2) met the criteria for intervention features these participants indicated were desirable in a previous survey. Participants used the apps for 4 weeks and completed baseline and 4-week self-assessments of depression, anxiety emotional regulation, and suicide risk.

**Results:** We found no differences between the apps in any outcome but did find significant changes in depression and anxiety over time (Patient Health Questionnaire [PHQ]-9: estimate=-1.5, SE 0.2, 95% CI -1.1 to -1.8,  $P<.001$ ; Generalized Anxiety Disorder Scale [GAD]-7: estimate=-1.3, SE 0.2, 95% CI -1.0 to -1.6,  $P<.001$ ). We found no significant changes in suicidal behavior (Suicide Behaviors Questionnaire-Revised [SBQ-R]) or emotional regulation (Difficulties in Emotion Regulation Scale – Short Form [DERS-SF]) for the 4 weeks. We did find a significant dose-response pattern for changes in depression and anxiety. Using the app at least once a week resulted in greater improvements in treatment conditions over time on depression (estimate=-0.6, SE 0.2, 95% CI 1.0-0.2,  $P=.003$ ) and anxiety (estimate=0.1, SE 0.2, 95% CI 0.4-0.6,  $P=.78$ ). There was no association between app frequency and changes in suicidal behavior (SBQ-R) or emotional regulation (DERS-SF). We further found a significant difference between the conditions with regard to app usability, with the control app being the most usable (mean<sub>Beautiful Mood</sub> 72.9, SD 16.7; mean<sub>COVID Coach</sub> 71.2, SD 15.4; mean<sub>Calm</sub> 66.8, SD 17.3; mean<sub>7 Cups</sub> 65.2, SD 17.7). We found no significant differences for app acceptability or appropriateness.

**Conclusions:** Few studies have evaluated prospectively the utility and usability of commercial apps for mood. This study found that free, self-guided commercial mobile mental health apps are seen as usable, but no one app is superior to the other. Although

we found that regular use is indicated for effects on depression and anxiety to occur in those who are more symptomatic, regression to the mean cannot be ruled out.

**Trial Registration:** ClinicalTrials.gov NCT04536935; <https://tinyurl.com/mr36zx3s>

(*JMIR Mhealth Uhealth* 2022;10(11):e41689) doi: [10.2196/41689](https://doi.org/10.2196/41689)

## KEYWORDS

COVID-19; COVID; coronavirus; pandemic; SARS-CoV-2; essential worker; suicide; suicidal; commercial app; mental health apps; health app; mental health; mHealth; mobile health; occupational health; employee; employment; unemployed; worker; job; depression; anxiety; stress; distress; mobile app; RCT; pragmatic trial; randomized; health care worker; health care provider; frontline staff

## Introduction

### Background

Access to mental health care by essential workers and the people unemployed due to COVID-19–related business closures and social distancing policies has been challenging [1-3]. To address this problem, health care organizations have created free mobile apps for stress related to COVID-19. Although overall app use during COVID-19 has been low (16%) [4,5], technology companies report substantial increases in the use of their tools.

There is limited information about the effectiveness of mental health apps, particularly free, self-guided commercial apps. Research on self-guided apps is mixed, with some studies finding them to be minimally effective [6-8] and others reporting beneficial effects; we note here that most evidence points to the superiority of coach-based apps for depression and anxiety outcomes, but effect sizes for self-guided apps are still notable [9] and offer an opportunity for stress management in populations that do not have the financial or time resources to avail themselves of coaching services [10]. It is important to also note here that most studies that find positive effects use research grade tools with a paid participant pool and are typically not available to the public. Many commercial apps do include principals and features that are similar to research grade tools; however, there remains skepticism about the effectiveness of these derivations [11]. This has led to the need to create app review resources, such as One Mind PsyberGuide [12] and the American Psychiatry Association’s App Advisor [13], which provide ratings of app effectiveness, transparency, and usability. Still, evidence for free commercial apps is limited, and calls for additional research [14,15], particularly in the context of COVID-19 [16], have been made.

### Objective

We previously reported on a large-scale survey of essential workers and people unemployed due to COVID-19 for their preferences for mobile apps for mood management [4]. In this study, we found that participants had strong preferences for apps that focus on mindfulness approaches, information about coping with COVID-19, symptom tracking, and connection with others. In this pragmatic clinical trial, we randomized 838 of these participants who indicated they were depressed or anxious or had suicidal thoughts in order to use 1 of 4 commercial apps for 4 weeks. We selected the 4-week time frame because in our past research, we found that this is an optimal dose of digital mental health in a distressed sample [17]

and other research has found that this is the length of time participants tend to engage with these tools [18,19]. Thus, we are interested in addressing the issues of app use and outcomes pragmatically, as it would occur in actual practice. The main objectives of this study were:

- Determine whether users of these apps show significant improvement in anxiety, depression, emotion regulation, and suicide risk.
- Identify differences between the apps in use, usability, and acceptability.
- Determine whether there is a dose-response relationship such that the frequency of app use is positively associated with improvement in depression, anxiety, emotional regulation, and suicide risk.
- Identify outcome differences between apps in this dose-response relationship.

## Methods

### Recruitment and Safeguards Against Bad Actors

Participants were recruited nationally via Prolific, an online research platform that includes several safeguards to preserve data quality [20-22] and minimize bad actors and has been shown to be reliable, efficient, and affordable for remote data collection for behavioral research [23]. Participants provided electronic informed consent prior to study completion. Additional survey safeguards were an attention check [24] and a review of open-ended items to screen out autofilled and nonsensical responses.

### Ethical Approval

The study received ethical approval from the University of Washington Institutional Review Board (STUDY00010842). In the consent, participants were explained the purpose of the study, that it would be randomized to 1 of 4 mobile apps, and that they would be asked to complete surveys before treatment began and 4 weeks later. Participants were also told how data were stored and managed and approximately how long each survey would take.

### Participants

Participants for this study were recruited from a larger study [4], which included a convenience sample of approximately 2000 participants that self-identified as COVID-19–designated essential workers or unemployed due to COVID-19 social distancing policies or COVID-19–related business closures. To identify as 1 of these 2 groups, participants responded to the

following 2 questions: (1) Are you considered an essential worker during the COVID-19 pandemic? (2) Have you become unemployed as a result of the COVID-19 pandemic?

To be eligible for this study, inclusion criteria included (1) previously granting permission to be recontacted for future research; (2) age  $\geq 19$  years, living in the United States, and English speaking; (3) access to a mobile device; and (4) report of depression (Patient Health Questionnaire [PHQ]-2 score  $\geq 3$ ) [25], anxiety (Generalized Anxiety Disorder Scale [GAD]-2 score  $\geq 3$ ) [26], risk for suicidal behaviors (Suicide Behaviors Questionnaire-Revised [SBQ-R] score  $\geq 7$ ) [27], or a history of past suicide attempt [28]. Participants were offered crisis management resources when they endorsed the ninth item of the PHQ-9 or were over the cut-off for the SBQ-R.

### Study Timeline

Participants were recruited from October through December 2020, during the middle of the initial COVID-19 variant, and shortly after vaccines were available to the public. Additionally, most states (with few exceptions) were continuing to institute public closures of restaurants, gyms, and other enclosed public places, meaning the unemployment rate due to COVID-19 was still quite high. Hospital censuses were at historically high rates, and essential workers were still mandated to wear protective gear. Thus, the sample is representative of people living under peak pandemic conditions. Participants were randomized after completing a web-based baseline assessment of mood and paid US \$1 (see the Measures section). Participants were randomized to 1 of 4 apps and asked to use their assigned app as instructed by the developers. Participants completed a web-based posttreatment survey at 4 weeks postrandomization and app assignment. After completing follow-up, participants were compensated US \$4.

### Mobile Interventions and Attention Control

This remote pragmatic clinical trial used simple randomization with parallel assignment comparing 3 active apps to an attention control app. This study meets the definition of a pragmatic trial in that the study was designed to test the effects of mobile apps for depression, anxiety, emotion regulation, and suicide as they are typically used by the general public [29]. In pragmatic trials, the intent is to determine the effect of existing treatments in the context of real-world use compared to existing treatment options. In such trials, the control condition is not a placebo, which is not usually part of standard care [30]. Although a waitlist may be appropriate for a pragmatic trial, waitlist controls are appropriate only when this is part of usual practice and if they are ethically sound; however, previous research has found internal validity issues with waitlist controls, and in the context of self-guided commercial digital mental health, there is no waitlist control [31]. In our sample, which consisted of participants at risk for suicide, neither a placebo nor waitlist controls were ethical choices [32]. Thus, our decision to use an attention control app was based on what is considered appropriate for pragmatic trials of this nature in potentially high-risk populations [33,34].

We selected apps based on the following criteria: (1) they were free; (2) reflected desired app features during COVID-19, as

identified in the survey study [4]; and (3) had good ratings on PsyberGuide [4]. The 3 active app interventions included (1) meditation (Calm), (2) COVID-19 coping (COVID Coach), and (3) chat and positive psychology (7 Cups of Tea). The attention control app used only mood tracking (Beautiful Mood) and did not include any intervention elements the other apps possessed (mindfulness meditation, emotional coping skills, social connection, or positive psychology approaches). Participants were randomized by study staff using random allocation functions in Microsoft Excel and received their app assignment through a URL to Google Play Store or Apple App Store. Participants confirmed app download prior to receiving compensation. Participants were blinded to the study hypotheses but not condition.

### Rationale for the 4-Week Intervention Timeline

We feel it is important to note that although mental health apps are based on evidence-based treatment approaches, people use apps differently than the way they use traditional mental health services [35]. The optimal dose of mobile mental health apps is measured in the frequency of use rather than the number of weeks of use, and research shows that considerable improvement in mood and function can occur rapidly with digital mental health tools and as early as after 2 weeks of use [17,36,37]. We acknowledge that although other randomized clinical studies do show the greatest impact at 8 weeks [38], the general population tends to initially engage with digital mental health apps frequently over the course of 2 weeks, with notable disengagement by 4 weeks [11,18,19]. Based on the literature from the informatics field on typical engagement patterns with digital health tools in general, this is a common pattern of engagement and may mean the user has met their goal [18].

