Contents

Original Papers

Association Between LGB Sexual Orientation and Depression Mediated by Negative Social Media Experiences: National Survey Study of US Young Adults (e23520)
César Escobar-Viera, Ariel Shensa, Jaime Sidani, Brian Primack, Michael Marshal. ................................................................. 3

Repeated Digitized Assessment of Risk and Symptom Profiles During Inpatient Treatment of Affective Disorder: Observational Study (e24066)
Maike Richter, Michael Storck, Rogério Blitz, Janik Goltermann, Juliana Seipp, Udo Dannlowski, Bernhard Baune, Martin Dugas, Nils Opel. . 8

Mobile-Assisted Cognitive Behavioral Therapy for Negative Symptoms: Open Single-Arm Trial With Schizophrenia Patients (e24406)
Eric Granholm, Jason Holden, Kristen Dwyer, Tanya Mikhail, Peter Link, Colin Depp. ................................................................. 50

Usage and Acceptability of the iBobbly App: Pilot Trial for Suicide Prevention in Aboriginal and Torres Strait Islander Youth (e14296)
Joseph Tighe, Fiona Shand, Kathy McKay, Taylor-Jai Mcalister, Andrew Mackinnon, Helen Christensen. ................................................................. 62

Visual and Personalized Quality of Life Assessment App for People With Severe Mental Health Problems: Qualitative Evaluation (e19593)
David Buitenweg, Dike van de Mheen, Jean-Paul Grund, Hans van Oers, Chijs van Nieuwenhuizen. ................................................................. 71

Evaluation of a Mobile App to Enhance Relational Awareness and Change During Cognitive Analytic Therapy: Mixed Methods Case Series (e19888)
Stephen Kellett, Katherine Easton, Martin Cooper, Abigail Millings, Melanie Simmonds-Buckley, Glenys Parry. ................................................................. 82

Changes in Stress, Anxiety, and Depression Levels of Subscribers to a Daily Supportive Text Message Program (Text4Hope) During the COVID-19 Pandemic: Cross-Sectional Survey Study (e22423)
Vincent Agyapong, Marianne Hrabok, Wesley Vuong, Reham Shalaby, Jasmine Noble, April Gusnowski, Kelly Mrklias, Daniel Li, Liana Uruchuk, Mark Snaterse, Shireen Surood, Bo Cao, Xin-Min Li, Russell Greiner, Andrew Greenshaw. ................................................................. 95

Associations Among Internet Addiction, Genetic Polymorphisms, Family Functioning, and Psychopathological Risk: Cross-Sectional Exploratory Study (e17341)
Luca Cerniglia, Silvia Cimino, Eleonora Marzilli, Esterina Pascale, Renata Tambelli. ................................................................. 106

Health Care Providers’ Perceptions of Quality, Acceptance, and Satisfaction With Telebehavioral Health Services During the COVID-19 Pandemic: Survey-Based Study (e23245)
Jesse Wright, Shyam Dewan, Donald Hilty, Naakesh Dewan. ................................................................. 130
Web-Based Relaxation Intervention for Stress During Social Isolation: Randomized Controlled Trial (e22757)
Silvia Pizzoli, Chiara Marzorati, Davide Mazzoni, Gabriella Pravettoni. ................................................................. 136

Measuring COVID-19 Related Anxiety in Parents: Psychometric Comparison of Four Different Inventories (e24507)
Christian Kubb, Heather Foran. ................................................................. 149

Real-time Mental Health Impact of the COVID-19 Pandemic on College Students: Ecological Momentary Assessment Study (e24815)
Evan Kleiman, April Yeager, Jeremy Grove, John Kellerman, Joanne Kim. ................................................................. 188

Patient Attitudes Toward Telepsychiatry During the COVID-19 Pandemic: A Nationwide, Multisite Survey (e24761)
Daniel Guinart, Patricia Marcy, Marta Hauser, Michael Dwyer, John Kane. ................................................................. 196

Viewpoints

Ethics of Digital Mental Health During COVID-19: Crisis and Opportunities (e23776)
Nicole Martinez-Martin, Ishan Dasgupta, Adrian Carter, Jennifer Chandler, Philipp Kellmeyer, Karola Kreitmair, Anthony Weiss, Laura Cabrera. ................................................................. 1

The Digital Therapeutic Alliance and Human-Computer Interaction (e21895)
Simon D’Alfonso, Reeva Lederman, Sandra Bucci, Katherine Berry. ................................................................. 23

Videoconferencing-Based Telemental Health: Important Questions for the COVID-19 Era From Clinical and Patient-Centered Perspectives (e24021)
Emil Chiauzzi, Ashley Clayton, Jina Huh-Yoo. ................................................................. 168

Flip the Clinic: A Digital Health Approach to Youth Mental Health Service Delivery During the COVID-19 Pandemic and Beyond (e24578)
Tracey Davenport, Vanessa Cheng, Frank Iorfino, Blake Hamilton, Eva Castaldi, Amy Burton, Elizabeth Scott, Ian Hickie. ................................................................. 180
Association Between LGB Sexual Orientation and Depression Mediated by Negative Social Media Experiences: National Survey Study of US Young Adults

César G Escobar-Viera1, MD, PhD; Ariel Shensa2, MA, PhD; Jaime Sidani1, PhD; Brian Primack3, MD, PhD; Michael P Marshall4, PhD
1Center for Research on Behavioral Health, Media, and Technology, Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, United States
2Department of Physical Therapy, University of Pittsburgh, Pittsburgh, PA, United States
3College of Education and Health Professions, University of Arkansas, Fayetteville, AR, United States
4Suzanne Dworak-Peck School of Social Work, University of Southern California, Los Angeles, CA, United States

Corresponding Author:
César G Escobar-Viera, MD, PhD
Center for Research on Behavioral Health, Media, and Technology
Department of Psychiatry
University of Pittsburgh
3811 O’Hara Street
Pittsburgh, PA, 15213
United States
Phone: 1 412 246 5864
Email: escobar-viera@pitt.edu

Abstract

Background: Lesbian, gay, and bisexual (LGB) persons are disproportionately affected by depression and have high social media use rates. Negative social media experiences may modify depressive symptoms among LGB persons. We sought to assess the potential influence of negative social media experiences on the association between LGB orientation and depression.

Objective: The aim of this study was to assess the potential influence of negative social media experiences on the association between LGB orientation and depression.

Methods: We performed a web-based survey of a national sample of US young adults aged 18-30 years. We assessed the respondents’ LGB orientation, negative social media experiences, and depression using the 9-item Patient Health Questionnaire. We used generalized structural equation modeling to assess both the direct and indirect effects (via negative social media experiences) of LGB orientation on depression while controlling for relevant demographic and personal characteristics.

Results: We found a conditional indirect effect (ab path) of LGB orientation on depressive symptoms via negative social media experience (a: observed coefficient 0.229; P<.001; bias-corrected bootstrapped 95% CI 0.162-0.319, and b: observed coefficient 2.158; P<.001; bias-corrected bootstrapped 95% CI 1.840-2.494). The results show that among LGB respondents, for those who reported negative social media experiences in the past year, a 1 unit increase in these experiences was associated with a 0.494 unit increase in depressive symptomatology.

Conclusions: Our results suggest that higher rates of depression among LGB young adults are partially explained by negative social media experiences; these results could help inform future patient/provider conversations about mental health risk and protective factors related to social media use. Reducing these experiences and increasing positive social media experiences among LGB persons may mitigate depressive symptomatology in this population.

(JMIR Ment Health 2020;7(12):e23520) doi:10.2196/23520

KEYWORDS
social media; depression; mental health; sexual minorities; minority stress; GSEM; survey; young adult; adolescent; LGBTQ
**Introduction**

Lesbian, gay, and bisexual (LGB) persons are at higher risk of experiencing depression than non-LGB individuals [1-4]. Although the cause of depression is multifactorial, the minority stress model [5] posits that LGB people are exposed to a number of environment- and community-related stressors (eg, discrimination, violence) that interact with person-specific stressors and coping skills to influence mental health outcomes. Interestingly, the dramatic rise of social media use in the past decade has profoundly changed the nature of our social relationships, potentially influencing stressors and coping skills originally described in the minority stress framework. Thus, it is important to understand whether the tenets of the minority stress model generalize to other forms of social stressors, such as those experienced through social media use.

Social media use has been consistently increasing in the United States; approximately 92% of young adults currently report that they have at least one social media account [6,7]. Among LGB individuals, social media websites and mobile apps are now a primary way of connecting with others and building relationships [8,9], sometimes replacing interactions that used to mainly occur face-to-face. LGB young adults may turn to social media to compensate for perceived lack of resources and support in their immediate environment [10]. They may also use social media to learn about their own sexual orientation [11], find LGB-specific resources and services [12], find new friends or partners [13], and find social support [14].

Given the relatively recent growth in social media use in the United States, research conducted on the effects of social media on mental health outcomes found that the content and quality of social media experiences and interactions may be important to our understanding of the roles of social media in myriad mental health outcomes [15-18]. For LGB persons, negative social media experiences may play a critical role in depression disparities. These experiences may include exchanges that make individuals feel sad, depressed, or angry, or viewing content that negatively influences their emotional status. Despite their mostly positive motivations for using social media, a content analysis found that LGB individuals reported a higher frequency of negative effects of social media experiences on their well-being than non-LGB persons [19].

Assessing if and to what extent negative social media experiences are associated with depressive symptoms among LGB young adults may help identify intervention points for potential ways to improve or prevent worsening of depressive symptoms. It may also provide initial evidence that can enable mental health professionals to raise awareness among their patients about the importance of monitoring and improving their social media experiences. Therefore, this research used cross-sectional data from a national web-based survey of US young adults to examine (1) differences in depressive symptoms between LGB and heterosexual young adults and (2) determine whether negative social media experiences partially explain LGB depression disparities.

**Methods**

**Participants and Procedures**

In March 2018, we commissioned Qualtrics Sample Services to recruit a national sample of US adults aged 18-30 years. Recruitment sought to reflect the sociodemographic characteristics of the 2010 US Census. A total of 2408 study participants responded to a web-based survey developed using Qualtrics Online Surveys. The survey was active for 30 days after invitations were sent. Respondents received points that were redeemable for incentives of their choice, such as gift cards or charitable contributions. Qualtrics Sample Services, in conjunction with the research team, conducted a series of procedures to enhance data quality. First, the survey instrument was pilot-tested with a sample of 30 individuals who were not included in the final study sample to ensure that skip patterns were working correctly and to review data for inconsistencies. Next, a “soft launch” was conducted, during which 10% of the target sample (n=240) completed the survey. Again, the data were reviewed for inconsistencies and to ensure that quotas were being reached. Finally, as the final study sample was being recruited, data were reviewed to remove straight-liners, speeders, and responses with duplicate IP addresses. A total of 94 responses were removed due to low quality data based on these procedures. All study procedures were approved by the University of Pittsburgh Institutional Review Board.

**Measures**

**Demographics**

Sociodemographic characteristics included assigned sex at birth, age, highest education level achieved, relationship status, and living arrangement. These factors were included because they may be associated with different measures of social media use and depressive symptoms [16,20]. Assigned sex at birth was assessed as female or male. Age was measured in years. Educational level was assessed with five categories: some high school, high school graduate, some college or technical institute, college graduate, and graduate school. For these analyses, the categories were collapsed into “some college or less” and “college graduate or more.” We used three categories to assess relationship status: single, member of an unmarried couple, and married. Living arrangement assessed the people with whom the respondent was living at the time of the survey, and it included four categories: by myself, parent or guardian, significant other, and other arrangement.

**LGB Orientation**

Self-reported sexual minority orientation was assessed with the single item “Do you consider yourself to be?” and the following options: exclusively straight/heterosexual, mostly straight/heterosexual, bisexual, mostly gay or lesbian, and queer. This item has been extensively used in the literature and is suggested as a best practice for population surveys [21,22]. For this analysis, respondents who described themselves as exclusively straight/heterosexual and mostly straight/heterosexual were categorized as non-LGB, and the remaining four groups were categorized as LGB.
**Depression**

The 9-item Patient Health Questionnaire (PHQ-9) was used to assess depressive symptoms over the last 14 days [23]. Items were scored using a 4-point Likert-type scale with scores ranging from 0 to 3 (“not at all” to “nearly every day”). We summed the scores for all responses and created a continuous scale ranging from 0-27.

**Social Media Use Time**

We asked respondents to provide an estimate, in hours and minutes, of the amount of time they used social media for personal use unrelated to work during a single day. The response options included 0 to 23 hours and 0, 15, 30, and 45 minutes [24,25]. For this analysis, the responses were combined to create a composite, continuous single measure of social media time use.