### Measures

All data collected for this study are considered sensitive. We did not collect or store names, addresses, locations, IP addresses, or other digital identifiers. All survey data, including demographics, were immediately stored behind secure firewalls on servers at the University of Washington School of Medicine. The survey was developed by the study's lead investigators (authors PAA and KAM), measures were selected for their validity and reliability, and we selected those measures that had been validated for online use. The survey was programmed into REDCap, a web-based survey program developed by Vanderbilt University [39]. It has been used extensively for clinical research and is Health Insurance Portability and Accountability Act (HIPAA) compliant, highly secure, and intuitive to use. After the survey was built, we tested it with research group members naive to the study for readability, programming bugs, and time to completion.

### Demographics

Participants provided information about age, race and ethnicity, gender identity, sexual orientation, education, income, and living situation. We used similar questions to those in the US Census categories [40]. This survey has been used successfully in other online studies [4,41]. See [Multimedia Appendix 1](#). Race, ethnicity, and gender were assessed because mental health disparities were present in these groups [42,43].

### Primary Clinical Outcomes

Participants completed measures of depression (PHQ-9) [25,44,45], anxiety (GAD-7) [26,46,47], emotion dysregulation (Difficulties in Emotion Regulation – Short Form [DERS-SF]) [48], and suicidal behaviors (SBQ-R) [27] at baseline and follow-up. The PHQ-9 score ranges from 0 to 27, with 0-4 indicating no depression, 5-9 indicating mild depression, 10-20 indicating moderate depression, and a score >20 indicating severe depression. The GAD-7 is scored from 0 to 21, with 0-4 indicating no anxiety, 5-9 indicating mild anxiety, 10-14 indicating moderate anxiety, and 15-21 indicating severe anxiety. The DER-SF is scored from 1 to 180 and, while showing strong psychometric properties in clinical populations, does not have a clinical cut-off. The SBQ-R is scored from 3 to 18 and has a nonclinical cut-off of 7 and a clinical cut-off of 8 for elevated suicide risk [49].

The PHQ-9 and GAD-7 have been used successfully as online survey instruments and have been validated as online instruments [50,51]. Although the DERS and SBQ-R have been used in online survey research [52,53], to the best of our knowledge, no formal tests of validity have been conducted. We still elected to use these scales as there is no existing validated instrument for emotion regulation and suicide behaviors and because, of the existing scales, these have the best psychometric properties, are valid and reliable across demographic groups, and are least burdensome to administer owing to a shorter length and ease of understanding [27,54-66].

### App Use

As we were not able to collect in-app use data, participants were asked how often they used the app that they were assigned to over the past 4 weeks on a scale of 1 (*never downloaded the app*) to 8 (*multiple times per day*). To ease interpretation, results presented here are for response options collapsed into 4 categories, with findings highly similar in both categorization schemes. Categories included 1 (*never downloaded the app and downloaded but did not use the app*), 2 (*rarely [1-3 times in the past month] and infrequently [less than weekly]*), 3 (*weekly and more than weekly but less than daily*), and 4 (*daily and multiple times per day*).

### Fidelity

A dichotomized fidelity measure was created in accordance with each app's recommended use found on its website. Daily use was recommended for the apps Beautiful Mood [67], COVID Coach [68], and Calm [69], while weekly use was recommended for 7 Cups of Tea [70].

### Usability

App usability was assessed with the Intervention Usability Scale (IUS) [71], a 10-item measure that assesses psychosocial intervention usability through its likeability, learnability, difficulty, need for support, system integration, and efficiency. This measure is based on the System Usability Scale [72], a standardized, normed measure in industry for digital tools, and has been validated for online research [73,74]. The IUS is scored from 0 to 100, and a score of 85 or more is considered to be excellent usability [75].

### Acceptability and Appropriateness

The degree to which the app was satisfactory and appropriate (ie, the fit and relevance of the intervention) was measured with the Acceptability Intervention Measure (AIM) and the Intervention Appropriateness Measure (IAM) [76]. These scales' scores range from 0 to 20, with higher scores indicating great acceptability (AIM) and appropriateness (IAM). Both measures contain 4 items that exhibit good psychometric properties, and the items have been validated by implementation scientists and mental health professionals. Although these measures have not been validated for online use, they are the only validated instruments for intervention acceptability and appropriateness, are brief, and face-valid [76].

### Survey Administration

The survey was completed by volunteers identified through our initial sample [4]. Participation in the study, including survey completion, was voluntary. After volunteering on Prolific and providing consent, participants completed all measures in a web-based REDCap interface. The survey included 14 measures, each on a separate page, with 4-18 questions per measure. Participants were able to review and change their responses on each measure before proceeding to the next measure. The survey items were not randomized, as each scale used must be delivered in the order it was validated. We did not use skip patterns or other survey logic; participants were asked to complete all survey questions and had the option to not answer certain questions.

### Statistical Methods

Prior to analyses, we examined the data and eliminated participants for not meeting inclusion or good-actor criteria. Good-actor criteria required participants to correctly answer an attention check item, answer at least 50% of the items on the survey, complete the survey faster than 33% of the median length of time, and not have any problems with the Prolific ID (not being an approved ID, being a duplicate ID, or having a missing ID). We performed *t* tests and cross-tabulations with chi-square tests to compare demographic and baseline clinical outcomes by missing data status at the follow-up time point. Chi-square tests examined the association between condition assignment and compliance (whether a participant used the app they were assigned or used an alternative app). All analyses were of the intention-to-treat (ITT) type. Analyses of variance were used to compare conditions on AIM, IAM, and IUS scores at follow-up, with Tukey honestly significant difference (HSD) tests making pairwise post hoc comparisons among all apps. Mixed effects models using restricted maximum likelihood estimation were built to test the linear time change on the PHQ-9, GAD-7, SBQ-R, and DERS-SF and to test for condition differences at follow-up and change slope. We applied mixed effects models with 2 time points with within-person nesting. We used this rather than alternatives, such as a simple regression, because mixed effects models efficiently (1) simplified simultaneous testing of within-person changes via testing slope coefficients and between-condition differences via testing condition coefficients; (2) facilitated testing as we built models progressively adding time trends, condition effects, and dosage effects; and (3) permitted the inclusion of random

intercepts to account for variance at baseline, which was particularly important for testing models of dosage. Models were built and tested in an outwardly nested fashion, such that an initial null model was computed, followed by models that added a random intercept, random time component, condition assignment with Beautiful Mood as the reference variable (the attention control condition), and condition × time interaction effects. To test whether there was a dose-response relationship such that the app use frequency was associated with the rate of change on the PHQ-9, GAD-7, SBQ-R, and DERS-SF scores, we computed another series of mixed effects models for each outcome. An initial model included variables for time and frequency of use, a second model added condition terms for each app using Beautiful Mood as the reference, a third model added condition × frequency interaction terms for each app, and a fourth model included time, frequency, and a time × frequency interaction term. Model comparisons applied  $-2 \log$  likelihood ( $-2LL$ ), the Akaike information criterion (AIC), and Bayesian information criterion (BIC) deviance statistics. To test the impact of app use frequency on change over time, similar nested model testing was applied using an initial null model, followed by models that added dosage, dosage × time interaction, and condition assignment.

**Sample Size and Power**

A priori power analysis for an ANOVA *F* test indicated that a sample size of 800 (n=200 participants in each of the active and control app conditions) would be sufficient with power=0.80

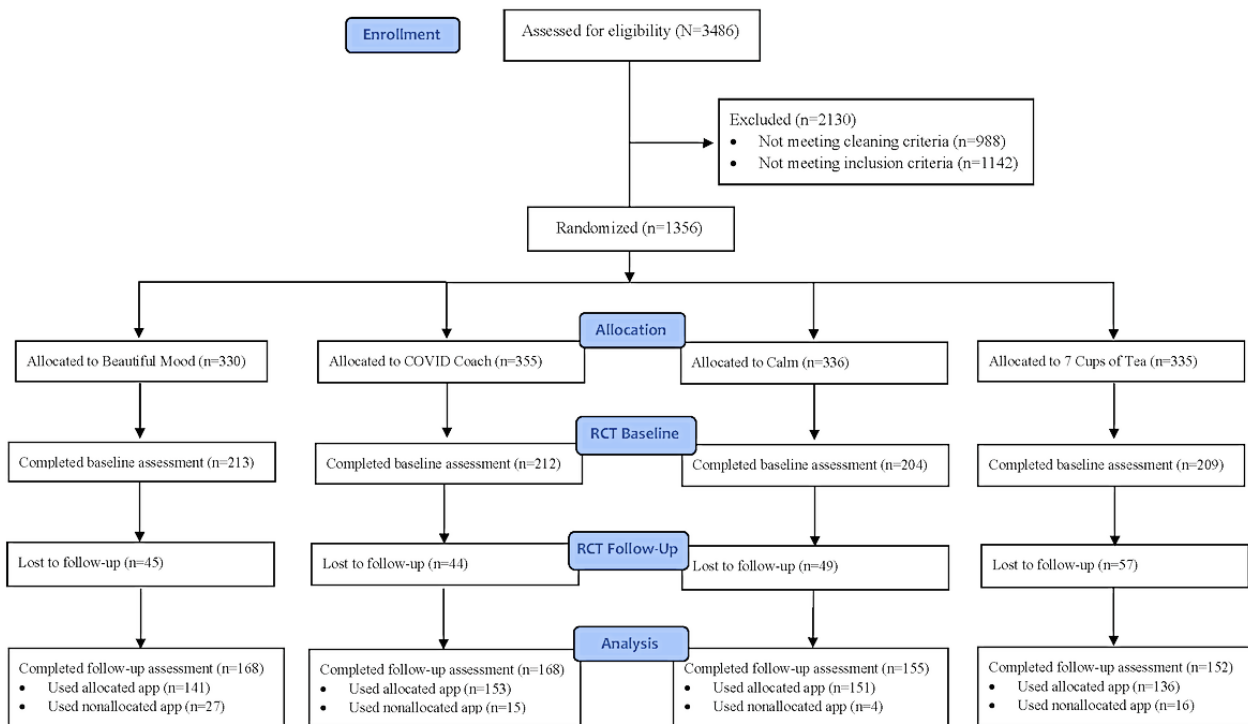
and  $\alpha=.05$  for a minimum detectable effect size (MDES) of Cohen  $d=0.24$  for main effect comparisons between any 2 conditions. The post hoc power analysis for 643 participants with complete data found an MDES of Cohen  $d=0.29$  for main effect comparisons between any 2 conditions. Previous research has found an average Hedges *g* effect size (a comparable effect size to Cohen *d* but corrected for small samples) on self-guided mental health apps to be 0.50 and for self-guided tools to be 0.24 [77].