**Negative Social Media Experiences**

We used four items to assess negative social media experiences. These were adapted from prior research that assessed sexual minority victimization [26,27]. Respondents were asked: “Please think about your experiences on social media over the past year. How often: (a) were you called out or hurt by one of your social media contacts/friends? (b) have you posted something and received negative feedback? (c) have you posted something and received no feedback at all? and (d) have you seen posts or pictures that made you realize you were not invited to a peer’s activity/party?” Response options used a Likert-type scale that included 1, never; 2, once or twice; 3, a few times a month; 4, about once a week, and 5, more often. Scores were summed and averaged, generating a continuous score from 1 to 5. Given that these items were adapted for this study, we conducted further principal component and reliability analysis prior to model inclusion.

**Data Analysis**

Our initial sample (N=2408) was described using mean (SD) for variables measured on a continuous scale and n (%) for variables measured on a categorical scale. To assess if negative social media experiences, depression, and all covariates varied by LGB orientation (yes/no), we used one-way analysis of variance (ANOVA) for continuous variables and chi-square tests for categorical variables.

While previous research on victimization of LGB individuals [26,27] informed the construct of negative social media experiences, we conducted a principal component analysis (PCA) to assess the underlying factor structure of our particular set of adapted items. The PCA of the negative social media experiences scale helped determine a single-factor structure. This factor accounted for 66% of the total variance observed, with an eigenvalue of 2.62. Factor loadings ranged from 0.86 (for “Were you called out or hurt by one of your social media contacts/friends?”) to 0.77 (for “Have you seen posts or pictures that made you realize you were not invited to a peer’s activity/party?”). The Cronbach alpha of .81 showed strong reliability of the scale overall.

After completing these initial steps, we used generalized structural equation modeling (GSEM) to construct two path models. The first model assessed a conditional direct path between LGB orientation and depressive symptoms while controlling for sociodemographic factors and social media time use. Among categorical control variables, referent groups were selected based on groups that are less likely to report depressive symptoms (eg, people who are married, are living with their significant other, and have a higher educational level). Then, we fitted a second path model, constructing a path between LGB orientation and negative social media experience as well as between negative social media experiences and depressive symptoms. For this model, we also controlled for sociodemographic factors and social media time use. We used full information maximum likelihood listwise deletion to account for missing data found within the variables included in our models. This resulted in 2336 responses to fit both the conditional direct and indirect models. Our choice of GSEM was appropriate because GSEM accounts for the Likert-type scales used to assess both depressive symptoms and negative social media experience. Because of the relatively recent use of GSEM to assess mediation paths, we obtained biased-corrected bootstrap estimates and CIs using 200 bootstrapped samples to assess the significance of both direct and indirect path models. Finally, we manually computed the proportion of the total effect that corresponded to both direct and indirect effects of LGB orientation on depressive symptoms. All statistical analyses were conducted using Stata 15 [28].

**Results**

Approximately half of our initial sample was female (1223/2408, 50.8%), and the majority were aged 25-30 years (1948/2408, 80.9%) and of White, non-Hispanic race/ethnicity (1607/2408, 68.2%). A total of 497/2408 (20.6%) of the sample reported LGB orientation. The characteristics of the complete sample are presented in Table 1. Depression and negative social media experiences scores differed significantly between LGB and non-LGB individuals (P<.001 in both cases), with higher mean scores for LGB participants for both scales. Additionally, time per day on social media, education, and relationship status differed between LGB and non-LGB individuals.
Table 1. Characteristics of the sample and associations with LGB orientation (N=2408).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Whole sample</th>
<th>LGB(^a) (n=497, 20.6%)</th>
<th>Non-LGB (n=1911, 79.4%)</th>
<th>(P) value(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression, mean (SD)</td>
<td>6.2 (5.6)</td>
<td>8.4 (6.3)</td>
<td>5.6 (5.2)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Negative social media experiences, mean (SD)</td>
<td>1.9 (0.8)</td>
<td>2.1 (0.8)</td>
<td>1.8 (0.8)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time per day on social media, mean (SD)</td>
<td>3.1 (2.9)</td>
<td>3.7 (3.1)</td>
<td>2.9 (2.9)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Sex assigned at birth, n (%)(^c)</td>
<td></td>
<td></td>
<td></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Male</td>
<td>1184 (49.2)</td>
<td>197 (39.6)</td>
<td>987 (51.7)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1223 (50.8)</td>
<td>300 (60.4)</td>
<td>923 (48.3)</td>
<td></td>
</tr>
<tr>
<td>Age (years), n (%)</td>
<td></td>
<td></td>
<td></td>
<td>.89</td>
</tr>
<tr>
<td>18-24</td>
<td>460 (19.1)</td>
<td>96 (19.3)</td>
<td>364 (19.1)</td>
<td></td>
</tr>
<tr>
<td>25-30</td>
<td>1948 (80.9)</td>
<td>401 (80.7)</td>
<td>1547 (81.0)</td>
<td></td>
</tr>
<tr>
<td>Race/ethnicity, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>.15</td>
</tr>
<tr>
<td>White, non-Hispanic</td>
<td>1607 (68.2)</td>
<td>326 (67.6)</td>
<td>1281 (68.4)</td>
<td></td>
</tr>
<tr>
<td>Black, non-Hispanic</td>
<td>180 (7.6)</td>
<td>35 (7.3)</td>
<td>145 (7.7)</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>339 (14.4)</td>
<td>84 (17.4)</td>
<td>255 (13.6)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>198 (8.4)</td>
<td>32 (6.6)</td>
<td>166 (8.9)</td>
<td></td>
</tr>
<tr>
<td>Other(^d)</td>
<td>32 (1.4)</td>
<td>5 (1.0)</td>
<td>27 (1.4)</td>
<td></td>
</tr>
<tr>
<td>Education, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>.002</td>
</tr>
<tr>
<td>Some college or less</td>
<td>1078 (44.8)</td>
<td>252 (50.9)</td>
<td>826 (43.3)</td>
<td></td>
</tr>
<tr>
<td>College graduate or more</td>
<td>1327 (55.2)</td>
<td>243 (49.1)</td>
<td>1084 (56.8)</td>
<td></td>
</tr>
<tr>
<td>Relationship status, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Married</td>
<td>737 (30.6)</td>
<td>111 (22.3)</td>
<td>626 (32.8)</td>
<td></td>
</tr>
<tr>
<td>Member of unmarried couple</td>
<td>623 (25.9)</td>
<td>167 (33.6)</td>
<td>456 (23.9)</td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>1046 (43.5)</td>
<td>219 (44.1)</td>
<td>827 (43.3)</td>
<td></td>
</tr>
<tr>
<td>Living with, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>.46</td>
</tr>
<tr>
<td>Significant other</td>
<td>1103 (45.8)</td>
<td>213 (42.9)</td>
<td>890 (46.6)</td>
<td></td>
</tr>
<tr>
<td>Parent/guardian</td>
<td>505 (21.0)</td>
<td>114 (22.9)</td>
<td>391 (20.5)</td>
<td></td>
</tr>
<tr>
<td>Alone</td>
<td>444 (18.5)</td>
<td>93 (18.7)</td>
<td>351 (18.4)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>355 (14.8)</td>
<td>77 (15.5)</td>
<td>278 (14.6)</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)LGB: lesbian, gay, and bisexual.

\(^b\)Significance derived from analysis of variance for characteristics measured on a continuous scale and chi-square test for characteristics measured on a categorical scale.

\(^c\)Column percentages may not equal 100 due to rounding.

\(^d\)Included American Indian or Alaskan Native, Native Hawaiian or other Pacific Islander, and multiracial, non-Hispanic.

Our first path c model sought to assess LGB disparities by examining the relationship between LGB orientation and depression while controlling for respondents’ social media time use, sex assigned at birth, age, race, educational level, relationship status, and living arrangement. The results are presented in Table 2. This path c model showed a positive, statistically significant relationship between LGB orientation and higher scores of depressive symptoms, with an observed coefficient of 2.160 (\(P<.001\)) and a bias-corrected bootstrapped 95% CI of 1.590-2.768.
Table 2. Generalized structural equation model (GSEM) results for conditional direct effects of LGB orientation and depressive symptoms (path c model) (N=2336).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Observed coefficient</th>
<th>Bootstrapped standard error</th>
<th>BCB (95% CI)(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LGB(^b) orientation</td>
<td>2.159</td>
<td>0.287</td>
<td>1.590 to 2.768</td>
</tr>
<tr>
<td>Social media time use</td>
<td>0.286</td>
<td>0.047</td>
<td>0.212 to 0.391</td>
</tr>
<tr>
<td>Female sex assigned at birth(^c)</td>
<td>0.636</td>
<td>0.227</td>
<td>0.182 to 1.046</td>
</tr>
<tr>
<td>Age 18-24 years(^d)</td>
<td>–0.025</td>
<td>0.293</td>
<td>–0.587 to 0.548</td>
</tr>
<tr>
<td>Race/ethnicity(^e)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black, non-Hispanic</td>
<td>–0.651</td>
<td>0.408</td>
<td>–1.340 to 0.274</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.045</td>
<td>0.340</td>
<td>–0.594 to 0.765</td>
</tr>
<tr>
<td>Asian</td>
<td>0.331</td>
<td>0.372</td>
<td>–0.334 to 1.188</td>
</tr>
<tr>
<td>Other</td>
<td>–0.476</td>
<td>0.859</td>
<td>–2.154 to 1.251</td>
</tr>
<tr>
<td>Education level(^f)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some college or less</td>
<td>1.470</td>
<td>0.249</td>
<td>0.908 to 1.968</td>
</tr>
<tr>
<td>Relationship status(^g)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Member of unmarried couple</td>
<td>0.257</td>
<td>0.295</td>
<td>–0.371 to 0.822</td>
</tr>
<tr>
<td>Single</td>
<td>0.251</td>
<td>0.411</td>
<td>–0.466 to 1.056</td>
</tr>
<tr>
<td>Living with(^h)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent/guardian</td>
<td>0.935</td>
<td>0.389</td>
<td>0.306 to 1.834</td>
</tr>
<tr>
<td>Alone</td>
<td>1.002</td>
<td>0.425</td>
<td>0.247 to 1.979</td>
</tr>
<tr>
<td>Other</td>
<td>0.359</td>
<td>0.416</td>
<td>–0.496 to 1.090</td>
</tr>
</tbody>
</table>

\(^a\)Bias-corrected bootstrapped 95% CI (200 bootstrapped samples).
\(^b\)LGB: lesbian, gay, and bisexual.
\(^c\)Reference was Male.
\(^d\)Reference was 25-30 years of age.
\(^e\)Reference was White.
\(^f\)Reference was College graduate or more.
\(^g\)Reference was Married.
\(^h\)Reference was Living with significant other.

In the second path model (Table 3), we sought to determine whether negative social media experiences significantly explained potential higher rates of LGB depression while controlling for the same demographic and personal variables specified in the previous model. We found a statistically significant, conditional indirect effect (ab path) of LGB orientation on depressive symptoms via negative social media experiences (a: observed coefficient 0.229; \(P<.001\); bias-corrected bootstrapped 95% CI 0.162-0.319, and b: observed coefficient 2.158; \(P<.001\); bias-corrected bootstrapped 95% CI 1.840-2.494).
Table 3. Generalized structural equation model (GSEM) results for conditional indirect effects of LGB orientation on depressive symptoms (path ab-c’ model) (N=2336).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Observed coefficient</th>
<th>Bootstrapped standard error</th>
<th>BCB (95% CI)a</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Path a (Negative social media experiences)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGB orientation</td>
<td>0.229</td>
<td>0.039</td>
<td>0.162 to 0.319</td>
</tr>
<tr>
<td>Social media time use</td>
<td>0.070</td>
<td>0.008</td>
<td>0.055, 0.084</td>
</tr>
<tr>
<td>Female sex assigned at birthc</td>
<td>-0.190</td>
<td>0.030</td>
<td>-0.251 to 0.135</td>
</tr>
<tr>
<td>Age 18 to 24d</td>
<td>0.146</td>
<td>0.048</td>
<td>0.057 to 0.251</td>
</tr>
<tr>
<td><strong>Race/ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black, non-Hispanic</td>
<td>0.127</td>
<td>0.076</td>
<td>-0.001 to 0.311</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.045</td>
<td>0.049</td>
<td>-0.129 to 0.054</td>
</tr>
<tr>
<td>Asian</td>
<td>0.072</td>
<td>0.059</td>
<td>-0.038 to 0.203</td>
</tr>
<tr>
<td>Other</td>
<td>-0.050</td>
<td>0.145</td>
<td>-0.314 to 0.260</td>
</tr>
<tr>
<td><strong>Education level</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some college or less</td>
<td>0.071</td>
<td>0.037</td>
<td>0.007 to 0.135</td>
</tr>
<tr>
<td><strong>Relationship status</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Member of unmarried couple</td>
<td>-0.041</td>
<td>0.043</td>
<td>-0.142 to 0.033</td>
</tr>
<tr>
<td>Single</td>
<td>-0.021</td>
<td>0.054</td>
<td>-0.130 to 0.066</td>
</tr>
<tr>
<td><strong>Living with</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent/guardian</td>
<td>-0.008</td>
<td>0.058</td>
<td>-0.110 to 0.126</td>
</tr>
<tr>
<td>Alone</td>
<td>0.145</td>
<td>0.062</td>
<td>0.035 to 0.286</td>
</tr>
<tr>
<td>Other</td>
<td>-0.051</td>
<td>0.054</td>
<td>-0.150 to 0.070</td>
</tr>
<tr>
<td>Parent/guardian</td>
<td>-0.008</td>
<td>0.058</td>
<td>-0.110 to 0.126</td>
</tr>
<tr>
<td><strong>Path bc’ (Depressive symptoms)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGB orientation</td>
<td>1.658</td>
<td>0.314</td>
<td>1.068 to 2.252</td>
</tr>
<tr>
<td>Negative social media experiences</td>
<td>2.158</td>
<td>0.176</td>
<td>1.840 to 2.494</td>
</tr>
<tr>
<td>Social media time use</td>
<td>0.134</td>
<td>0.047</td>
<td>0.054 to 0.233</td>
</tr>
<tr>
<td>Female sex assigned at birthc</td>
<td>1.040</td>
<td>0.221</td>
<td>0.691 to 1.536</td>
</tr>
<tr>
<td>Age 18 to 24d</td>
<td>-0.359</td>
<td>0.308</td>
<td>-0.943 to 0.189</td>
</tr>
<tr>
<td><strong>Race/ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black, non-Hispanic</td>
<td>-0.856</td>
<td>0.447</td>
<td>-1.635 to 0.183</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.165</td>
<td>0.352</td>
<td>-0.438 to 1.048</td>
</tr>
<tr>
<td>Asian</td>
<td>0.191</td>
<td>0.362</td>
<td>-0.527 to 1.090</td>
</tr>
<tr>
<td>Other</td>
<td>-0.366</td>
<td>1.064</td>
<td>-2.468 to 1.697</td>
</tr>
<tr>
<td><strong>Education level</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some college or less</td>
<td>1.326</td>
<td>0.237</td>
<td>0.891 to 1.872</td>
</tr>
<tr>
<td><strong>Relationship status</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Member of unmarried couple</td>
<td>0.342</td>
<td>0.292</td>
<td>-0.190 to 0.878</td>
</tr>
<tr>
<td>Single</td>
<td>0.297</td>
<td>0.379</td>
<td>-0.345 to 1.036</td>
</tr>
<tr>
<td><strong>Living with</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent/guardian</td>
<td>0.932</td>
<td>0.388</td>
<td>0.224 to 1.748</td>
</tr>
<tr>
<td>Alone</td>
<td>0.725</td>
<td>0.399</td>
<td>-0.244 to 1.401</td>
</tr>
</tbody>
</table>
Moreover, the product of multiplying the observed coefficients of paths a and b in our ab-c’ model (Figure 1) showed a conditional indirect effect of 0.494 (approximately 23% of the total effect of LGB orientation on depressive symptomatology), with a direct effect (path c’) equal to 1.658. This indirect effect suggests that among LGB respondents in our national sample of US young adults, and for those who reported negative social media experiences in the past year, a 1 unit increase in these experiences is associated with a 0.494 unit increase in reporting of depressive symptomatology (Figure 1).

Figure 1. Path model (ab-c’) presenting observed coefficients of the conditional effect of LGB orientation on depressive symptoms via negative social media experiences. LGB: lesbian, gay, and bisexual. *P<.001.

Discussion

Principal Findings

Sexual minority individuals face well-known mental health disparities, including disparities in depression. In this national web-based survey of US young adults, we found a positive, strong association of LGB orientation with depressive symptoms when controlling for demographic and personal characteristics. We also found that compared to their non-LGB peers, LGB respondents reported significantly more negative social media experiences in the previous 12 months. Importantly, our results suggest that the association of LGB orientation with depression might have an indirect path via this negative social media experiences. Therefore, reducing negative social media experiences among LGB persons may mitigate depressive symptomatology in this population and potentially reduce LGB depression disparities.

Main Findings

Our analyses showed a direct relationship between LGB orientation and higher PHQ-9 depression scores. These results echo a plethora of previous research that found disparities in depression between this population and their heterosexual peers [4,29,30]. According to the minority stress framework [5], external LGB-related stressors such as social rejection and victimization are associated with more individual-based, proximal stressors (eg, internalized homonegativity), which in turn interact with an individual’s coping skills and social support gathered from the community to buffer the negative impact on their mental health. Interestingly, the defining affordances of social media platforms deeply transformed social interactions. For example, asynchronicity of social media conversations and interactions greatly modifies the ability of the people involved in those interactions to take cues, interpret subtleties, and process the social media experience itself [31]. Therefore, researchers must continue to study the influence of these social media experiences on individuals’ mental health, especially among groups already recognized as mental health disparity populations, such as sexual and gender minority groups (eg, LGB and transgender individuals).

In our national sample of US young adults, we found that LGB respondents had significantly higher scores of negative social media experiences in the previous 12 months compared to non-LGB peers. Our results also showed an indirect relationship between LGB orientation and severity of depressive symptoms via negative social media experiences. While these results

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Observed coefficient</th>
<th>Bootstrapped standard error</th>
<th>BCB (95% CI)&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other</td>
<td>0.467</td>
<td>0.355</td>
<td>-0.119 to 1.355</td>
</tr>
</tbody>
</table>

<sup>a</sup>Bias-corrected bootstrapped 95% CI (200 bootstrapped samples).

<sup>b</sup>LGB: lesbian, gay, and bisexual.

<sup>c</sup>Reference was Male.

<sup>d</sup>Reference was 24-30 years of age.

<sup>e</sup>Reference was White.

<sup>f</sup>Reference was College graduate or more.

<sup>g</sup>Reference was Married.

<sup>h</sup>Reference was Living with significant other.
confirm previous research reporting higher frequency of web-based discrimination, cyberbullying, and rejection among LGB populations [32], they expand previous findings by identifying a potential mediating role of social media experiences in depression among this group. While one of the items included in our negative social media experiences scale (“How often were you called out or hurt by one of your social media contacts/friends?”) alluded to situations that might occur both on the web or in the physical environment, the other three items (“How often have you posted something and received negative feedback?” “How often have you posted something and received no feedback at all?” and “Have you seen posts or pictures that made you realize you were not invited to a peer’s activity/party?”) not only highlight an experience of rejection but also allude to the unique asynchronous nature of the social media experience.

Given our results and the prevalence of depression and social media use among LGB persons, we suggest interventions that seek to empower LGB young people to reduce negative social media experiences or even eliminate experiences these before they happen. Although to our knowledge, these interventions are currently few or nonexistent, results of prior research [33] suggest that sexual and gender minority youth go through a series of decision points while navigating negativity on the web that could be responsive to behavioral interventions. Given the affordances of social media [34], future research must carefully assess a wider array of LGB-specific negative social media experiences to determine whether these experiences represent a distinctive stressor in the minority stress framework. Research should also determine the extent to which the risk posed by social media experiences to the mental health of LGB persons is comparable to the risk posed by discriminatory experiences in the physical environment. For example, permanent accessibility of social media content, as well as the ability to quantify content popularity, may render a seeming “one-time” negative social media interaction more permanent, searchable, shareable, and “likeable” far beyond the original intent. Potential educational interventions could help social media users become familiar with social media account management, vetting of new connections, and use of privacy settings and platform features to protect users’ privacy and make the social media experience less stressful and more enjoyable.

Interventions could also focus on increasing positive social media experiences. This could be achieved by creating and joining supportive social media communities with strong community guidelines, enhancing the ability of LGB youth to make new and safer social media connections, and providing LGB-specific mental health resources, which are among the top needs of LGB youth [35,36]. Prior research has described the importance of synchronous, text-based internet-based interactions for sexual and gender minority youth to find safe, affirming, and supportive spaces on the web that may not exist in their offline lives [37,38]. From a clinical standpoint, although these findings are preliminary, mental health professionals may want to specifically address social media experiences among youth and emerging adults in general and sexual minority persons in particular. Our findings indicate that it may be important to ask about apparently subtle circumstances in which an LGB person may feel ignored and discriminated against (eg, lack of positive reinforcement after posting one’s photograph with a same-sex partner), which could have a lasting impact in an individual’s emotional well-being.

Limitations
Our study has a number of limitations to consider. While the proposed direction of our analysis was supported by the tenets of the minority stress framework, the lack of longitudinal individual-level data on the variables used in our analysis precludes us from making strong causal inferences from our findings. Moreover, our analyses relied on data from questions that asked participants to report about their social media experiences over a period of 12 months before the survey; because of this, some respondents might have faced limitations when thinking back to the social media experiences they had. However, this is the first study assessing the role of negative social media experiences on mental health outcomes among LGB individuals, a well-known health disparity population. This study could help guide future longitudinal research looking at inter- and intra-individual change over time in both negative social media experiences and depression among LGB populations.

Our conceptualization of negative social media experiences was focused on personal interactions (or the lack thereof), and the questions were not specifically developed to thoroughly assess the components of minority stress. Thus, the items might have not captured the entirety of this construct or its relevance in the broader context of minority stress. While our scale was reliable in assessing negative social media experiences related to interactions between individuals and their social media contacts, we did not assess experiences related to interactions with LGB-specific social media content (eg, LGB-specific negative or discriminatory content within friends’ posts, photographs, or videos posted on individuals’ social media newsfeeds) or the potential differential impact of negative social media experiences on a person-to-person basis compared to a person-to-many (ie, social media groups) basis. However, our results confirm and expand previous research showing higher frequency of negative experiences on the web among LGB persons than among non-LGB individuals. Because of this, and given the unique affordances of social media [31,39], such as asynchronicity and permanence, it is important to further explore these nuances as they apply specifically to the social media experiences of LGB persons and to account for similar experiences in the offline environment.

Although the larger study from which this study was conducted recruited nationally, the respondents’ age skewed toward older young adults, and recruitment did not focus on a nationally representative sample of LGB respondents. Thus, the results of this study are not necessarily generalizable to all groups of young adults or LGB young adults across the United States. Future research in this area should aim to recruit a nationally representative sample focused solely on LGB young adults.

Additionally, we did not assess other social media experiences that may also influence mental health. Previous research has identified a number of social media–related experiences and factors that are associated with positive mental health outcomes.
among LGB individuals [14,40,41]. These include, but are not limited to, seeking and perceiving the availability of social support on social media, establishing new connections and friendships on the web, and creating network structures of social media friends. Given that negative social media experiences do not occur isolated from other experiences related to social media, future research will benefit from accessing these complex interactions not only longitudinally but potentially also in real time. This will be an important component of understanding the role of social media interactions in LGB individuals’ mental health, specifically depression.

Conclusions

Our study is part of a growing body of literature focused on analyzing the potential effects of negative social media experiences (one aspect of social media use) on depressive symptoms. The authors’ perspective was one of understanding negative social media experiences, as opposed to previous research that solely focused on the volume of negative social media experience use (ie, amount of time spent on negative social media experiences). This research is critical for two reasons: First, there are no signs that young adults’ negative social media use will decrease in the foreseeable future. Therefore, it is necessary to grasp the subtleties of negative social media experience interactions and their influence on individuals’ mental health. Second, understanding these nuances among minority groups, such as LGB persons, will help inform potential expansion of existing theories that seek to explain mental health disparities within subpopulations of young adults. The results of this study will help inform future patient/provider conversations about mental health risk and protective factors related to negative social media experience use as well as the development of interventions that seek to improve the experiences and mental health of sexual minority populations.

Acknowledgments

We gratefully acknowledge funding from an NIH/NIMHD K99/R00 “Pathway to Independence” Award (grant number 5K99MD012813-02) and The Fine Foundation (grant number 171R32SM-2018) for the completion of this research. The funding sources played no role in the design of the study; collection, analysis, and interpretation of data; writing of the report; or decision to submit the article for publication.

Conflicts of Interest

None declared.