**Results**

**Recruitment**

Figure 1 presents the Consolidated Standards of Reporting Trails (CONSORT) diagram. There were 3486 individuals assessed for eligibility, 2130 (61.1%) were excluded for not meeting good-actor criteria (n=988, 46.4%) or not meeting randomized clinical trial (RCT) inclusion criteria (n=1142, 53.6%). A total of 1356 (38.9%) individuals were randomized to Beautiful Mood (n=330, 24.3%), COVID Coach (n=355, 26.2%), Calm (n=336, 24.8%), or 7 Cups of Tea (n=335, 24.7%). Among those allocated to a condition, 838 (61.8%) participants completed the RCT baseline assessment, while 643 (47.4%) participants completed the follow-up assessment. In addition, 581 (90.4%) participants reported using the assigned app and 62 (9.6%) reported using a nonassigned app. For this ITT trial, all randomized participants were included in the primary analysis.

Figure 1. Consolidated Standards of Reporting Trails (CONSORT) table.



**Sample Description**

Table 1 presents individual-level demographic data. The analytic sample consisted of 838 adults, with a mean age of 31.1 (SD 9.5) years. Most patients identified as women (467/833, 56.1%)

and White (616/838, 73.5%). Participants self-identified as unemployed due to COVID-19 (428/838, 51.1%) or COVID-19–designated essential workers (410/838, 48.9%). There were no significant differences in demographics by condition.

**Table 1.** Sample descriptive statistics for 4 conditions using different self-guided mobile apps.

Demographics	Beautiful Mood (n=213)	COVID Coach (n=212)	Calm (n=204)	7 Cups of Tea (n=209)	Total (N=838)
<b>Race, n (%), P=.28<sup>a</sup></b>					
African American/Black	15 (7.0)	20 (9.4)	11 (5.4)	12 (5.7)	58 (6.9)
American Indian/Alaska Native	3 (1.4)	2 (0.9)	3 (1.5)	0	8 (1.0)
Asian	22 (10.3)	23 (10.8)	26 (12.7)	13 (6.2)	84 (10.0)
Multiracial	13 (6.1)	18 (8.5)	10 (4.9)	12 (5.7)	53 (6.3)
Other race <sup>b</sup>	4 (1.9)	7 (3.3)	5 (2.5)	3 (1.4)	19 (2.3)
White	156 (73.2)	142 (67.0)	149 (73.0)	169 (80.9)	616 (73.5)
<b>Ethnicity, n (%), P=.79<sup>a</sup></b>					
Hispanic/Latinx	23 (10.9)	17 (8.2)	21 (10.3)	19 (9.1)	80 (9.6)
Not Hispanic/Latinx	188 (89.1)	190 (91.8)	183 (89.7)	189 (90.9)	750 (90.4)
Missing	2 (0.9)	5 (2.4)	0	1 (0.5)	8 (1.0)
<b>Age (years), P=.16<sup>c</sup></b>					
Mean (SD)	31.8 (10.1)	29.9 (8.2)	31.4 (10.0)	31.4 (9.4) <sup>d</sup>	31.1 (9.5) <sup>e</sup>
<b>Gender, n (%), P=.86<sup>a</sup></b>					
Women	116 (54.7)	118 (56.5)	118 (57.8)	115 (55.3)	467 (56.1)
Gender diverse	0	1 (0.5)	1 (0.5)	2 (1.0)	4 (0.5)
Men	85 (40.1)	80 (38.3)	78 (38.2)	83 (39.9)	326 (39.1)
Nonbinary	10 (4.7)	7 (3.3)	7 (3.4)	7 (3.4)	31 (3.7)
Transgender	1 (0.5)	3 (1.4)	0	1 (0.5)	5 (0.6)
Missing	1 (0.5)	3 (1.4)	0	1 (0.5)	5 (0.6)
<b>Marital status, n (%), P=.41<sup>a</sup></b>					
Divorced	14 (6.7)	11 (5.3)	10 (4.9)	17 (8.2)	52 (6.3)
Married (including same-sex partnership)	61 (29.0)	56 (26.8)	56 (27.6)	61 (29.5)	234 (28.2)
Never married	134 (63.8)	136 (65.1)	131 (64.5)	123 (59.4)	524 (63.2)
Separated	1 (0.5)	6 (2.9)	4 (2.0)	3 (1.4)	14 (1.7)
Widowed	0	0	2 (1.0)	3 (1.4)	5 (0.6)
Missing	3 (1.4)	3 (1.4)	1 (0.5)	2 (1.0)	9 (1.1)
<b>Education, n (%), P=.47<sup>a</sup></b>					
High school, General Educational Development (GED), or less	22 (10.3)	28 (13.3)	27 (13.2)	23 (11.0)	100 (11.9)
Some college	74 (34.7)	50 (23.7)	54 (26.5)	51 (24.4)	229 (27.4)
Trade/technical/vocational	19 (8.9)	27 (12.8)	25 (12.3)	32 (15.3)	103 (12.3)
Bachelor's degree	64 (30.0)	73 (34.6)	65 (31.9)	67 (32.1)	269 (32.1)
Higher education	34 (16.0)	33 (15.6)	33 (16.2)	36 (17.2)	136 (16.2)
Missing	0	1 (0.5)	0	0	1 (0.1)
<b>Income (US \$), n (%), P=.28<sup>a</sup></b>					
<10,000	30 (14.2)	37 (18.0)	31 (15.3)	30 (14.5)	128 (15.5)
10,000-31,199	60 (28.3)	56 (27.2)	58 (28.6)	53 (25.6)	227 (27.4)
31,200-33,280	17 (8.0)	17 (8.3)	5 (2.5)	8 (3.9)	47 (5.7)
33,281-49,999	29 (13.7)	31 (15.0)	31 (15.3)	31 (15.0)	122 (14.7)

Demographics	Beautiful Mood (n=213)	COVID Coach (n=212)	Calm (n=204)	7 Cups of Tea (n=209)	Total (N=838)
50,000-59,999	11 (5.2)	16 (7.8)	23 (11.3)	17 (8.2)	67 (8.1)
60,000-69,999	10 (4.7)	10 (4.9)	14 (6.9)	14 (6.8)	48 (5.8)
70,000-99,999	34 (16.0)	19 (9.2)	20 (9.9)	27 (13.0)	100 (12.1)
100,000-149,999	14 (6.6)	15 (7.3)	11 (5.4)	14 (6.8)	54 (6.5)
≥150,000	7 (3.3)	5 (2.4)	10 (4.9)	13 (6.3)	35 (4.2)
Missing	1 (0.5)	6 (2.8)	1 (0.5)	2 (1.0)	10 (1.2)
<b>Employment, n (%), <math>P=.47^a</math></b>					
Essential worker	112 (52.6)	107 (50.5)	95 (46.6)	96 (45.9)	410 (48.9)
Unemployed	101 (47.4)	105 (49.5)	109 (53.4)	113 (54.1)	428 (51.1)

<sup>a</sup>Chi-square  $P$  value.

<sup>b</sup>Other race: most common responses for race were Hispanic, Mexican, and mixed.

<sup>c</sup>ANOVA  $F$  test  $P$  value.

<sup>d</sup> $N=208$ .

<sup>e</sup> $N=837$ .

### Missing Data

A total of 643/838 (76.7%) participants completed the follow-up assessment. There were no significant differences between those missing or not missing the follow-up assessment in the demographic data in Table 1 or clinical measures at baseline.

### Randomization Adherence and Compliance

At follow-up, 62/643 (9.6%) participants reported being nonadherent to condition assignment and reported that they use a different app than the one they were randomly assigned to

use. Participants who were randomized to Beautiful Mood were less likely to use their assigned app, while individuals randomized to Calm were more likely to use their assigned app ( $P<.001$ ).

### App Use

A cross-tabulation with the chi-square test found significant differences between the apps in the amount of use the participants reported; participants used Beautiful Mood more frequently and COVID Coach and 7 Cups of Tea less frequently (see Table 2).

**Table 2.** App compliance and use frequency<sup>a</sup> by condition ( $P<.001^b$ ).

App use and compliance	Beautiful Mood (n=168), n (%)	COVID Coach (n=168), n (%)	Calm (n=155), n (%)	7 Cups of Tea (n=152), n (%)	Total (N=643), n (%)
Adherent to app assignment	141(83.9)	153 (91.1)	151 (97.4)	136 (89.5)	581 (90.4)
<b>App use</b>					
Never downloaded/no use	18 (10.7)	14 (8.3)	15 (9.7)	21 (13.8)	68 (10.5)
Rarely/infrequently	<i>45 (26.8) <sup>c</sup></i>	<i>80 (47.6) <sup>c</sup></i>	62 (40.0)	<i>72 (47.4) <sup>c</sup></i>	259 (40.3)
Weekly or more	56 (33.3)	55 (32.7)	55 (35.5)	52 (34.2)	218 (33.9)
Daily/multiple times per day	<i>49 (29.2) <sup>c</sup></i>	19 (11.3)	23 (14.8)	<i>7 (4.6) <sup>c</sup></i>	98 (15.2)

<sup>a</sup>According to their websites, Beautiful Mood, COVID Coach, and Calm apps recommend daily use, while 7 Cups of Tea recommends weekly use.

<sup>b</sup>Chi-square  $P$  value.

<sup>c</sup>Italicized values indicate a significant difference indicated by standardized residuals.

### Usability, Acceptability, and Appropriateness

ANOVA found a significant difference on the IUS between the conditions (mean<sub>Beautiful Mood</sub> 72.9, SD 16.7; mean<sub>COVID Coach</sub> 71.2, SD 15.4; mean<sub>Calm</sub> 66.8, SD 17.3; mean<sub>7 Cups</sub> 65.2, SD 17.7). Tukey HSD post hoc tests indicated Beautiful Mood is significantly more usable than Calm (mean difference 6.0, 95% CI 1.2-10.8,  $P=.01$ ) and 7 Cups of Tea (mean difference 7.7, 95% CI 2.9-2.5,  $P<.001$ ). COVID Coach was significantly more

usable than 7 Cups of Tea (mean difference 6.1, 95% CI 1.2-10.9,  $P=.01$ ). We found no significant differences in app acceptability (overall AIM mean 3.5, SD 1.0, 95% CI 3.4-3.6,  $P=.22$ ) or appropriateness (overall IAM mean 3.6, SD 0.9, 95% CI 3.6-3.7,  $P=.48$ ).