References


https://mental.jmir.org/2020/12/e23520

JMIR Ment Health 2020 | vol. 7 | iss. 12 | e23520 | p.11

(None for citation purposes)


Abbreviations

LGB: lesbian, gay, and bisexual
GSEM: generalized structured equation model
PHQ-9: 9-item Patient Health Questionnaire
Ethics of Digital Mental Health During COVID-19: Crisis and Opportunities

Nicole Martinez-Martin1*, JD, PhD‡; Ishan Dasgupta2*, JD, MPH; Adrian Carter3, PhD; Jennifer A Chandler4, LLM; Philipp Kellmeyer5, MPhil, MD; Karola Kreitmair6, PhD, MSc; Anthony Weiss7, MD; Laura Y Cabrera8, PhD

1Department of Pediatrics, Center for Biomedical Ethics, School of Medicine, Stanford University, Stanford, CA, United States
2Department of Philosophy, University of Washington, Seattle, WA, United States
3School of Psychological Sciences and the Turner Institute for Brain and Mental Health, Monash University, Melbourne, Australia
4Faculty of Law, Centre for Health Law, Policy & Ethics, University of Ottawa, Ottawa, ON, Canada
5Neuroethics and AI Ethics Lab Department of Neurorsurgery, University Medical Center Freiburg, Freiburg, Germany
6Department of Medical History and Bioethics, School of Medicine and Public Health, University of Wisconsin, Madison, WI, United States
7Department of Psychiatry and Center for Bioethics, Harvard Medical School, Boston, MA, United States
8Center for Ethics & Humanities in the Life Sciences, Department of Translational Neuroscience, Michigan State University, East Lansing, MI, United States

‡Emerging Issues Task Force, International Neuroethics Society

*these authors contributed equally

Corresponding Author:
Nicole Martinez-Martin, JD, PhD
Department of Pediatrics, Center for Biomedical Ethics, School of Medicine
Stanford University
1215 Welch Road
Modular A
Stanford, CA, 94305
United States
Phone: 1 6507235760
Email: nicolemz@stanford.edu

Abstract

Social distancing measures due to the COVID-19 pandemic have accelerated the adoption and implementation of digital mental health tools. Psychiatry and therapy sessions are being conducted via videoconferencing platforms, and the use of digital mental health tools for monitoring and treatment has grown. This rapid shift to telehealth during the pandemic has given added urgency to the ethical challenges presented by digital mental health tools. Regulatory standards have been relaxed to allow this shift to socially distanced mental health care. It is imperative to ensure that the implementation of digital mental health tools, especially in the context of this crisis, is guided by ethical principles and abides by professional codes of conduct. This paper examines key areas for an ethical path forward in this digital mental health revolution: privacy and data protection, safety and accountability, and access and fairness.

(JMIR Ment Health 2020;7(12):e23776) doi:10.2196/23776

KEYWORDS
ethics; digital mental health; neuroethics; mental health; COVID-19; crisis; opportunity; implementation; online tool; telehealth

Introduction

COVID-19 is presenting a mental health crisis of unprecedented scale, due to the social and psychological burdens stemming from the pandemic, including social isolation, widespread unemployment, worries over contracting the virus, insomnia, social media exposure, and the rising death toll [1,2]. Mental health studies conducted during the pandemic have confirmed that symptoms of acute stress, anxiety, and depression, as well as suicidality, have been increasing [3-5]. Pre-existing mental health and health issues, secondary stressors such as job loss, and greater exposure to pandemic-related media coverage are the factors most strongly associated with increases in depressive and acute stress symptoms [6]. Social distancing measures also contribute to cognitive decline, substance abuse, and other mental health problems [7,8]. As is often the case, these negative
effects disproportionately affect vulnerable groups in society, such as older adults, racial and ethnic minorities, people living with disabilities, people who are neurologically atypical, children, and people who are homeless [9-11]. People whose illness is serious enough to require admission to an intensive care unit may experience lingering trauma [12,13], and health care workers are experiencing extreme distress [14].

To address this mental health crisis and provide physically distanced care, there has been a major, accelerated move toward adoption and implementation of digital tools for mental health care [15-17]. Psychiatry and therapy sessions are being conducted via videoconferencing platforms, and digital tools for diagnosis, monitoring, and treatment, such as mental health apps, are increasing in use [18]. As the pandemic continues, telehealth is increasingly becoming the “new normal” in mental health care. Many in the mental health field have noted the upside to the increased use of digital mental health tools, such as broader access and lower rates of “no-shows” to appointments. Yet, given that privacy and safety regulations in the United States were relaxed to facilitate this new stage of mental health care [19] and the large scale in which these tools are being used, there is renewed urgency to assess and address ethical issues presented by digital mental health tools, such as lack of evidence of efficacy; privacy and data protection; access; and fairness, transparency, and accountability [20-22].

The urgent need for socially distanced mental health care should not be used to erode the regulations and practices that protect people from substandard mental health care or unwanted uses of their personal information. During this public health emergency, the need to protect people from infection by increasing tools for socially distanced care could certainly be seen as having greater weight than protecting privacy or cautious oversight regarding safety or effectiveness. Nevertheless, it remains important to examine the burdens, as well as benefits, and assess the appropriateness of the trade-offs being made to permit expanded digital mental health care. This paper examines key areas for addressing ethical digital mental health care in the wake of COVID-19: privacy and data protection, safety and accountability, and access and fairness.

Privacy and Data Protection

Although privacy and data protection are important ethical considerations for health technologies, in digital mental health technologies, they become major concerns due to the sensitive personal information people share through digital mental health tools, as well as the potential for government and corporate misuse and surveillance. Mental health data is considered more sensitive than other health data [23] and can often be more personal or stigmatizing in nature. In the United States, to facilitate the provision of telehealth, the Office of Civil Rights at the Department of Health and Human Services issued a notification on March 30, 2020, altering the Health Information Portability and Accountability Act (HIPAA) Privacy Rule to eliminate penalties for health care providers for violations during the good faith provision of telehealth [24]. This rule change is limited to the provision of telehealth through non–public-facing communication products such as Zoom (Zoom Video Communications, Inc) [25]. Privacy issues that arose in Zoom, a videoconferencing platform used by some health providers for telemedicine, since the pandemic [26] highlight the need to assess and address data practices of digital platforms before using them for sensitive purposes such as the provision of health care.

In effect, after this rule change, a large portion of health personal information that is communicated through or stored on digital tools for mental health care is now not covered by HIPAA. HIPAA protections apply to “covered entities” including health care providers and institutions but not to telehealth technologies including apps that collect information directly from consumers [27]. Given the major shift toward the use of telehealth and apps for mental health care, an increasing portion of sensitive mental health information is now shared and stored by technologies that do not provide protections for that information. Even nonhealth digital data collected on a smartphone, such as location, can be used to generate highly personal behavioral and health information such as evidence of depressive or manic episodes, psychosis, or onset of Parkinson Disease [28-30]. Developers of mental health apps have been shown to engage in misleading privacy and data practices, sharing users’ personal data with third parties without their consent [31]. Behavioral health information is a valuable commodity, and it is likely that companies will take further advantage of the lax security and privacy landscape [32]. The potentially unchecked collection of user data through consumer mental health technologies, allowing for personal information to connect individual users to mental health concerns (eg, heightened anxiety for fear of infection), could especially violate a user’s right to privacy. For example, people could be targeted for advertisements that take advantage of their anxieties around COVID-19 by drawing behavioral inferences from their personal information [33]. At a time when the need for trust in telehealth and mental health apps has intensified, it is important to ensure accountability and mechanisms to mitigate unauthorized or unpredictable use of mental health data.

Appropriate informed consent regarding data practices are often invoked as part of ethical approaches to privacy and data protection for digital mental health in psychiatric and clinical care. It is therefore critical to note that the changes to HIPAA, in effect, mean that patients, as consumers of the technologies used for telehealth, are responsible for availing themselves of the consumer technology company’s privacy policy [34]. In the consumer domain, data protection and privacy information are generally presented through dense legal terms and conditions that are difficult for users to parse. Improvements to consent and terms and conditions regarding privacy and data use are needed. The General Data Protection Regulation in Europe and California’s Consumer Privacy Act provide some models for giving individuals clearer information regarding personal data use and options to consent to certain uses [35,36].

Yet, the HIPAA rule change raises broader questions regarding whether it is appropriate to place the burden on patients and consumers to understand and respond to how technologies are using their mental health data. Expecting patients, particularly those who may have pressing mental health needs, to understand and weigh issues, such as additional health inferences drawn...
from their data, the third parties to whom the data may be sold, and the various implications and repercussions of having their data sold, is unreasonable. In the consumer context, the lack of transparency in privacy policies has sometimes been justified by the argument that the use of these technologies is usually not necessary for the health and well-being of the user. However, this is not the case in the clinical context. If patients have no choice but to use the consumer technologies to obtain health care, then they have no choice but to acquiesce to the exploitative privacy practices of technology companies. In addition to focusing on consent and transparency as ways to address deficiencies in data protection, regulations to limit how data may be used for targeted marketing, actions by the Federal Trade Commission to address unfair or deceptive practices, and additional regulations in lieu of HIPAA may be needed to protect patients.

Data collected by digital technology and social media can be valuable for research into the mental health impact of COVID-19 and for developing evidence-based interventions [37]. HIPAA already contains provisions that allow for covered entities to convey personal health information to relevant health authorities for public health purposes. Even when such research is performed outside of HIPAA-covered entities, there is just as much of an imperative to protect private health information. Neglecting to put strong regulations in place now might risk a new normal of lowered privacy standards in the future. Although the context of a pandemic justifies sharing of confidential information for goals such as contact tracing, the practice becomes more problematic when the threat to society is less urgent. To avoid inappropriate government or commercial tracking of data, regulations and guidelines must be developed to address tracking of personal data for health research purposes, clearly identifying the types of data that may be tracked, the length of time such data may be stored, and the entities that may access this data [38].

**Safety and Accountability**

Increased access to digital mental health tools, when paired with appropriate clinical oversight, can improve and expand care, especially to underserved areas. At the same time, the large-scale reliance on digital tools during the pandemic has underscored and exacerbated the existing gaps in accountability and oversight in digital mental health. There are not yet broadly accepted ethical guidelines for the provision of digital mental health care. Digital technology can have disparate risks and benefits for research and treatment in different populations. For example, although some apps and digital tools have demonstrated benefit for people with moderate anxiety or depression, there is an urgent need to determine best practices and digital tools for providing socially distanced care for people with more severe mental illness [39]. Furthermore, the majority of consumer apps are not evidence-based and some even contain harmful content [40].

The population most likely to use digital mental health apps are people with pre-existing or newly diagnosed mental health problems such as anxiety or depression. Mental health issues add specific layers of vulnerability to a person and may compound and exacerbate coexisting psychosocial crises. If a person in severe emotional crisis, for example, due to depression, is already in a financially precarious position and then loses their job (and health insurance) because of the pandemic, they might be prone to substituting (costly) quality in-person mental health care with a cheaper, yet often unvetted, mental health app. In such a situation, the compounding effects of pathogenic (depression) and socioeconomic (joblessness) vulnerabilities might both interfere with their capacity to evaluate or use a mental health app appropriately or make an informed decision regarding the protection of their mental privacy.

In the United States, the Food and Drug Administration (FDA) relaxed regulation of mental health apps for depression, anxiety, and insomnia to facilitate expanded use of digital health tools during the pandemic [41-43]. Lowering standards may lead to efficiency in the short-term, but the widespread adoption of low-quality technology during the pandemic could lead to a long-term substandard tier of service. One specific concern is that the change in FDA oversight for digital mental health tools further opens the gate for unvetted apps and other services that are put on the market purely for profiteering off of the ongoing mental health crisis rather than providing actual relief for patients. The lack of clear regulatory oversight and guidance during the pandemic will also make it difficult to hold developers and companies liable and responsible for potential harms a user may sustain from interacting with insufficiently vetted digital mental health. Research to identify the benefits and burdens of digital mental health tools, as well as best practices in their application, is urgently needed. Therefore, we emphasize the need for the FDA or other centralized body to coordinate the systematic gathering of data on possible adverse effects of digital mental health tools. This could include establishing a secure channel for user feedback and other means of research.

With the shift to telehealth, many mental health practitioners found themselves encountering new terrain with insufficient guidance or training. Difficulties include the need for more careful safety planning for patients who are at high risk, maintaining professional boundaries in the newly informal virtual space, and designing the physical space to both frame the patient encounter and maintain work-life balance for the therapist [44]. The use of video in teleconference presents challenging questions regarding managing what a client might see of one’s home environment and vice versa. Therapists also need to consider how and when to handle situations where a client might not be alone for confidential one-on-one therapy. There are anecdotal observations supporting some advantages to telehealth for mental health care, such as clients being more likely to show up for an online therapy session [45,46]. However, it will be important to track these gains to see if they are retained after the pandemic and as patients adjust their expectations for care.

Even before the pandemic, many mental health professionals had concerns regarding their legal liability when using digital mental health tools, particularly when it came to assessment and treatment of individuals with severe mental illness or high suicide risk [47]. In the wake of the pandemic, there is added
tension for professionals as they balance the need to ensure the continued care for their clients with caution regarding additional liability from their use of digital tools. Physicians are accountable through the law of tort and under the regulatory licensing systems of the jurisdiction in which they provide care [48]. The standard of “reasonable care” creates a duty to exercise a level of care, skill, and judgment that falls within local professional norms. The relative novelty of many digital health tools means there are not yet established standards that courts can use to determine when clinicians fail to meet their duty to patients. There are unclear legal risks involved in cases where a professional relies on a mental health app in an unsuitable case nor is it established how much due diligence a professional is expected to do into the quality and functioning of mental health apps before adoption of them [49]. This is complicated by the context of a pandemic in which clinicians have had to shift quickly to digital tools without the opportunity for training or gradual implementation. Furthermore, not all malpractice insurance providers cover claims arising from telehealth services [50].

In addition to the provision of reasonable care, physicians are expected to obtain informed consent to treatment and to safeguard the confidentiality of medical information. These legal responsibilities can be challenging in relation to poorly understood apps with opaque data management and the potentially large volume of incoming information from digital mental health tools. Given the current need to use digital mental health tools to provide mental health care during the pandemic, the lacuna of norms and regulations leaves providers and patients without clear guidance regarding accountability. Local state-based licensing boards, which serve an oversight function for physician practice, will have a difficult time managing the national and global nature of digital health. Professional organizations such as the American Psychiatric Association (APA) and the American Academy of Clinical Psychology play an important role in establishing guidelines and standards of care for practitioners in these new circumstances. The development of consistent policies around the use of digital mental health tools is vital to ensuring clinicians are not discouraged from using new therapies out of fear of litigation. Even after the pandemic, we may see a rise in mental health app use if it becomes normalized during the pandemic. If insurers similarly shift to encouraging mental health apps as a replacement for in-person services due to cost, it will be vital that apps are only reimbursed if they provide effective care. Germany’s Digital Health Act, intended to accelerate the use of digital health tools during the pandemic, provides a model for navigating these concerns by requiring companies to submit evidence of safety and efficacy before they are allowed to receive reimbursement [51,52]. Similar regulation could help to provide a more consistent system for evaluation of digital health tools and ensure that users have access to safe products.

Although the pandemic has increased the number of people turning to apps to address mental health issues, there remains a need for a clear framework to allow users to understand the risks and benefits of selecting a digital mental health tool from the sea of options. The APA’s database evaluating mental health apps is a step in the right direction but is geared toward clinicians rather than consumers [53]. In contrast, Australia has a consumer-facing online government portal that provides evidence-based information regarding mental health apps developed by universities, government, and public health agencies [54]. The issue of vulnerability in relation to mental health might be a useful lens to better understand what the optimal balance could be between the benefits of continual access to therapy and the associated possible trade-offs such as involuntary exposure or violation of a patient’s privacy (eg, regarding their home environment). Guidance for therapists will need to be informed by further research and include frameworks for evaluating and handling the trade-offs that may be presented when conducting digital mental health therapy. The context of the pandemic forced much of therapy online, but a true revolution in digital mental health will require evidence-based interventions that are more than simply traditional practice in a virtual space.

### Access and Fairness

Even as the social and psychological burdens of the pandemic are expected to increase rates of mental health issues [55,56], the economic fallout of the pandemic is leading to further slashing of already underresourced mental health budgets [57,58]. The impact of COVID-19 has laid bare systemic health inequities and exerted a disproportionate impact on vulnerable populations such as older adults, racial and ethnic minorities, people living with disabilities, and people who are homeless [59,60]. Although a majority of people using telehealth have reported satisfaction [61], there are indications that telehealth has not served Black Americans or Latinx populations as well. Community mental health centers, which disproportionately serve Black and Latinx people, are much less likely to be prepared to implement digital mental health technologies [62]. Black and Latinx patients are less likely than White patients to use online health services and more likely to express reservations regarding privacy and suspicion regarding the quality of telehealth [63]. The conditions of the pandemic underscore the need to prioritize approaches that can provide improved care for groups that have repeatedly seen the least benefits from the mental health system and emerging technologies.

Since the pandemic, there has been renewed focus on how the design of health technologies may unfortunately reflect and reinforce existing biases and health disparities [64]. As with other areas of health research and technology, non-White populations are likely to be underrepresented in the data used to develop digital mental health algorithms and tools [65]. Informational and treatment apps also may not be tailored appropriately for people of different racial, linguistic, ethnic, or cultural backgrounds. Even with digital tools that are simply meant to connect patients with mental health providers, the lack of racial and cultural diversity in the pool of clinicians and therapists can create barriers to mental health care [66]. In the United States, allowing mobility of professional licensure could help create more diverse care teams and improve access to clinicians who speak the same language as patients.

Furthermore, digital mental health tools require proper training and oversight to use effectively. Insufficient resources for
adequate training of mental health professionals serving low-income demographics may mean that quality mental health care will still not be accessible despite expansion of digital mental health [67,68]. It is vital for developers, researchers, and clinicians to address potential areas of bias and plan for how to engage culturally diverse populations as well as vulnerable populations. Given that low-income groups may face worse health outcomes from contracting COVID-19 and greater economic uncertainty after the pandemic, they can least afford this result.

In the United States, where there is a patchwork of public and private health care options for mental health care, reimbursement mechanisms are a key issue for fair access. During the pandemic, reimbursement for telehealth has expanded due to social distancing requirements, but postpandemic, reimbursement for digital tools may be lower than for traditional approaches. Vulnerable populations offered mental health access during the pandemic may only have them taken away once the main crisis has ended. Inadequate evaluation mechanisms for digital health tools can also make it difficult to identify and reimburse the tools that provide quality mental health care.

As with other digital tools for health-related applications, it is important to ensure fair access, particularly to vulnerable user groups. At the same time, low-access barriers for end users often goes along with low priorities for the security and data privacy of digital tools. With the continuing crisis of mental health care provision during the pandemic, users might trade off privacy and access in ways that are not commensurate with accepted norms of human dignity and fairness. One pathway to take off pressure from individuals in dire need of digital mental health provision could be to prioritize the development and dissemination of FDA-vetted tools that conform to high standards of psychiatric care and scientific evidence. It will also be imperative to invest resources in improving the digital tools available through community mental health centers and provide for training and outreach programs for digital mental health tools in underresourced communities.

Conclusion

The COVID-19 pandemic has forced rapid adoption of digital mental health tools. However, the ethical challenges regarding privacy, fairness, transparency, and accountability remain unresolved. There is a pressing need for interdisciplinary, coordinated research efforts to understand the effects of this large-scale shift to digital mental health tools. Multidisciplinary efforts should also incorporate the input of people with lived experience of mental health issues [69]. Policy makers must assess whether to adopt new regulations to protect privacy and ensure transparency or whether modification of existing standards will be sufficient. Policy must be developed in conjunction with service users to avoid creating new inequalities in access to mental health [70]. Ultimately, the current crisis may be an opportunity to unpack the great potential for digital mental health tools to improve public health.

Acknowledgments

All authors are members of the International Neuroethics Society Emerging Issues Task Force. The views expressed in this paper are the authors’ own and do not reflect the views of the International Neuroethics Society.

Authors’ Contributions

NMM and ID both conceptualized, drafted, and edited the manuscript. AC, JC, PK, KK, AW, and LC conceptualized and edited the manuscript.

Conflicts of Interest

None declared.

References


30. Wachter S, Mittelstadt B. A right to reasonable inferences: re-thinking data protection law in the age of big data and AI. Columbia Business Law Rev 2019(2) [FREE Full text]

Abbreviations

APA: American Psychiatric Association
FDA: Food and Drug Administration
HIPAA: Health Information Portability and Accountability Act

©Nicole Martinez-Martin, Ishan Dasgupta, Adrian Carter, Jennifer A Chandler, Philipp Kellmeyer, Karola Kreitmair, Anthony Weiss, Laura Y Cabrera. Originally published in JMIR Mental Health (http://mental.jmir.org), 22.12.2020. 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 Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.
The Digital Therapeutic Alliance and Human-Computer Interaction

Simon D’Alfonso, PhD; Reeva Lederman, PhD; Sandra Bucci, ClinPsyD; Katherine Berry, PhD, ClinPsyD

1School of Computing and Information Systems, University of Melbourne, Parkville, Australia
2Division of Psychology and Mental Health, School of Health Sciences, University of Manchester, Manchester, United Kingdom

Corresponding Author:
Simon D’Alfonso, PhD
School of Computing and Information Systems
University of Melbourne
Doug McDonell Building
Parkville, 3010
Australia
Phone: 61 3 9035 5511
Email: dalfonso@unimelb.edu.au

Abstract

The therapeutic alliance (TA), the relationship that develops between a therapist and a client/patient, is a critical factor in the outcome of psychological therapy. As mental health care is increasingly adopting digital technologies and offering therapeutic interventions that may not involve human therapists, the notion of a TA in digital mental health care requires exploration. To date, there has been some incipient work on developing measures to assess the conceptualization of a digital TA for mental health apps. However, the few measures that have been proposed have more or less been derivatives of measures from psychology used to assess the TA in traditional face-to-face therapy. This conceptual paper explores one such instrument that has been proposed in the literature, the Mobile Agnew Relationship Measure, and examines it through a human-computer interaction (HCI) lens. Through this process, we show how theories from HCI can play a role in shaping or generating a more suitable, purpose-built measure of the digital therapeutic alliance (DTA), and we contribute suggestions on how HCI methods and knowledge can be used to foster the DTA in mental health apps.

(JMIR Ment Health 2020;7(12):e21895) doi:10.2196/21895

KEYWORDS

therapeutic alliance; digital mental health; affective computing; persuasive computing; positive computing; mobile phone; mHealth

Introduction

Background

The therapeutic alliance (TA), the relationship that develops between a therapist and a client/patient, is a critical factor in the outcome of psychological therapy [1,2]. As mental health care is increasingly adopting digital technologies and offering therapeutic interventions that may not involve human therapists, the notion of the TA in digital mental health care requires exploration. Although work on the TA is largely the province of clinical psychology, questions pertaining to the relationship between a human user and a therapeutic computing system presumably offer a significant opportunity for input from the field of human-computer interaction (HCI).

The term digital therapeutic alliance (DTA) is a broad one that can be applied to a range of types of digital mental health care or interventions, including computer-mediated teletherapy [3,4], web/mobile apps, and therapy agents driven by artificial intelligence [5-8]. This paper focuses on the notion of a DTA in terms of web and particularly mobile apps for mental health, which predominate the work currently carried out under the banner of digital mental health. It is also where work in HCI can be most directly applied, particularly in the case of smartphone mental health apps. Research on smartphone interfaces and the psychological aspects of interaction between a user and their smartphone as a technological object could inform the development of mental health app features that are conducive to DTA formation. Although some of the work covered in this paper could benefit digital health and behavior change technologies more generally, given the motivations behind this paper and the TA as a psychological or mental health concept, this paper focuses solely on digital mental health interventions.

Objectives

As our starting point in this discussion piece, we consider the recent efforts to develop quantitative measures of the DTA from measures of the traditional TA. Given the incipience of this
topic, little work has been conducted on devising or employing measures of the DTA. Early efforts have more or less taken an existing measure and simply modified the items such that therapist is replaced with the word app or program [9,10]. Even the so-called Working Alliance Inventory for Technology-Based Interventions essentially takes this approach [11]. Although perhaps a convenient starting point, such an approach ultimately seems unsatisfactory as it cannot account for certain nuances, particularities, and complexities that could arise in the context of digital interventions. Furthermore, although there will surely be an overlap between traditional and digital therapy, not all components of the traditional TA will necessarily apply to a DTA. There may also be dimensions of alliance in the digital contexts that are not accounted for in traditional models of the TA. To date, perhaps the most considered and detailed attempt to construct a customized measure of the DTA that does not simply mirror traditional measures comes from the study by Berry et al [12], which adapts the Agnew Relationship Measure (ARM) [13] of the traditional TA for use with mental health apps by appropriately modifying and removing items in accordance with consultations with mental health professionals and clients.

In this paper, we use this Mobile Agnew Relationship Measure (mARM) as a specific starting point by discussing its items in terms of themes or topics in HCI. Despite the positive gains made with the mARM in terms of attempting to devise a custom measure of the DTA, this attempt is solely based on applying user feedback and considerations, obtained from a clinical psychology environment, to inform modifications to an existing measure from clinical psychology. Given the significance of the interaction between humans and machines in digital mental health interventions, we show that scrutinizing the mARM items through a lens of HCI theories can provide a valuable complementary approach to considering the DTA, which could inform further work on modifying existing measures or even generating measures from scratch.

The (Digital) Therapeutic Relationship/Alliance

Conceptualizations

Work on formulating the notion of a therapeutic relationship between a client and a human therapist emerged over the course of the 20th century with the development of psychotherapeutic practice. For example, Carl Rogers, a pioneer of the humanistic approach to psychotherapy, argued that the primary task of the therapist was to embody 3 core conditions required for therapeutic change to occur: empathy, unconditional positive regard (acceptance) for the client, and congruence (what a therapist says and does matches what they think and feel) [14-16]. Although for the purpose of this paper, the terms relationship and alliance have been used more or less interchangeably, in one sense, relationship encompasses all aspects of the client-therapist relationship, whereas alliance refers to a specific aspect of the relationship by which the client and therapist hope to engage with each other to produce positive therapeutic outcomes [1]. Bordin [17] conceptualized this therapeutic or working alliance as consisting of 3 parts: (1) goals (mutual understanding of what the client hopes to achieve with therapy), (2) tasks (what the therapist and client agree needs to be done to achieve the goals), and (3) bond (the bond of trust and confidence between the client and therapist). The conceptualization by Bordin forms the basis of the Working Alliance Inventory (WAI) scale [1,18,19], the most commonly used measure of the TA in face-to-face therapy. Another conceptualization of the TA, consisting of bond, partnership, confidence, openness, and client initiative categories, forms the basis of the commonly used ARM [13].