### Clinical Outcomes

Table 3 displays the reporting sample size, mean scores, and SDs at each time point for the PHQ-9, GAD-7, SBQ-R, and DERS-SF for each app.

**Table 3.** Pretest and posttest scores on clinical outcomes by condition.

App and time	PHQ <sup>a</sup> -9		GAD <sup>b</sup> -7		SBQ-R <sup>c</sup>		DERS-SF <sup>d</sup>	
	n (%)	Mean (SD)	n (%)	Mean (SD)	n (%)	Mean (SD)	n (%)	Mean (SD)
<b>Beautiful Mood (n=168)</b>								
Pretest	165 (98.2)	10.6 (6.6)	167 (99.4)	8.8 (5.7)	154 (91.7)	7.1 (3.6)	165 (98.2)	44.5 (13.0)
Posttest	165 (98.2)	9.1 (6.5)	167 (99.4)	7.8 (5.6)	154 (91.7)	7.0 (3.8)	165 (98.2)	44.6 (14.2)
<b>COVID Coach (n=168)</b>								
Pretest	166 (98.8)	11.2 (6.3)	168 (100)	9.2 (5.6)	159 (94.6)	6.9 (3.8)	163 (97.0)	44.6 (12.5)
Posttest	166 (98.8)	9.8 (6.7)	168 (100)	7.8 (5.6)	159 (94.6)	6.7 (3.7)	163 (97.0)	44.6 (14.0)
<b>Calm (n=155)</b>								
Pretest	155 (100)	10.1 (5.7)	153 (98.7)	7.9 (4.8)	144 (92.9)	6.6 (3.3)	153 (98.7)	42.9 (11.7)
Posttest	155 (100)	8.5 (5.9)	153 (98.7)	6.7 (5.1)	144 (92.9)	6.4 (3.3)	153 (98.7)	42.2 (13.1)
<b>7 Cups of Tea (n=152)</b>								
Pretest	151 (93.3)	11.0 (6.5)	151 (93.3)	8.9 (5.6)	139 (91.4)	7.1 (3.9)	148 (97.4)	44.6 (12.8)
Posttest	151 (93.3)	9.7 (6.6)	151 (93.3)	7.3 (5.8)	139 (91.4)	7.1 (3.9)	148 (97.4)	44.9 (13.0)

<sup>a</sup>PHQ: Patient Health Questionnaire.

<sup>b</sup>GAD: Generalized Anxiety Disorder Scale.

<sup>c</sup>SBQ-R: Suicide Behaviors Questionnaire-Revised.

<sup>d</sup>DERS-SF: Difficulties in Emotion Regulation Scale – Short Form.

Examining the  $-2LL$ , AIC, and BIC deviance statistics for each of the 4 analyses revealed that the best-fitting model was a random intercept model with a linear time slope in 2 cases: the PHQ-9 ( $-2LL=9118$ ,  $dev_1=63.2$ ,  $P<.001$ ; AIC=9128,  $dev_1=61.2$ ,  $P<.001$ ; BIC=9137,  $dev_1=71$ ,  $P<.001$ ; parameters=5) and GAD-7 ( $-2LL=8686$ ,  $dev_1=71.7$ ,  $P<.001$ ; AIC=8696,  $dev_1=69.7$ ,  $P<.001$ ; BIC=8720,  $dev_1=64.9$ ,  $P<.001$ ; parameters=5). From baseline to follow-up, participants improved by an estimated  $-1.5$  points on the PHQ-9 (SE 0.2, 95% CI  $-1.1$  to  $-1.8$ ,  $P<.001$ ) and  $-1.3$  points on the GAD-7 (SE 0.2, 95% CI  $-1.0$  to  $-1.6$ ,  $P<.001$ ). Models that included condition main effects centered at the follow-up time point and condition  $\times$  time were not significantly better fitting than the random intercept and time model. For the other 2 analyses, the SBQ-R and DERS-SF, the best-fitting models were the null models, with no random terms or interaction variables (SBQ-R:  $-2LL=6695$ , AIC=6703, BIC=6722, parameters=4; DERS-SF:  $-2LL=11,078$ , AIC=11,086, BIC=11,105, parameters=4). Thus, there were significant mean improvements in the PHQ-9 and GAD-7 scores but not the SBQ-R and DERS-SF scores, and the app condition was not associated with differences in any of the 4 analyses.

### Dosage

Mixed effects models were computed to examine the relation of frequency of app use with change over time on the PHQ-9, GAD-7, SBQ-R, and DERS-SF, over all conditions, controlling for the condition, for a condition  $\times$  app use interaction, and for condition  $\times$  time. The best-fitting model for the PHQ-9 included time, frequency of app use, and frequency  $\times$  time interaction, as indicated by 2 of the 3 fit indices ( $-2LL=7868$ ,  $dev_1=8.9$ ,  $P=.002$ ; AIC=7882,  $dev_1=6.9$ ,  $P=.01$ ; BIC=7914,  $dev_1=2.4$ ,  $P=.12$ ; parameters=7,  $P<.001$ ). The BIC statistic, which penalizes for model complexity, was not statistically significant;

therefore, these results should be viewed with some caution. All 3 fit statistics indicated that this same model structure was the best fit for the GAD-7 ( $-2LL=7481$ ,  $dev_1=14.2$ ,  $P<.001$ ; AIC=7495,  $dev_1=12.2$ ,  $P<.001$ ; BIC=7527,  $dev_1=7.8$ ,  $P=.01$ ; parameters=7). None of the models that included condition or condition  $\times$  frequency of app use was a significantly better fit, meaning we found no differences between the treatment groups on the impact that app use frequency had on change over time. For the SBQ-R and DERS-SF, none of the more complex models improved on the fit of the initial model that included the time and frequency of app use.

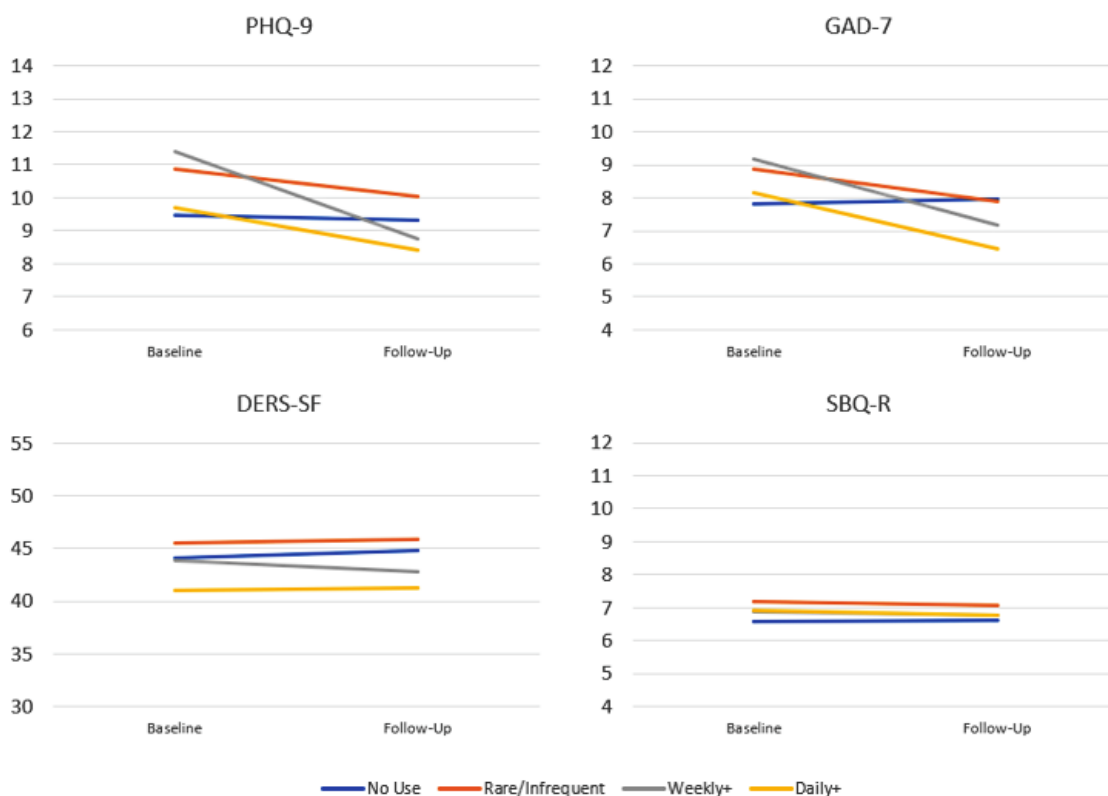
Salient model parameters for the best-fitting models were as follows. For the PHQ-9, when time, frequency of app use, and time  $\times$  frequency interaction were included in the best-fitting model, the frequency of app use was not significant (estimate=0.05, SE 0.3, 95% CI 0.5-0.6,  $P=.86$ ) and time was not significant (estimate=0.1, SE 0.6, 95% CI 1.0-1.3,  $P=.80$ ), but the time  $\times$  frequency interaction was significant (estimate= $-0.6$ , SE 0.2, 95% CI 1.0 to  $-0.2$ ,  $P=.003$ ). There were similar findings for the GAD-7 such that when time, frequency of app use, and time  $\times$  frequency interaction were included in the best-fitting model, the frequency of app use was not significant (estimate=0.1, SE 0.3, 95% CI 0.4-0.6,  $P=.78$ ) and time was not significant (estimate=0.4, SE 0.5, 95% CI 0.6-1.3,  $P=.43$ ), but the time  $\times$  frequency interaction was significant (estimate= $-0.7$ , SE  $-0.7$ , 95% CI 1.0 to  $-0.3$ ,  $P<.001$ ). Figure 2 depicts the actual mean score for each condition by frequency. For the PHQ-9 and GAD-7, those who did not use the app had no significant change on that measure over time; those who used the app more frequently improved more quickly than those who used the app less frequently. By the 4-week follow-up, however, there were no significant differences in outcome by frequency of app use (dose).