Despite a history of research showing that the quality of the client-therapist alliance is a significant factor in the successful outcome of therapy [2], an underexamined point in determining the efficacy of digital mental health apps has been if, and to what extent, a user might develop a therapeutic connection with a mental health app. Even if such a DTA does not directly predict treatment outcomes, the formation of a DTA may support the user persisting with the app rather than prematurely discontinuing its use [20]. Research in the digital mental health field has, until recently, largely ignored the concept of a DTA when running clinical trials and developing digital mental health tools; however, given the implications of the impact of the DTA on engagement and outcomes, it is vital that researchers explore this concept in further detail. In fact, the DTA was voted as one of the top 10 research priorities in a 2018 national study in the United Kingdom involving over 600 mental health stakeholders [21].

Quantitative Measures of the DTA

As has been established, it is only recently that interest in the DTA has seen the emergence of a couple of efforts to devise measures of the DTA that go beyond simply using existing measures of the TA and modifying items such that therapist is replaced with the word app or program. In considering the DTA, Henson et al [22] took the WAI and its 3 categories of goals, tasks, and bond and informally constructed a short 6-item Digital Working Alliance Inventory, breaking each of the 3 categories into 2 app features judged essential for a client-app alliance to be formed (Table 1).

http://mental.jmir.org/2020/12/e21895/
Table 1. Items of the Digital Working Alliance Inventory.

<table>
<thead>
<tr>
<th>Number</th>
<th>Item</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>“I trust this app to guide me toward my personal goals”</td>
<td>Goals</td>
</tr>
<tr>
<td>2</td>
<td>“I believe these app tasks will help me to address my problem”</td>
<td>Tasks</td>
</tr>
<tr>
<td>3</td>
<td>“This app encourages me to accomplish tasks and make progress”</td>
<td>Bond</td>
</tr>
<tr>
<td>4</td>
<td>“I agree that the tasks within this app are important for my goals”</td>
<td>Goals</td>
</tr>
<tr>
<td>5</td>
<td>“This app is easy to use and operate”</td>
<td>Tasks</td>
</tr>
<tr>
<td>6</td>
<td>“This app supports me to overcome challenges”</td>
<td>Bond</td>
</tr>
</tbody>
</table>

To date, however, the mARM [12], an adaptation of the ARM, is to the best of our knowledge the most considered and detailed attempt to derive a custom measure specifically for digital mental health apps. The development of the measure involved 3 stages:

1. Interviews with mental health clients about the concept of TA in the context of a digital health intervention to derive key themes from interview transcripts using thematic analysis.

2. Rating scales and open-ended questions to elicit views from clients and mental health staff about the content and face validity of the original ARM scale that replaced the word therapist with the word app.

3. Findings from stages 1 and 2 used to develop the mARM, employing a decision-making algorithm about the items to be dropped, retained or adapted.

A list of the items in the mARM is provided in Table 2. It is worth noting that the following original ARM items were deemed irrelevant and removed from the mARM:

- “I am worried about embarrassing myself when using the app”
- “The app feels persuasive”
- “The app seems bored”
- “The app and I have difficulty working jointly as a partnership”

These omissions seem to make sense. Clients might be worried about embarrassing themselves with a human therapist, yet it does not seem possible to be concerned about embarrassing oneself in the eyes of an app per se. Similarly, smartphones cannot become bored and one does not enter into partnerships as such with an app. The item concerning persuasion, however, is an interesting one, since, as we will discuss below, there is a whole field that concerns itself with persuasive technology design.
Table 2. Items of the Mobile Agnew Relationship Measure.

<table>
<thead>
<tr>
<th>Number</th>
<th>Item</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>“I feel free to express the things that worry me”</td>
<td>Openness</td>
</tr>
<tr>
<td>2</td>
<td>“I feel friendly towards the app”</td>
<td>Bond</td>
</tr>
<tr>
<td>3</td>
<td>“I take the lead when using the app”</td>
<td>Client initiative</td>
</tr>
<tr>
<td>4</td>
<td>“I hold back some important things about myself from the app”</td>
<td>Openness</td>
</tr>
<tr>
<td>5</td>
<td>“I have confidence in the app and the things it suggests”</td>
<td>Confidence</td>
</tr>
<tr>
<td>6</td>
<td>“I feel optimistic about my progress”</td>
<td>Confidence</td>
</tr>
<tr>
<td>7</td>
<td>“I feel I can openly express my thoughts and feelings when using the app”</td>
<td>Openness</td>
</tr>
<tr>
<td>8</td>
<td>“I feel disappointed in the app”</td>
<td>Confidence</td>
</tr>
<tr>
<td>9</td>
<td>“I can share personal matters I am normally ashamed or afraid to reveal”</td>
<td>Openness</td>
</tr>
<tr>
<td>10</td>
<td>“I look to the app for solutions to my problem”</td>
<td>Client initiative</td>
</tr>
<tr>
<td>11</td>
<td>“I have confidence in the app and how it works”</td>
<td>Confidence</td>
</tr>
<tr>
<td>12</td>
<td>“The app accepts me no matter how I respond”</td>
<td>Bond</td>
</tr>
<tr>
<td>13</td>
<td>“The suggestions the app makes are important to me”</td>
<td>Confidence</td>
</tr>
<tr>
<td>14</td>
<td>“The app seems to understand me”</td>
<td>Bond</td>
</tr>
<tr>
<td>15</td>
<td>“The app feels warm and friendly with me”</td>
<td>Bond</td>
</tr>
<tr>
<td>16</td>
<td>“The app does not give me the help I would like”</td>
<td>Confidence</td>
</tr>
<tr>
<td>17</td>
<td>“The app is supportive”</td>
<td>Bond</td>
</tr>
<tr>
<td>18</td>
<td>“The app seems to ignore my needs”</td>
<td>Partnership</td>
</tr>
<tr>
<td>19</td>
<td>“The app confidently presents its information”</td>
<td>Confidence</td>
</tr>
<tr>
<td>20</td>
<td>“I am responsible for my recovery, not the app”</td>
<td>Client initiative</td>
</tr>
<tr>
<td>21</td>
<td>“The more I use the app, the more I get out of it”</td>
<td>Partnership</td>
</tr>
<tr>
<td>22</td>
<td>“The app gives me the confidence to take the lead in my recovery”</td>
<td>Client initiative</td>
</tr>
<tr>
<td>23</td>
<td>“I agree with the direction the app is taking me”</td>
<td>Partnership</td>
</tr>
<tr>
<td>24</td>
<td>“The app is like having a member of my care team in my pocket”</td>
<td>A novel addition (not resulting from retention or adaption) to capture a key theme from interviews</td>
</tr>
<tr>
<td>25</td>
<td>“I am clear about what the app can and cannot offer me”</td>
<td>Not categorized in the Agnew Relationship Measure</td>
</tr>
</tbody>
</table>

Applications of HCI

The field of HCI concerns the design of computer technology and how humans interact with such technology, particularly how it can be best designed to facilitate its use [23]. When considering digital mental health apps and the role HCI might play in DTA formation, it is not necessarily about ways in which the smartphone or computer can be anthropomorphized. Rather, we are interested in the smartphone or computer as a device per se and the ways in which apps can be given features, including those that make use of certain capacities, particularly in the case of smartphones, to foster the DTA. For example, a theme that will be considered as this paper unfolds is how the power afforded by modern smartphones to infer user behavior and context [24,25] might offer new opportunities to personalize content and tailor responses to support fostering of a DTA. We will now take a look at several areas of HCI germane to the DTA, before an extended discussion of how these areas can apply to questions relevant to the DTA via mARM items.

Persuasive System Design

A persuasive computing technology is “a computing system, device, or application intentionally designed to change a person’s attitudes or behaviour in a predetermined way” [26]. Fogg coined the term captology from the phrase Computers as Persuasive Technologies [26-28] to reflect this idea. As we will now briefly elucidate, persuasive design principles are relevant to several DTA criteria, not just the direct matter of whether the app feels persuasive.

Informed by Fogg’s conceptualization of persuasive technology Oinas-Kukkonen and Harjumaa [29] have developed a concrete framework that transforms persuasive design principles into software requirements and system features. According to their persuasive systems design (PSD) model, there are 4 categories for persuasive system design, each consisting of several principles:

1. Primary task support: the design principles in this category support the execution of the user’s primary task and consist...
of reduction, tunneling, tailoring, personalization, self-monitoring, simulation, and rehearsal.

2. Dialogue support: the design principles in this category are about the feedback an interactive system provides to its users to help them move toward their goal or a target behavior. This category consists of praise, rewards, reminders, suggestion, similarity, liking, and social roles.

3. System credibility support: the design principles in this category describe how to design a system so that it is more credible and thus more persuasive. The category consists of trustworthiness, expertise, surface credibility, real-world feel, authority, third-party endorsements, and verifiability.

4. Social support: the design principles in this category describe how to design the system so that it motivates users by leveraging social influence. The category consists of social facilitation, social comparison, normative influence, social learning, cooperation, competition, and recognition.

Of these 4 categories, the fourth one is not fully relevant, as this investigation focuses on the connection between an individual user and an app. If a second human were involved, it would generally be a therapist accompanying the app user. However, there are cases of digital mental health interventions involving dedicated social components [30], particularly social networking, in which case social support becomes a significant factor. Several of the mARM items, such as “I feel free to express the things that worry me” and “The more I use the app, the more I get out of it” would vary in meaning given a social (networking) component. We will also be touching upon social relatedness as a psychological principle further on.

Of the first 3 pertinent categories, a selection of principles is particularly relevant when considering the DTA. From Primary task support, the principles of personalization and tailoring require that systems provide personalized content and services and tailored information to users and user groups. Not surprisingly, “personalising tasks or goals to the individual is likely to support the formation of a relationship with the technology” [16], and in subsequent sections, we will discuss how approaches in HCI can be used to foster these aspects of the DTA and items such as “I have confidence in the app and the things it suggests” and “the suggestions the app makes are important to me.”

Several of the dialogue support principles are particularly relevant. Praise (offering praise), rewards (rewarding target behaviors), and reminders (reminding users of their target behavior) are principles whose implementation would support the mARM item the app is supportive. The principle of similarity, which says that “people are more readily persuaded through systems that remind them of themselves in some meaningful way” [29] is not just conducive to system persuasiveness but can also make a contribution to the DTA by supporting items such as the app seems to understand me. Finally, several of the system credibility support principles are also particularly relevant. Trustworthiness, expertise (system should provide information showing knowledge, experience, and competence), and surface credibility (system should have competent look and feel) clearly connect with DTA criteria such as “I have confidence in the app and how it works” and “I have confidence in the app and the things it suggests.”

In light of this discussion on persuasive system design, the choice to remove the item the app feels persuasive from the mARM because of “low relevancy” and “no alternative options were suggested or agreed upon” [12] is seriously brought into question. Indeed, the issue with this item is a prime example of how investigating the DTA from an HCI perspective can help to shape its conceptualization and measurement.

Affective Computing

Affective computing is a subfield of HCI that concerns systems and devices that can recognize, interpret, process, and simulate human affects/emotions [31,32]. Advances in smartphones have paved the way for rich opportunities in using data acquired from embedded smartphone sensors and smartphone use to infer a user's affective states [33]. Regarding the DTA, the capacities of affective computing can work in 2 ways.

The first method is to detect a user’s state and tailoring components of the app, such as therapy recommendations and screen messages or information accordingly. States or difficulties such as low mood, anxiety, and stress can be inferred with a variety of technological modalities, including phone interactions, movement sensors, facial analysis, voice analysis, and text analysis [34,35]. Whether it is in immediate response to a momentary signal given off by an individual or from behavior inferred over a longer time period such as a day or a week, an app can use such information to deliver a momentary interventional exercise suggestion, strategy, or message of encouragement [36,37]. For example, the detection of relatively high levels of negative emotional states such as anxiety or stress on a given day, using computational linguistic and acoustic analysis of an individual’s textual and vocal smartphone communications [38,39], could trigger an evening push notification with stress or anxiety management exercises.

The second way concerns simulating affect, in particular via screen content and messages, in such a way that they are appropriate or attempt to induce the right affective state in a user. In general terms, beyond quality content, the form in which the content is delivered is also a factor. Particularly in the cases of virtual agents and bots, there is a reasonable assumption that users would prefer an agent that exhibits in greater quantities the emotional intelligence and general anthropomorphic characteristics of a human therapist. This generalization is challenged, however, by the uncanny valley phenomenon [40], whereby people develop a sense of unease and discomfort at robots that fall within a certain range of human likeness, neither too human-like or human-like enough. Furthermore, some research suggests that it is not unequivocal whether users prefer an emotionally demonstrative system over one that is affectively neutral and that this might depend on the personality type of the user [41].