For the SBQ-R, the best-fitting model indicated that the frequency of app use was not associated with lower scores overall (estimate=-0.03, SE 0.2, 95% CI 0.4-0.3,  $P=.16$ ), although time was significant (estimate=-0.1, SE 0.1, 95% CI 0.3 to -0.003,  $P=.05$ ); interaction terms were not included. Therefore, when statistically controlling for the frequency of app use, SBQ-R scores decreased over time but there was no association between app frequency and change on the SBQ-R.

For the DERS-SF, the best-fitting model indicated that the frequency of app use was associated with lower scores overall (estimate=-1.5, SE 0.6, 95% CI 2.6 to -0.4,  $P=.01$ ), but time was not significant (estimate=-0.1, SE 0.3, 95% CI 0.8 to -0.6,  $P=.84$ ), and interaction terms were not included. Those who used their app frequently had lower scores on the DERS-SF at baseline and follow-up, with no change on the DERS-SF over time and no association between frequency of app use and change on the DERS-SF.

**Figure 2.** Clinical outcome means. DERS-SF: Difficulties in Emotion Regulation Scale – Short Form; GAD: Generalized Anxiety Disorder Scale; PHQ: Patient Health Questionnaire; SBQ-R: Suicide Behaviors Questionnaire-Revised.



## Discussion

### Principal Findings

To the best of our knowledge, this is 1 of the first pragmatic trials of free commercial apps among essential workers or those unemployed due to COVID-19 experiencing emotional distress and suicide risk. Our primary findings were that commercial mobile mental health apps are found to be usable and acceptable and have a positive impact on depression and anxiety but not emotional regulation or suicide risk. Although we did not find any significant difference between the 3 active apps on outcomes, nor between the active apps compared to the control app, we did find that the frequency of app use during the 4 weeks had a significant and positive impact on depression and anxiety outcomes. However, we offer here that 4 weeks of engagement may not be sufficient to show changes in emotional regulation or suicidal behavior or that online interventions may not be potent enough to manage these mental health challenges, given that those with greater emotion dysregulation used the apps less throughout the period. Indeed, a recent study offering online interventions aimed at suicide prevention not only found

no effects but also demonstrated more adverse events in those offered online care versus those offered care as usual [78]. We also cannot rule out regression to the mean, as people who used the attention control app had outcomes similar to participants who used the active apps.

### Comparison With Previous Work

Our previous work suggests that apps that focus on mindfulness, pandemic information, mood tracking, and connection with others are an acceptable means of managing stress during COVID-19 [4]. Our findings on the lack of differential clinical impact between apps is not surprising, given the data from other pragmatic trials of research-grade mobile apps. In large-scale remote, pragmatic clinical trials of mobile apps for depression, all apps found significant improvements in mood over an 8-week period but no differences between groups [17]. Our findings regarding the importance of app use on clinical outcomes have also been found in previous studies on research-grade mobile apps, where frequent use of a mobile mood app early in care resulted in better depression outcomes for those who were more severely distressed [17,79]. Although smaller, controlled trials of self-guided apps in a research context do find small but

statistically significant differences in outcomes compared to waitlist controls or no treatment [76], this study, and other pragmatic trials to date, have not demonstrated that active apps are more effective than attention control apps [7,78].

### Study Strengths and Limitations

The strength of this study is in its design: It is 1 of the first studies to evaluate free commercial mental health apps prospectively and independently in a large-scale, pragmatic RCT and to assess their impact on emotional distress, emotional regulation, and suicide risk in 2 suicide-vulnerable populations at the peak of the COVID-19 pandemic. The lessons learned from this study can be useful to people seeking free and readily accessible help for emotional distress. For the field, more work is needed to understand what role commercial apps can play for emotional distress, and the data from this study serve as a good starting point for understanding what is acceptable and effective and what optimal engagement should be.

Study limitations include the following:

- Our sample consisted of participants from Prolific and thus may be most representative of essential workers or those unemployed due to COVID-19 who are proactively seeking other sources of income to offset financial stress. Although this sample may be more comfortable with technology, we believe that people seeking mobile mental health apps are also comfortable with technology, and thus the results from this study are representative of this population.
- Because we did not partner with the technology companies who created the study apps, we relied on self-reported app

use, which may be subject to self-report bias. However, incentives for participating in this study were not tied to app use, and data from numerous intervention studies find that people are highly accurate in their reports of intervention adherence [80-82].

- Although we justify our timeline for measuring outcomes after 4 weeks of intervention use based on what is typical for most mental health app users, we do not have information on the lasting effects of treatment outcomes or on continued app use. Thus, although we can report on the immediate effects of the intervention, future research is needed to determine the permanence of treatment effects.
- We did not ask about potential adverse events related to app use. This is an interesting area of research that to date has not been explored. Understanding the risks of using commercial apps is as important as determining their impacts and should be explored in future studies.

### Conclusion

There are several papers calling for more research to study the effectiveness of commercial mental health app interventions [7,10,83,84], specifically for COVID-19 [85-88], but published studies to date report only app downloads, aesthetics, and app use [6,83,89]. Our data suggest that essential workers and those unemployed who want self-guided mental health care found 4 commercially available apps both acceptable and usable and might receive emotional benefit from a variety of self-guided mental health apps, particularly if they use the apps frequently, but that regression to the mean cannot be ruled out, so improvement in symptoms may not be attributable to app use.

### Acknowledgments

This study was supported by the National Institute of Mental Health (P50MH115837). The funding body has no role in study design, writing of the manuscript, or decision to submit the article for publication.

### Data Availability

The raw data supporting the conclusions of this paper will be made available by the authors, without undue reservation.

### Authors' Contributions

PA, KAC, and MDP contributed to the conception and design of the study. PA, KAC, BM, MDP, and FM-G contributed to the acquisition, analysis, and interpretation of the data. MJ and MDP performed the statistical analysis. All authors contributed to writing, reviewing, and reading the manuscript, and all approved the submitted version.

### Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

### Multimedia Appendix 1

Demographics survey.

[\[DOCX File, 24 KB-Multimedia Appendix 1\]](#)

### Multimedia Appendix 2

CONSORT-eHEALTH checklist (V 1.6.1).

[\[PDF File \(Adobe PDF File\), 1195 KB-Multimedia Appendix 2\]](#)

### References

1. Chaiton M, Musani I, Pullman M, Logie CH, Abramovich A, Grace D, et al. Access to mental health and substance use resources for 2SLGBTQ+ youth during the covid-19 pandemic. *Int J Environ Res Public Health* 2021 Oct 28;18(21):11315 [FREE Full text] [doi: [10.3390/ijerph182111315](https://doi.org/10.3390/ijerph182111315)] [Medline: [34769836](https://pubmed.ncbi.nlm.nih.gov/34769836/)]
2. Lucas D, Bamber J. Pandemics and maternal health: the indirect effects of COVID-19. *Anaesthesia* 2021 Apr;76(Suppl 4):69-75 [FREE Full text] [doi: [10.1111/anae.15408](https://doi.org/10.1111/anae.15408)] [Medline: [33682091](https://pubmed.ncbi.nlm.nih.gov/33682091/)]
3. Saltzman LY, Lesen AE, Henry V, Hansel TC, Bordnick PS. COVID-19 mental health disparities. *Health Secur* 2021 Jun 01;19(S1):S5-S13. [doi: [10.1089/hs.2021.0017](https://doi.org/10.1089/hs.2021.0017)] [Medline: [34014118](https://pubmed.ncbi.nlm.nih.gov/34014118/)]
4. Mata-Greve F, Johnson M, Pullmann MD, Friedman EC, Griffith Fillipo I, Comtois KA, et al. Mental health and the perceived usability of digital mental health tools among essential workers and people unemployed due to covid-19: cross-sectional survey study. *JMIR Ment Health* 2021 Aug 05;8(8):e28360 [FREE Full text] [doi: [10.2196/28360](https://doi.org/10.2196/28360)] [Medline: [34081592](https://pubmed.ncbi.nlm.nih.gov/34081592/)]
5. Sams N, Fisher DM, Mata-Greve F, Johnson M, Pullmann MD, Raue PJ, et al. Understanding psychological distress and protective factors amongst older adults during the covid-19 pandemic. *Am J Geriatr Psychiatry* 2021 Sep;29(9):881-894 [FREE Full text] [doi: [10.1016/j.jagp.2021.03.005](https://doi.org/10.1016/j.jagp.2021.03.005)] [Medline: [33867224](https://pubmed.ncbi.nlm.nih.gov/33867224/)]
6. Martinengo L, Stona A, Griva K, Dazzan P, Pariante CM, von Wangenheim F, et al. Self-guided cognitive behavioral therapy apps for depression: systematic assessment of features, functionality, and congruence with evidence. *J Med Internet Res* 2021 Jul 30;23(7):e27619 [FREE Full text] [doi: [10.2196/27619](https://doi.org/10.2196/27619)] [Medline: [34328431](https://pubmed.ncbi.nlm.nih.gov/34328431/)]
7. Saad A, Bruno D, Camara B, D'Agostino J, Bolea-Alamanac B. Self-directed technology-based therapeutic methods for adult patients receiving mental health services: systematic review. *JMIR Ment Health* 2021 Nov 26;8(11):e27404 [FREE Full text] [doi: [10.2196/27404](https://doi.org/10.2196/27404)] [Medline: [34842556](https://pubmed.ncbi.nlm.nih.gov/34842556/)]
8. Malakouti S, Rasouli N, Rezaean M, Nojomi M, Ghanbari B, Shahraki Mohammadi A. Effectiveness of self-help mobile telephone applications (apps) for suicide prevention: a systematic review. *Med J Islam Repub Iran* 2020 Oct 30;34:85. [doi: [10.47176/mjiri.34.85](https://doi.org/10.47176/mjiri.34.85)]
9. Moshe I, Terhorst Y, Philippi P, Domhardt M, Cuijpers P, Cristea I, et al. Digital interventions for the treatment of depression: a meta-analytic review. *Psychol Bull* 2021 Aug;147(8):749-786. [doi: [10.1037/bul0000334](https://doi.org/10.1037/bul0000334)] [Medline: [34898233](https://pubmed.ncbi.nlm.nih.gov/34898233/)]
10. Rathbone AL, Prescott J. The use of mobile apps and SMS messaging as physical and mental health interventions: systematic review. *J Med Internet Res* 2017 Aug 24;19(8):e295 [FREE Full text] [doi: [10.2196/jmir.7740](https://doi.org/10.2196/jmir.7740)] [Medline: [28838887](https://pubmed.ncbi.nlm.nih.gov/28838887/)]
11. 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](https://doi.org/10.2196/14567)] [Medline: [31573916](https://pubmed.ncbi.nlm.nih.gov/31573916/)]
12. One Mind PsyberGuide. URL: <https://onemindpsyberguide.org/> [accessed 2022-08-03]
13. American Psychiatric Association. App Advisor: An American Psychiatric Association Initiative. URL: <https://www.psychiatry.org/psychiatrists/practice/mental-health-apps> [accessed 2022-08-03]
14. Mohr DC, Azocar F, Bertagnolli A, Choudhury T, Chrisp P, Frank R, Banbury Forum on Digital Mental Health. Banbury forum consensus statement on the path forward for digital mental health treatment. *Psychiatr Serv* 2021 Jun 20;72(6):677-683 [FREE Full text] [doi: [10.1176/appi.ps.202000561](https://doi.org/10.1176/appi.ps.202000561)] [Medline: [33467872](https://pubmed.ncbi.nlm.nih.gov/33467872/)]
15. 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](https://doi.org/10.1002/wps.20592)] [Medline: [30600619](https://pubmed.ncbi.nlm.nih.gov/30600619/)]
16. Sharma M, Creutzfeldt C, Lewis A, Patel PV, Hartog C, Jannotta GE, et al. Health-care professionals' perceptions of critical care resource availability and factors associated with mental well-being during coronavirus disease 2019 (covid-19): results from a US survey. *Clin Infect Dis* 2021 May 18;72(10):e566-e576 [FREE Full text] [doi: [10.1093/cid/ciaa1311](https://doi.org/10.1093/cid/ciaa1311)] [Medline: [32877508](https://pubmed.ncbi.nlm.nih.gov/32877508/)]
17. Arean PA, Hallgren KA, Jordan JT, Gazzaley A, Atkins DC, Heagerty PJ, et al. The use and effectiveness of mobile apps for depression: results from a fully remote clinical trial. *J Med Internet Res* 2016 Dec 20;18(12):e330 [FREE Full text] [doi: [10.2196/jmir.6482](https://doi.org/10.2196/jmir.6482)] [Medline: [27998876](https://pubmed.ncbi.nlm.nih.gov/27998876/)]
18. Lin Z, Althoff T, Leskovec J. I'll be back: on the multiple lives of users of a mobile activity tracking application. 2018 Apr Presented at: WWW '18: The Web Conference 2018; April 23-27, 2018; Lyon, France p. 1501-1511 URL: <https://europepmc.org/abstract/MED/29780978> [doi: [10.1145/3178876.3186062](https://doi.org/10.1145/3178876.3186062)]
19. Epstein D, Ping A, Fogarty J, Munson S. A lived informatics model of personal informatics. 2015 Presented at: UbiComp '15: The 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 7-11, 2015; Osaka, Japan p. 731-742. [doi: [10.1145/2750858.2804250](https://doi.org/10.1145/2750858.2804250)]
20. 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](https://doi.org/10.1016/j.jbef.2017.12.004)]
21. Peer E, Brandimarte L, Samat S, Acquisti A. Beyond the Turk: alternative platforms for crowdsourcing behavioral research. *J Exp Soc Psychol* 2017 May;70:153-163. [doi: [10.1016/j.jesp.2017.01.006](https://doi.org/10.1016/j.jesp.2017.01.006)]
22. Prolific. A Higher Standard of Online Research. URL: <https://www.prolific.co/> [accessed 2022-10-05]
23. Lettmann H, Lumsden J. Prolific's Participant Pool - the Present and the Future. 2018. URL: <https://blog.prolific.co/prolifics-participant-pool-its-present-and-its-future/> [accessed 2022-10-05]