Despite smartphones and personal computers not having an anthropomorphic form, apps on such devices can still incorporate affective qualities. For example, the type of language that an app uses, and for those apps that include voice, the paralinguistic properties of that voice would influence responses to the mARM such as “I feel friendly towards the app” and “The app feels warm and friendly with me.”
Eudemonic Psychology and Positive Computing

The notion of human psychological well-being is accompanied by a variety of definitions and approaches to measurement. In the tradition of ethical hedonism [42], hedonic approaches to psychological well-being define it broadly as the experience of positive affect. Termed subjective well-being, measures based on this approach generally consist “of three components: life satisfaction, the presence of positive mood, and the absence of negative mood, together often summarised as happiness” [43]. Positive emotions and pleasure seeking are undoubtedly important elements of the human condition, but beyond this sense of well-being lies one with ties to the Aristotelian notion of eudemonia, a notion of well-being that goes “beyond the experience of positive emotion into the realms of engagement, meaning, relationships, and human potential” [44].

This broader sense of well-being, termed eudemonic or psychological well-being, encompasses aspects of positive human functioning and flourishing, such as purposeful engagement in life, realization of personal talents and capacities, and enlightened self-knowledge. [45] aspects neglected by accounts that narrowly focus on satisfaction, feeling good, and contentment.

There are several prominent accounts and frameworks based on this conception of well-being. The Self-Determination Theory (SDT) by Ryan and Deci [43,46] posits 3 basic elements that typically foster subjective as well as eudemonic well-being:

1. Autonomy: feeling agency and acting in accordance with one’s goals and values
2. Competence: feeling able and effective
3. Relatedness: feeling connected to others and a sense of belonging.

Similarly, the framework for eudemonic well-being by Ryff and Singer [47] is concerned with 6 core components: self-acceptance, autonomy, personal growth, positive relationships, environmental mastery, and purpose in life.

The positive psychology movement perhaps most conspicuously embodies the ethos of eudemonic or psychological well-being and the promotion of positive function and flourishing [48]. At the base of positive psychology is the PERMA model, which stands for positive emotions, engagement, relationships, meaning, and achievement. Positive psychology also identifies the importance of using “signature strengths every day to produce authentic happiness and abundant gratification” [49], strengths such as connectedness, gratitude, kindness, open-mindedness, perseverance, honesty, and courage [50].

The incorporation of this conception of well-being into the design and development of computing and information systems is embodied in the emerging field of positive computing, which addresses how technology can “support wellbeing that encompasses more than just immediate hedonic experience, but also its longer-term eudaimonia, or true flourishing” [51]. This is achieved through the integration of well-being theories and techniques from frameworks such as SDT and positive psychology into such technologies.

For example, the autonomy component of SDT can be supported by offering options and choices over use and not in turn demanding actions from users without their assent [51]. The component of competence can be enhanced by including optimal challenges that are neither too difficult nor too easy, positive feedback, and opportunities for learning [51]. Finally, an aim to foster relatedness can determine approaches taken in the development of digital systems for social connection. For example, direct communication such as wall posts, comments, and web chat is associated with greater relatedness over mere passive consumption of friends’ content [51]. Research [52] suggests that users develop a quality relationship or bond with health apps that are sensitive to their needs for autonomy and relatedness. Furthermore, listed below are the 5 identified dimensions of autonomy [53] “that are useful for understanding the mediating role that health and wellbeing apps have on the communication of information” [54]:

1. Degree of control and involvement that the user has within the app
2. Degree of personalization over the app’s functionality
3. Degree of truthfulness and reliability related to the information presented to the user and how this affects their decisions
4. User’s self-understanding of the goal pursuit and whether the app promotes or hinders a user’s awareness of their own agency
5. Whether the app promotes some form of moral deliberation or moral values in the actions it recommends.

The implementation of features conducive to the strengths of positive psychology is another example of positive computing. For example, designers might add a thanks button based on the evidence that expressing gratitude promotes overall well-being [55]. Furthermore, apps and software built from scratch to promote well-being, particularly digital mental health interventions, can be exemplars of positive computing. For example, the moderated online social therapy (MOST) mental health platform has been built on a basis significantly influenced by positive psychology [56]. Previous work on MOST suggests that platform design informed by the principles of SDT supports the emergence of a DTA between users of a digital mental health platform and the platform itself [57].

In our subsequent discussion, we will point out where approaches to well-being, such as SDT and positive psychology, and hence their HCI embodiment in positive computing, can promote certain DTA items as given in the mARM.

The Human-Smartphone Connection

The relationships we have with our technological devices such as the smartphone, whether positive or negative, is another relevant field of inquiry, which will primarily relate to the bond category of the TA. Smartphone attachment theories may play a role in our understanding of the extent to which humans can develop relationships, such as a DTA, with digital mental health apps. In developing smartphone mental health apps, “it is important to consider that the quality of an individual’s relationship to his/her mobile phone may influence their receptivity to, and ultimately the efficacy of, mobile health (mHealth) programs and interventions” [58].
Problematic mobile phone use and smartphone addiction are phenomena that are gaining some diagnostic currency [58,59]. There is even a purported phenomenon that goes by the neologism nomophobia, a portmanteau derived from NO MOBILE PHone PhoBIA, which also has a questionnaire to quantitatively measure it [60]. Smartphone addiction scale items such as “having my smartphone in my mind even when I am not using it” [59] are indicators of a negative relationship. However, they suggest the possibility of a deep bond between the user and phone, which, if combined with positive goal development, could be harnessed for beneficial, therapeutic ends.

The study by Ribak [61] has described how mobile phones can act as transitional objects for adolescents. Furthermore, Vincent [62] explores and examines the concept of emotional attachment to mobile phones, and Melumad and Pham [63] show that smartphones can serve as attachment objects for consumers:

Results from two experiments show that smartphones provide greater comfort and faster recovery from stress (vs. PCs), defining characteristics of attachment objects. A third study shows that smartphone use becomes pronounced among consumers particularly susceptible to stress – those who recently quit smoking.

As discussed in the study by Li et al [52], there is a tendency for emotional bonding and attachment behaviors toward a health app to occur when the app user is experiencing something negative and the app attends to their basic needs, such as providing help with ill-health. Such emotional bonding is conceptualized as an affectionate response when individuals use health apps, which manifests in 3 aspects: “warm feelings when using mHealth apps; they become aroused with intense and positive moods about mHealth apps; and they sense close connections with mHealth apps” [52]. This connection between a user and their smartphone would pertain to bond mARM items such as “I feel friendly towards the app.” Furthermore, it would also seem to pertain to the novel item “The app is like having a member of my care team in my pocket.”

Before moving on to the next section, one final point to gather from the discussion of the 4 HCI areas in this section is that the psychology or personality of the individual user is likely to play a role in which apps or app features work for them in terms of DTA formation and app adoption more generally. This indicates an advantage for recruitment systems whereby a preunderstanding of the user can be used to establish the suitability of an app for them. Beyond this possibility, which will most often not necessarily be the case, data-based profiling techniques embedded into app technologies provide another way to learn about the user on the fly and establish their relevant personality characteristics.

Assessing mARM Items in Light of HCI

We now examine each of the 25 mARM items listed below and discuss, where applicable, what the HCI topics of PSD, affective computing, positive computing, and the human-smartphone connection have to say about them, thus facilitating an exploratory discussion of the DTA and HCI via the structure of the mARM. A rough classification of 3 item types emerged, consisting of the following:

1. Items which can be supported by HCI theories.
2. Items for which HCI considerations are not directly relevant or which are not linked to specific HCI topics, as they are more so questions gauging the characteristics of the app user.
3. Items whose inclusion or exclusion in a DTA measure is brought into question in light of HCI considerations.

1. I Feel Free to Express the Things That Worry Me

An app to which this item applied would by definition be one in which users can express their worries. For example, the app might contain a simple journaling feature or questions in an exercise designed to elicit responses from the user expressing their worries. Another more involved possibility is the interaction with a conversational feature or agent in the app. In such cases, the suitability of the journaling medium or questions asked would influence the quality of what the user expresses. Trust and what the app does with what the user shares could also influence their expression, including data privacy and security, and whether the responses will be seen by another human. Whatever the case, however, it does seem that a core part of this question does not apply in the case of an app, namely, the freedom or inhibition a client may feel in expressing their worries depending upon the relationship they have with their human therapist. A related, perhaps more apt question could be “I find it beneficial expressing the things that worry me.”

Although the PSD category of social support is not a focus of individual user apps, this item would come to have another significance were the app to have a function through which worries could be expressed with peers and/or clinicians.

2. I Feel Friendly Towards the App

This item involves and is an opportunity to emphasize an important point that pertains to several of the mARM items and the DTA in general. Despite interacting with a nonhuman, nonconscious agent, people demonstrate a willingness to form human-like relationships with technology. As early as the 1960s, the tendency to anthropomorphize computers, ascribing to them human traits and intentions that they do not actually have, was observed with ELIZA, an early natural language processing computer program that simulated a Rogerian psychotherapist (dubbed the ELIZA effect) [64]. Furthermore, research [65] indicates that “the sophisticated interactions people have with computers engage many of the same cognitive schema and patterns of behaviour found in human social interactions” [16].

The fact that “people reciprocate positive behaviours from computers by behaving similarly in return” [16] suggests that the incorporation of certain affective computing characteristics or qualities in an app would engender a system that supports this item by providing friendly cues and language. However, what in fact this item would be measuring is brought into question by research suggesting that such interactions are mindless, that is, people are simply mindlessly following triggered social scripts and responding to computers as social

http://mental.jmir.org/2020/12/e21895/
actors via a relatively automatic process beyond their awareness, rather than a conscious choice [16,66].

3. I Take the Lead When Using the App

Although this item is largely dependent on the nature of the user, the facilitation of autonomy with supportive, positive computing design and features would be conducive to this item. Facilitating concordance, where an individual can modify an intervention to suit the way they prefer to use it, rather than just adherence, where the system might prescribe a strict therapy pathway the user should stick to, would also give users more opportunity to take the lead [57]. Of the 5 dimensions of app autonomy listed earlier in the section on Eudemonic Psychology and Positive Computing, dimensions 1 (degree of control and involvement that the user has within the app) and 2 (degree of personalization over the app’s functionality) would also contribute to this item.

4. I Hold Back Some Important Things About Myself From the App

Responses to this item would largely be determined by an individual’s psychology and attitude toward app therapy. However, the extent to which an app demonstrates PSD principles such as trustworthiness and expertise will perhaps influence how many important things about themselves the user is willing to share.

Inducing users to share things with PSD incentives such as personalization (eg, the more you share with the app, the better tailored the app will be for you) and self-monitoring (eg, sharing with the app will provide you with monitoring snapshots about yourself) are other options to promote this item. Further options include employing praise or rewarding users when they share things. However, these latter possibilities, in particular, raise consideration of the spectrum between intrinsic motivations and extrinsic motivations; there is a qualitative difference between a user sharing important things about themselves because it has intrinsic therapeutic value for them versus sharing things because they receive some extrinsic reward in doing so. However, gaining useful information from and about the user so that an app may better serve them, even if the information is obtained via extrinsic reward incentives, is generally better than nothing.

5. I Have Confidence in the App and the Things It Suggests

Trustworthiness, expertise, and surface credibility (system should have competent look and feel), design principles in the PSD category of system credibility, would naturally support this DTA item. More interesting are apps that aim to deliver accurate and relevant personalized therapy suggestions for the user. The PSD principle of personalization can be defined as “the ability to provide contents and services tailored to individuals based on knowledge about their needs, expectations, preferences, constraints, and behaviours” [67]. The ubiquity of digital technologies that are equipped with sensors for inferring user behaviors, situations, and contexts, combined with advances in data processing and science, has augmented the possibilities of personalized recommendations and personalized HCI more generally [37,68].

If an app does deliver personalized therapy suggestions based on a user’s app use history or their smartphone sensor information, then an explanation of why the suggestion was made would presumably help to promote a sense of confidence [69]:

Recommendation systems and explainability are fascinating and complex topics. Some basic example forms of explained recommendation will suffice to convey this idea, which will already be familiar to those who use websites that deliver content such as Netflix and Amazon:

- Therapy item X was suggested because you have recently completed therapy item Y (with X and Y having a predefined relevance connection).
- Therapy item X was suggested because you told us fact Y about yourself.
- Therapy item X was suggested because users similar to you have benefited from it.