24. Oppenheimer DM, Meyvis T, Davidenko N. Instructional manipulation checks: detecting satisficing to increase statistical power. *J Exp Soc Psychol* 2009 Jul;45(4):867-872. [doi: [10.1016/j.jesp.2009.03.009](https://doi.org/10.1016/j.jesp.2009.03.009)]
25. Kroenke K, Spitzer R, Williams J. The Patient Health Questionnaire-2: validity of a two-item depression screener. *Med Care* 2003;41(11):1284-1292. [doi: [10.1097/01.mlr.0000093487.78664.3c](https://doi.org/10.1097/01.mlr.0000093487.78664.3c)]
26. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Ann Intern Med* 2007 Mar 06;146(5):317. [doi: [10.7326/0003-4819-146-5-200703060-00004](https://doi.org/10.7326/0003-4819-146-5-200703060-00004)]
27. Osman A, Bagge CL, Gutierrez PM, Konick LC, Kopper BA, Barrios FX. The Suicidal Behaviors Questionnaire-Revised (SBQ-R): validation with clinical and nonclinical samples. *Assessment* 2001 Dec 26;8(4):443-454. [doi: [10.1177/107319110100800409](https://doi.org/10.1177/107319110100800409)] [Medline: [11785588](https://pubmed.ncbi.nlm.nih.gov/11785588/)]
28. Millner AJ, Lee MD, Nock MK. Single-item measurement of suicidal behaviors: validity and consequences of misclassification. *PLoS One* 2015;10(10):e0141606 [FREE Full text] [doi: [10.1371/journal.pone.0141606](https://doi.org/10.1371/journal.pone.0141606)] [Medline: [26496707](https://pubmed.ncbi.nlm.nih.gov/26496707/)]
29. Patsopoulos NA. A pragmatic view on pragmatic trials. *Dialogues Clin Neurosci* 2022 Apr 01;13(2):217-224. [doi: [10.31887/dcms.2011.13.2/npatsopoulos](https://doi.org/10.31887/dcms.2011.13.2/npatsopoulos)]
30. Dal-Ré R, Janiaud P, Ioannidis J. Real-world evidence: how pragmatic are randomized controlled trials labeled as pragmatic? *BMC Med* 2018 Apr 03;16(1):49 [FREE Full text] [doi: [10.1186/s12916-018-1038-2](https://doi.org/10.1186/s12916-018-1038-2)] [Medline: [29615035](https://pubmed.ncbi.nlm.nih.gov/29615035/)]
31. Patterson B, Boyle MH, Kivlenieks M, Van Ameringen M. The use of waitlists as control conditions in anxiety disorders research. *J Psychiatr Res* 2016 Dec;83:112-120. [doi: [10.1016/j.jpsychires.2016.08.015](https://doi.org/10.1016/j.jpsychires.2016.08.015)] [Medline: [27585425](https://pubmed.ncbi.nlm.nih.gov/27585425/)]
32. Bailey E, Mühlmann C, Rice S, Nedeljkovic M, Alvarez-Jimenez M, Sander L, et al. Ethical issues and practical barriers in internet-based suicide prevention research: a review and investigator survey. *BMC Med Ethics* 2020 May 13;21(1):37 [FREE Full text] [doi: [10.1186/s12910-020-00479-1](https://doi.org/10.1186/s12910-020-00479-1)] [Medline: [32404098](https://pubmed.ncbi.nlm.nih.gov/32404098/)]
33. Mohr DC, Spring B, Freedland KE, Beckner V, Arean P, Hollon SD, et al. The selection and design of control conditions for randomized controlled trials of psychological interventions. *Psychother Psychosom* 2009;78(5):275-284. [doi: [10.1159/000228248](https://doi.org/10.1159/000228248)] [Medline: [19602916](https://pubmed.ncbi.nlm.nih.gov/19602916/)]
34. Areán P, Alvidrez J. Ethical considerations in psychotherapy effectiveness research: choosing the comparison group. *Ethics Behav* 2002 Jan;12(1):63-73. [doi: [10.1207/s15327019eb1201\\_4](https://doi.org/10.1207/s15327019eb1201_4)]
35. Carlo AD, Hosseini Ghomi R, Renn BN, Areán PA. By the numbers: ratings and utilization of behavioral health mobile applications. *NPJ Digit Med* 2019 Jun 17;2(1):54 [FREE Full text] [doi: [10.1038/s41746-019-0129-6](https://doi.org/10.1038/s41746-019-0129-6)] [Medline: [31304400](https://pubmed.ncbi.nlm.nih.gov/31304400/)]
36. Pratap A, Renn BN, Volponi J, Mooney SD, Gazzaley A, Arean PA, et al. Using mobile apps to assess and treat depression in Hispanic and Latino populations: fully remote randomized clinical trial. *J Med Internet Res* 2018 Aug 09;20(8):e10130 [FREE Full text] [doi: [10.2196/10130](https://doi.org/10.2196/10130)] [Medline: [30093372](https://pubmed.ncbi.nlm.nih.gov/30093372/)]
37. Schlosser DA, Campellone TR, Truong B, Anguera JA, Vergani S, Vinogradov S, et al. The feasibility, acceptability, and outcomes of PRIME-D: a novel mobile intervention treatment for depression. *Depress Anxiety* 2017 Jun 18;34(6):546-554 [FREE Full text] [doi: [10.1002/da.22624](https://doi.org/10.1002/da.22624)] [Medline: [28419621](https://pubmed.ncbi.nlm.nih.gov/28419621/)]
38. Graham AK, Greene CJ, Kwasny MJ, Kaiser SM, Lieponis P, Powell T, et al. Coached mobile app platform for the treatment of depression and anxiety among primary care patients: a randomized clinical trial. *JAMA Psychiatry* 2020 Sep 01;77(9):906-914 [FREE Full text] [doi: [10.1001/jamapsychiatry.2020.1011](https://doi.org/10.1001/jamapsychiatry.2020.1011)] [Medline: [32432695](https://pubmed.ncbi.nlm.nih.gov/32432695/)]
39. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap): a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009 Apr;42(2):377-381 [FREE Full text] [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
40. U.S. Census Bureau. The 2017 National Population Projections. 2017. URL: <https://www.census.gov/programs-surveys/popproj.html> [accessed 2022-08-03]
41. Woerner M, Sams N, Rivera Nales C, Gorstein T, Johnson M, Mosser BA, et al. Generational perspectives on technology's role in mental health care: a survey of adults with lived mental health experience. *Front Digit Health* 2022;4:840169 [FREE Full text] [doi: [10.3389/fgdth.2022.840169](https://doi.org/10.3389/fgdth.2022.840169)] [Medline: [35224537](https://pubmed.ncbi.nlm.nih.gov/35224537/)]
42. National Research Council (US) Panel on DHHS Collection of Race and Ethnicity Data. Measuring race, ethnicity, socioeconomic position, and acculturation. In: Ver Ploeg M, Perrin W, editors. *Eliminating Health Disparities: Measurement and Data Needs*. Vol 3. Washington, DC: National Academies Press; 2004.
43. Rosenfield S, Mouzon D. Gender and mental health. In: *Handbook of the Sociology of Mental Health*. Berlin: Springer; 2013:277-296.
44. Kroenke K, Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. *Psychiatric Annals* 2002 Sep 01;32(9):509-515. [doi: [10.3928/0048-5713-20020901-06](https://doi.org/10.3928/0048-5713-20020901-06)]
45. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001 Sep;16(9):606-613 [FREE Full text] [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
46. Plummer F, Manea L, Trepel D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic metaanalysis. *Gen Hosp Psychiatry* 2016 Mar;39:24-31. [doi: [10.1016/j.genhosppsy.2015.11.005](https://doi.org/10.1016/j.genhosppsy.2015.11.005)] [Medline: [26719105](https://pubmed.ncbi.nlm.nih.gov/26719105/)]
47. Spitzer RL, Kroenke K, Williams JBW, 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](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]

48. Kaufman EA, Xia M, Fosco G, Yaptangco M, Skidmore CR, Crowell SE. The Difficulties in Emotion Regulation Scale Short Form (DERS-SF): validation and replication in adolescent and adult samples. *J Psychopathol Behav Assess* 2015 Nov 23;38(3):443-455. [doi: [10.1007/s10862-015-9529-3](https://doi.org/10.1007/s10862-015-9529-3)]
49. Hallion LS, Steinman SA, Tolin DF, Diefenbach GJ. Psychometric properties of the Difficulties in Emotion Regulation Scale (DERS) and its short forms in adults with emotional disorders. *Front Psychol* 2018 Apr 19;9:539 [FREE Full text] [doi: [10.3389/fpsyg.2018.00539](https://doi.org/10.3389/fpsyg.2018.00539)] [Medline: [29725312](https://pubmed.ncbi.nlm.nih.gov/29725312/)]
50. Erbe D, Eichert H, Rietz C, Ebert D. Interformat reliability of the Patient Health Questionnaire: validation of the computerized version of the PHQ-9. *Internet Interv* 2016 Sep;5:1-4 [FREE Full text] [doi: [10.1016/j.invent.2016.06.006](https://doi.org/10.1016/j.invent.2016.06.006)] [Medline: [30135800](https://pubmed.ncbi.nlm.nih.gov/30135800/)]
51. Seo J, Park S. Validation of the Generalized Anxiety Disorder-7 (GAD-7) and GAD-2 in patients with migraine. *J Headache Pain* 2015 Nov 23;16(1):97 [FREE Full text] [doi: [10.1186/s10194-015-0583-8](https://doi.org/10.1186/s10194-015-0583-8)] [Medline: [26596588](https://pubmed.ncbi.nlm.nih.gov/26596588/)]
52. Hesser H, Axelsson S, Bäcké V, Engstrand J, Gustafsson T, Holmgren E, et al. Preventing intimate partner violence via the internet: a randomized controlled trial of emotion-regulation and conflict-management training for individuals with aggression problems. *Clin Psychol Psychother* 2017 Sep;24(5):1163-1177. [doi: [10.1002/cpp.2082](https://doi.org/10.1002/cpp.2082)] [Medline: [28261923](https://pubmed.ncbi.nlm.nih.gov/28261923/)]
53. Tsaousides T, Spielman L, Kajankova M, Guetta G, Gordon W, Dams-O'Connor K. Improving emotion regulation following web-based group intervention for individuals with traumatic brain injury. *J Head Trauma Rehabil* 2017 Sep;32(5):354-365. [doi: [10.1097/htr.0000000000000345](https://doi.org/10.1097/htr.0000000000000345)]
54. Cassidy SA, Bradley L, Cogger-Ward H, Rodgers J. Development and validation of the suicidal behaviours questionnaire - autism spectrum conditions in a community sample of autistic, possibly autistic and non-autistic adults. *Mol Autism* 2021 Jun 21;12(1):46 [FREE Full text] [doi: [10.1186/s13229-021-00449-3](https://doi.org/10.1186/s13229-021-00449-3)] [Medline: [34154642](https://pubmed.ncbi.nlm.nih.gov/34154642/)]
55. Batterham PJ, Ftanou M, Pirkis J, Brewer JL, Mackinnon AJ, Beautrais A, et al. A systematic review and evaluation of measures for suicidal ideation and behaviors in population-based research. *Psychol Assess* 2015 Jun;27(2):501-512. [doi: [10.1037/pas0000053](https://doi.org/10.1037/pas0000053)] [Medline: [25496086](https://pubmed.ncbi.nlm.nih.gov/25496086/)]
56. Huen JMY, Yip PSF, Osman A, Leung ANM. The Suicidal Behaviors Questionnaire-Revised (SBQ-R) and its Chinese version (C-SBQ-R): further validity testing using the culture, comprehension, and translation bias procedure. *Psychol Assess* 2022 Jul;34(7):704-710. [doi: [10.1037/pas0001134](https://doi.org/10.1037/pas0001134)] [Medline: [35467908](https://pubmed.ncbi.nlm.nih.gov/35467908/)]
57. Amini-Tehrani M, Nasiri M, Jalali T, Sadeghi R, Ghotbi A, Zamanian H. Validation and psychometric properties of Suicide Behaviors Questionnaire-Revised (SBQ-R) in Iran. *Asian J Psychiatr* 2020 Jan;47:101856. [doi: [10.1016/j.ajp.2019.101856](https://doi.org/10.1016/j.ajp.2019.101856)] [Medline: [31704597](https://pubmed.ncbi.nlm.nih.gov/31704597/)]
58. Gómez-Romero MJ, Tomás-Sábado J, Montes-Hidalgo J, Brando-Garrido C, Limonero JT. The suicidal behaviors questionnaire-revised. Spanish form. *Death Stud* 2021 Oct 01;45(8):623-629. [doi: [10.1080/07481187.2019.1671544](https://doi.org/10.1080/07481187.2019.1671544)] [Medline: [31573435](https://pubmed.ncbi.nlm.nih.gov/31573435/)]
59. Bjureberg J, Ljótsson B, Tull MT, Hedman E, Sahlin H, Lundh L, et al. Development and validation of a brief version of the Difficulties in Emotion Regulation Scale: the DERS-16. *J Psychopathol Behav Assess* 2016 Jun 14;38(2):284-296 [FREE Full text] [doi: [10.1007/s10862-015-9514-x](https://doi.org/10.1007/s10862-015-9514-x)] [Medline: [27239096](https://pubmed.ncbi.nlm.nih.gov/27239096/)]
60. Semborski S, Henwood B, Rhoades H, Mason T, Wenzel S, Rice E. Construct, concurrent, and real-world predictive validity of the Difficulties in Emotion Regulation (DERS-18) among young adults with history of homelessness. *Psychol Assess* 2021 May;33(5):385-394 [FREE Full text] [doi: [10.1037/pas0000991](https://doi.org/10.1037/pas0000991)] [Medline: [33630634](https://pubmed.ncbi.nlm.nih.gov/33630634/)]
61. Xu L, Cheng F, Liu C, Jin R, Pan H, Zhang M, et al. Psychometric properties and factor structures of the CERQ, DERS, and RESE measures: a bifactor approach. *J Pers Assess* 2021 Mar 11;103(6):797-806. [doi: [10.1080/00223891.2021.1887201](https://doi.org/10.1080/00223891.2021.1887201)] [Medline: [33703970](https://pubmed.ncbi.nlm.nih.gov/33703970/)]
62. Li J, Han ZR, Gao MM, Sun X, Ahemaitijiang N. Psychometric properties of the Chinese version of the Difficulties in Emotion Regulation Scale (DERS): factor structure, reliability, and validity. *Psychol Assess* 2018 May;30(5):e1-e9. [doi: [10.1037/pas0000582](https://doi.org/10.1037/pas0000582)] [Medline: [29756796](https://pubmed.ncbi.nlm.nih.gov/29756796/)]
63. McVey AJ, Schiltz HK, Coffman M, Antezana L, Magnus B. A preliminary psychometric analysis of the Difficulties with Emotion Regulation Scale (DERS) among autistic adolescents and adults: factor structure, reliability, and validity. *J Autism Dev Disord* 2022 Mar 22;52(3):1169-1188. [doi: [10.1007/s10803-021-05018-4](https://doi.org/10.1007/s10803-021-05018-4)] [Medline: [33886035](https://pubmed.ncbi.nlm.nih.gov/33886035/)]
64. Goldstein BL, Briggs-Gowan MJ, Greene CC, Chang R, Grasso DJ. An item response theory examination of the original and short forms of the Difficulties in Emotion Regulation Scale (DERS) in pregnant women. *J Clin Psychol* 2021 Jul 10;77(7):1591-1606 [FREE Full text] [doi: [10.1002/jclp.23167](https://doi.org/10.1002/jclp.23167)] [Medline: [33971024](https://pubmed.ncbi.nlm.nih.gov/33971024/)]
65. Benfer N, Bardeen JR, Fergus TA, Rogers TA. Factor structure and incremental validity of the original and modified versions of the Difficulties in Emotion Regulation Scale. *J Pers Assess* 2019 Aug 24;101(6):598-608. [doi: [10.1080/00223891.2018.1492927](https://doi.org/10.1080/00223891.2018.1492927)] [Medline: [30142305](https://pubmed.ncbi.nlm.nih.gov/30142305/)]
66. Staples AM, Mohlman J. Psychometric properties of the GAD-Q-IV and DERS in older, community-dwelling GAD patients and controls. *J Anxiety Disord* 2012 Apr;26(3):385-392. [doi: [10.1016/j.janxdis.2012.01.005](https://doi.org/10.1016/j.janxdis.2012.01.005)] [Medline: [22306131](https://pubmed.ncbi.nlm.nih.gov/22306131/)]
67. Beautiful Mood. URL: <https://www.beautifulmood.app/> [accessed 2022-08-03]
68. Mobile Mental Health Apps Team at the National Center for PTSD. 30 Days of Self-Care with COVID Coach: Prioritize Your Well-Being and Mental Health with the COVID Coach. 2020. URL: <https://www.ptsd.va.gov/appvid/docs/30DaysSelfCareCOVIDCoach508.pdf> [accessed 2022-10-05]

69. Smith L. Getting Started with Calm. URL: <https://support.calm.com/hc/en-us/articles/115002474567-What-s-the-best-way-to-use-Calm-Where-do-I-start-> [accessed 2022-08-03]
70. 7 Cups. Need Someone to Talk to? Our Counselors and Listeners Are Standing By. URL: <https://www.7cups.com/> [accessed 2022-10-05]
71. Lyon A, Pullmann M, Jacobson J, Osterhage K, Al AM, Renn B. Assessing the usability of complex psychosocial interventions: the Intervention Usability Scale. *Implement Res Pr* 2021;2:2633489520987828. [doi: [10.21203/rs.3.rs-20758/v3](https://doi.org/10.21203/rs.3.rs-20758/v3)]
72. Brooke J. SUS: a quick and dirty usability scale. In: Weerdmeester B, McClelland IL, Thomas B, Jordan PW, editors. *Usability Evaluation in Industry*. London, UK: Taylor and Francis; 1996:189.
73. Sevilla-Gonzalez MDR, Moreno Loeza L, Lazaro-Carrera LS, Bourguet Ramirez B, Vázquez Rodríguez A, Peralta-Pedrero ML, et al. Spanish version of the System Usability Scale for the assessment of electronic tools: development and validation. *JMIR Hum Factors* 2020 Dec 16;7(4):e21161 [FREE Full text] [doi: [10.2196/21161](https://doi.org/10.2196/21161)] [Medline: [33325828](https://pubmed.ncbi.nlm.nih.gov/33325828/)]
74. Pal D, Vanijja V. Perceived usability evaluation of Microsoft Teams as an online learning platform during COVID-19 using system usability scale and technology acceptance model in India. *Child Youth Serv Rev* 2020 Dec;119:105535 [FREE Full text] [doi: [10.1016/j.childyouth.2020.105535](https://doi.org/10.1016/j.childyouth.2020.105535)] [Medline: [33020676](https://pubmed.ncbi.nlm.nih.gov/33020676/)]
75. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. *Int J Hum-Comput Interact* 2008 Jul 30;24(6):574-594. [doi: [10.1080/10447310802205776](https://doi.org/10.1080/10447310802205776)]
76. Weiner BJ, Lewis CC, Stanick C, Powell BJ, Dorsey CN, Clary AS, et al. Psychometric assessment of three newly developed implementation outcome measures. *Implement Sci* 2017 Aug 29;12(1):108 [FREE Full text] [doi: [10.1186/s13012-017-0635-3](https://doi.org/10.1186/s13012-017-0635-3)] [Medline: [28851459](https://pubmed.ncbi.nlm.nih.gov/28851459/)]
77. 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](https://doi.org/10.4088/JCP.18r12188)] [Medline: [30900849](https://pubmed.ncbi.nlm.nih.gov/30900849/)]
78. Simon GE, Shortreed SM, Rossom RC, Beck A, Clarke GN, Whiteside U, et al. Effect of offering care management or online dialectical behavior therapy skills training vs usual care on self-harm among adult outpatients with suicidal ideation: a randomized clinical trial. *JAMA* 2022 Feb 15;327(7):630-638 [FREE Full text] [doi: [10.1001/jama.2022.0423](https://doi.org/10.1001/jama.2022.0423)] [Medline: [35166800](https://pubmed.ncbi.nlm.nih.gov/35166800/)]
79. Moberg C, Niles A, Beermann D. Guided self-help works: randomized waitlist controlled trial of Pacifica, a mobile app integrating cognitive behavioral therapy and mindfulness for stress, anxiety, and depression. *J Med Internet Res* 2019 Jun 08;21(6):e12556 [FREE Full text] [doi: [10.2196/12556](https://doi.org/10.2196/12556)] [Medline: [31199319](https://pubmed.ncbi.nlm.nih.gov/31199319/)]
80. Stirratt MJ, Dunbar-Jacob J, Crane HM, Simoni JM, Czajkowski S, Hilliard ME, et al. Self-report measures of medication adherence behavior: recommendations on optimal use. *Transl Behav Med* 2015 Dec 9;5(4):470-482 [FREE Full text] [doi: [10.1007/s13142-015-0315-2](https://doi.org/10.1007/s13142-015-0315-2)] [Medline: [26622919](https://pubmed.ncbi.nlm.nih.gov/26622919/)]
81. Simoni JM, Huh D, Wang Y, Wilson IB, Reynolds NR, Remien RH, et al. The validity of self-reported medication adherence as an outcome in clinical trials of adherence-promotion interventions: findings from the MACH14 study. *AIDS Behav* 2014 Dec 4;18(12):2285-2290 [FREE Full text] [doi: [10.1007/s10461-014-0905-x](https://doi.org/10.1007/s10461-014-0905-x)] [Medline: [25280447](https://pubmed.ncbi.nlm.nih.gov/25280447/)]
82. Kassavou A, Mirzaei V, Brimicombe J, Edwards S, Massou E, Prevost AT, et al. A highly tailored text and voice messaging intervention to improve medication adherence in patients with either or both hypertension and type 2 diabetes in a UK primary care setting: feasibility randomized controlled trial of clinical effectiveness. *J Med Internet Res* 2020 May 19;22(5):e16629 [FREE Full text] [doi: [10.2196/16629](https://doi.org/10.2196/16629)] [Medline: [32427113](https://pubmed.ncbi.nlm.nih.gov/32427113/)]
83. Lattie EG, Cohen KA, Hersch E, Williams KD, Kruzan KP, MacIver C, et al. Uptake and effectiveness of a self-guided mobile app platform for college student mental health. *Internet Interv* 2022 Mar;27:100493 [FREE Full text] [doi: [10.1016/j.invent.2021.100493](https://doi.org/10.1016/j.invent.2021.100493)] [Medline: [35036331](https://pubmed.ncbi.nlm.nih.gov/35036331/)]
84. Lau N, O'Daffer A, Yi-Frazier JP, Rosenberg AR. Popular evidence-based commercial mental health apps: analysis of engagement, functionality, aesthetics, and information quality. *JMIR Mhealth Uhealth* 2021 Jul 14;9(7):e29689 [FREE Full text] [doi: [10.2196/29689](https://doi.org/10.2196/29689)] [Medline: [34259639](https://pubmed.ncbi.nlm.nih.gov/34259639/)]
85. Marshall JM, Dunstan DA, Bartik W. Smartphone psychological therapy during covid-19: a study on the effectiveness of five popular mental health apps for anxiety and depression. *Front Psychol* 2021 Dec 13;12:775775 [FREE Full text] [doi: [10.3389/fpsyg.2021.775775](https://doi.org/10.3389/fpsyg.2021.775775)] [Medline: [35069357](https://pubmed.ncbi.nlm.nih.gov/35069357/)]
86. Di Carlo F, Sociali A, Picutti E, Pettoroso M, Vellante F, Verrastro V, et al. Telepsychiatry and other cutting-edge technologies in COVID-19 pandemic: bridging the distance in mental health assistance. *Int J Clin Pract* 2021 Jan;75(1):ijcp.13716 [FREE Full text] [doi: [10.1111/ijcp.13716](https://doi.org/10.1111/ijcp.13716)] [Medline: [32946641](https://pubmed.ncbi.nlm.nih.gov/32946641/)]
87. Torous J, Jän Myrick K, Rauseo-Ricupero N, Firth J. Digital mental health and covid-19: using technology today to accelerate the curve on access and quality tomorrow. *JMIR Ment Health* 2020 Mar 26;7(3):e18848 [FREE Full text] [doi: [10.2196/18848](https://doi.org/10.2196/18848)] [Medline: [32213476](https://pubmed.ncbi.nlm.nih.gov/32213476/)]
88. Kondylakis H, Katehakis DG, Kouroubali A, Logothetidis F, Triantafyllidis A, Kalamaras I, et al. COVID-19 mobile apps: a systematic review of the literature. *J Med Internet Res* 2020 Dec 09;22(12):e23170 [FREE Full text] [doi: [10.2196/23170](https://doi.org/10.2196/23170)] [Medline: [33197234](https://pubmed.ncbi.nlm.nih.gov/33197234/)]
89. Terry NP, Gunter TD. Regulating mobile mental health apps. *Behav Sci Law* 2018 Mar 16;36(2):136-144. [doi: [10.1002/bsl.2339](https://doi.org/10.1002/bsl.2339)] [Medline: [29659069](https://pubmed.ncbi.nlm.nih.gov/29659069/)]

## Abbreviations

**-2LL:** -2 log likelihood  
**AIC:** Akaike information criterion  
**AIM:** Acceptability Intervention Measure  
**BIC:** Bayesian information criterion  
**DERS-SF:** Difficulties in Emotion Regulation Scale – Short Form  
**GAD:** Generalized Anxiety Disorder Scale  
**HSD:** honestly significant difference  
**IAM:** Intervention Appropriateness Measure  
**ITT:** intention-to-treat  
**IUS:** Intervention Usability Scale  
**MDES:** minimum detectable effect size  
**PHQ:** Patient Health Questionnaire  
**RCT:** randomized clinical trial  
**SBQ-R:** Suicide Behaviors Questionnaire-Revised

*Edited by G Eysenbach, L Buis; submitted 04.08.22; peer-reviewed by L Martinengo; comments to author 26.08.22; revised version received 15.09.22; accepted 30.09.22; published 07.11.22*

*Please cite as:*

*Comtois KA, Mata-Greve F, Johnson M, Pullmann MD, Mosser B, Arean P  
Effectiveness of Mental Health Apps for Distress During COVID-19 in US Unemployed and Essential Workers: Remote Pragmatic  
Randomized Clinical Trial*

*JMIR Mhealth Uhealth 2022;10(11):e41689*

*URL: <https://mhealth.jmir.org/2022/11/e41689>*

*doi: [10.2196/41689](https://doi.org/10.2196/41689)*

*PMID: [36191176](https://pubmed.ncbi.nlm.nih.gov/36191176/)*

©Katherine Anne Comtois, Felicia Mata-Greve, Morgan Johnson, Michael D Pullmann, Brittany Mosser, Patricia Arean. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 07.11.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.