Furthermore, the application of smartphone sensing for contextual awareness and personal sensing insights [35] raises a range of rich recommendation possibilities and explainability challenges. For example, suppose that an individual who generally goes to bed before midnight during weeknights is up at 3 AM using social media on their smartphone for a third consecutive weeknight. This fact, coupled with other recent smartphone use patterns such as keystroke dynamics, might be indicative of stress-related insomnia and could be an opportunity for their mental health app to offer push notifications for a real-time therapy exercise to help with their condition. If so, an appropriately worded explanation for such a recommendation would also need to be considered.

These points serve to make the case that an app, powered by smartphone technology and algorithmic intelligence, would inspire confidence by successfully generating accurate personalized suggestions that resonate with the user and, furthermore, by accompanying them with a good explanation (especially for more sophisticated suggestions).

6. I Feel Optimistic About My Progress

In terms of PSD, the dialogue support principles of praise (offering praise) and rewards (rewarding target behaviors) would foster this item.

In addition, unless the mental health app involves social networking or human moderation, the relatedness component of SDT has no direct import. However, certain indirect features could be incorporated into an app to encourage use and engagement with therapy. One can imagine a feature that enables certain successes or milestones within the app to be shared with
an individual’s social network channels. Such a feature would support this item.

7. I Feel I Can Openly Express My Thoughts and Feelings When Using the App

Most of the points made above for item 1 apply to this item. Whatever means an app has for users to express their thoughts and feelings, it should be easy to do so, and the user must be confident that the content they share will be used appropriately. Similar to item 4, the informational value an app offers the user in response to sharing their thoughts and feelings can also be a fundamental incentive. For example, in the theme of self-tracking with technology and the PSD principle of self-monitoring, “app-based features that enable users to self-monitor their mood by periodically reporting their thoughts, behaviours, and actions can increase emotional self-awareness (ESA)” [70]. This ability to identify and understand one’s own emotions “has been shown to reduce symptoms of mental illness and improve coping skills” [70].

8. I Feel Disappointed in the App

A variety of matters, including HCI ones and those involving the theories introduced earlier, can influence responses to this item: usability of the app, quality of the content, nature of the individual using the app, and accuracy and reliability of the app. Another important aspect of this item can be considered a form of congruence for computing systems: how the app presents itself and the experience it delivers should be consistent and match the expectations and relationship the user forms with the app. For example, suppose that an app has an onboarding process asking the user for personal information, with messages that this information will be used to provide the user with relevant information and tailored help throughout their journey with the app. This will set up an expectation in the user that the app will do what it signals it will do, such that if the app fails to deliver on its promise and does not satisfy the user with relevant information or provides egregiously irrelevant personalized suggestions, it is likely to disappoint the user.

9. I Can Share Personal Matters I Am Normally Ashamed or Afraid to Reveal

This item obviously shares similarities and overlaps with items 1 and 7. In the case of traditional therapy, the score received for this item will be a function of the relationship developed between the client and therapist. In the case of an app, certain app features might foster this item; however, the score will largely be a function of the user’s attitude toward using mental health apps. Another factor positively contributing to this item is the fact that apps offer users, particularly those concerned about such a thing, a means to share personal matters without (directly) communicating with another person and feeling stigma or concern that what they share will be scrutinized. This aspect of computers has often been mentioned as one of their advantages as a mental health care solution [71,72].

10. I Look to the App for Solutions to My Problems

The score an app receives for this item will be a function of both the quality of the app and the user’s willingness to use it as a solution in their mental health care.

11. I Have Confidence in the App and How It Works

Trustworthiness, expertise, and surface credibility (system should have competent look and feel), design principles in the PSD category of system credibility, naturally support this DTA item. It should be noted that this mARM item seems related to item 5, and it is worth considering how the distinction between the 2 analogous original ARM items might be affected when translated into mARM form (the item 5 ARM equivalent is “I have confidence in the therapist and their techniques” and the item 11 ARM equivalent is “The therapist’s skills are impressive”). The ARM versions seem sufficiently independent, whereas with the mARM versions, the response to item 5 seems quite constitutive of the response to item 11; an app that does not offer good suggestions is in one important or even crucial sense not working well, despite the fact that the app may be technically impressive. If these 2 items are to remain distinctive in a measure of the DTA, then the item replacing the ARM notion of item 11 could perhaps be something like the app is technically impressive.

12. The App Accepts Me No Matter How I Respond

This item ties in with the Rogerian notion of unconditional positive regard. The first question for this item concerns the notion of acceptance; it is questionable if a computing device can accept the responses of a user in the same way as intended in the original ARM question, though they can provide responses indicating some form of programmed acceptance. However, an app should not provide blanket responses of acceptance to any user response. Although an app should not reject a user or provide responses containing unnecessary negativity, “negative or directive feedback provides guidance, leading people to become, over time, more certain about their behaviour and more confident in their competence” [73].

13. The Suggestions the App Makes Are Important to Me

This item shares some overlap with item 5, and it stands to reason that the PSD principle of personalization of app content delivery and the quality of that content will increase the chances of better scores for this item. Findings of the study by Duggan et al [74] suggest that personalizing tasks or goals to the individual is likely to support the formation of a relationship.

14. The App Seems to Understand Me

As with item 13, personalization of app content delivery will increase the chances of better scores for this item. In fact, one qualitative analysis identified that automated personalization helped one user feel understood and that intelligent responses from an app fostered the perception of a relationship for another user [75]. There is also a connection between personalization and user autonomy (first principle of SDT), as “personalization also creates a sense of ownership and choice beneficial to autonomy” [51]. However, there are 2 types of personalization that require distinction. The first type provides users with the ability to customize their experience of the system by giving them access to edit certain settings, in line with app autonomy dimension 2 introduced earlier. The second type is that of automated personalized content recommendation systems. The former of these 2 types is relatively straightforward; however,
the latter raises consideration of possible tensions between user autonomy and system automation.

Although most flagrantly problematic in the case of big commercial platforms such as YouTube and Facebook, whose recommendation systems, newsfeed, and advertising are fraught with consequences of political, social, and epistemological detriment, the design of recommendation systems for automated intervention suggestions in health apps, particularly mental health apps, warrants consideration. Beyond the issues of ensuring safe, accurate recommendations possibly accompanied by explanation, recommendation systems can negatively encroach on a users autonomy by nudging them in a particular direction or limiting the range of options which could be presented to them, even possibly in extreme cases addicting them to certain content or actions [76]. If autonomy and self-directedness are conducive to TA formation and positive therapeutic change, we need to create systems that strike a balance between providing personalized automation and facilitating the client decision making necessary for effective therapy. The app autonomy dimension 5 introduced in the section Eudemonic Psychology and Positive Computing (whether the app promotes some form of moral deliberation or moral values in the actions it recommends) is pertinent in these considerations.

15. The App Feels Warm and Friendly With Me
How can an app possibly be made to feel warm and friendly? One possibility is for the app to use language that is warm and friendly. However, perhaps more significant would be for the app to have an affective computing ability to detect a users affective state and tailor its responses accordingly. As has been noted by others, the Rogerian notion of congruence “is a particular challenge for technological interventions where it is trivial to programme expressions of empathy or positive regard, but not easy to imbue these expressions with genuineness or authenticity” [16].

A smartphone does not have intentionality, and any signaling of empathy is not an embodiment of some consciousness correlate. However, such signals can be generated in such a way that users, in the spirit of the ELIZA effect, treat them as though they do have a degree of genuineness or authenticity. This is by programming the device so that it clearly acts in a way that is sensitive to its environment and its user. An app that exhibits artificial emotional intelligence by responding in such a way will possibly give users a sense of empathy via a simulative effect.

16. The App Does Not Give Me the Help I Would Like
General failures of HCI principles, poor information quality, and unsuitable therapy content can all contribute to higher scores for this item. Responses to this item would also largely be a function of the users needs and the psychological content of the app.

17. The App Is Supportive
Praise (offering praise), rewards (rewarding target behaviors), and reminders (reminding users of their target behavior) are PSD dialogue support principles whose implementation would support this item.

18. The App Seems to Ignore My Needs
An app being affectively and effectively responsive will also decrease the chances of dissatisfaction with the app and correspondingly minimize values for this item. When users provide input indicating the need for help, an app needs to readily respond with information or therapy content relevant to the needs of the user.

19. The App Confidently Presents Its Information
This item overlaps with item 5. Apps can easily present information, but how can they present information in a manner that a user perceives as confident? Generally speaking, an app that promotes this item should incorporate PSD system credibility support principles such as expertise, surface credibility, authority and verifiability. Signalizing to users, where appropriate, that app content being delivered is evidence based is a relatively easy way to help achieve this. In terms of personalized therapy suggestions that an app may present, explanation would signal that the app is confident in the information/therapy suggestions that it is presenting.

20. I Am Responsible for My Recovery, Not the App
Although this is an item whose score would largely be a function of the users nature, the facilitation of user autonomy would also be conducive to this item. The app autonomy dimension 4, listed earlier in the Eudemonic Psychology and Positive Computing section, pertains to this item.

21. The More I Use the App, the More I Get Out of It
This item departs significantly from the original ARM item from which it was derived, which says that the therapist and client are willing to work hard. The resulting mARM item was due to a rewording following the suggestion of one participant in the research survey, given that the original version was deemed of low relevance.

The original item comes under the partnership component of the ARM, and although this modified mARM version relates to the connection between a user and an app, it does not seem to necessarily be an indicator of beneficial app use. It may very well be the case that a moderate amount of app use benefits a user, but that anything beyond this does not provide any additional therapeutic benefit; in some cases, more use could be detrimental. Therefore, even if the score for this item is not high, this does not necessarily mean that the user did not get a lot out of their connection with the app or that using the app did not result in significant, positive therapeutic change.

22. The App Gives Me the Confidence to Take the Lead in My Recovery
Competence, which is the second component of SDT and as discussed earlier can be fostered by positive computing factors, relates to this item. In this way, apps should include optimal challenges, positive feedback, and learning opportunities.
23. I Agree With the Direction the App Is Taking Me
Failures of HCI principles could contribute to lower scores for this item, for example, if an app fails to accurately tailor itself to the user or if an app operates in a way that conflicts with the user’s sense of autonomy or competence. However, responses to this item would also largely be a function of the user’s needs and the psychological content of the app.

24. The App Is Like Having a Member of My Care Team in My Pocket (A Novel Addition [Not Resulting From Retention or Adaption] to Capture a Key Theme From Interviews)
Two factors that will determine the score for this item are the extent to which the individual is in possession of their phone and the quality of care provided by the app. An individual being in possession of their phone for a large majority of the day and they deeming the app to be useful and supportive are likely to help the score for this item. The connection a user has with their phone, including its status as a potential attachment object, could also influence this item.

25. I Am Clear About What the App Can and Cannot Offer Me
The original ARM version of this item was about the therapist and client being clear about their roles and responsibilities in their interaction. Once again, this mARM version departs a fair bit from what the original item was measuring. If a user is clear about the app, this is likely to support other mARM items. For example, it is likely to reduce the chances of the user being disappointed in the app.

However, perhaps in departing from the original, this item unnecessarily removes elements that could be captured. It is certainly possible to ask whether a user is clear about their responsibilities in using and interacting with the app. On the other side, although an app cannot understand its roles and responsibilities, it can be judged on whether it was designed and functions appropriately to fulfill the roles and responsibilities it should have.

Conclusions
The growing presence of mental health apps and digital mental health interventions in general calls for research into the notion of a DTA. The significance of the TA in traditional mental health therapy suggests the importance of considering its translation in the digital context. At this early stage, the TA does not seem to be associated with therapeutic outcomes in digital interventions as it is with human-human therapy. It may, however, be conducive to increased engagement with and adherence to digital interventions, through which it could influence outcomes. However, when making such assessments at this stage, we must be mindful of the fact that the few existing studies on measuring the DTA have either been simply copied with minimal rewording or modestly adapted existing measures of the traditional TA. Irrespective of whether such attempts demonstrate some association between their alliance measures and app outcomes, a true conceptualization of the DTA may very well require its own type of measure that fundamentally differs in certain ways from the traditional TA, a radically customized or novel measure that better fits the contours of human engagement with computers and digital interventions. Thus, although work on the DTA has been largely confined to digital mental health within the field of clinical psychology, the field of HCI can, as we have shown in this paper, profitably play a part. First, as by definition, HCI studies the interaction between humans and computers that is central to the DTA, it can be applied to determine inadequacies in simply translating traditional measures of the TA and can help to shape a novel conceptualization of the DTA. Second, tools and techniques from HCI can be employed in app development to foster items in a suitable measure of the DTA.

Acknowledgments
This work was supported by The University of Melbourne and The University of Manchester Research Fund initiative.

Conflicts of Interest
Professor Bucci is Director of a not-for-profit Community Interest Company, Affigo.io, designed to make apps available in the NHS.

References


70. Chandrashekar P. Do mental health mobile apps work: evidence and recommendations for designing high-efficacy mental health mobile apps. Mhealth 2018;4:6 [FREE Full text] [doi: 10.21037/mhealth.2018.03.02] [Medline: 29682510]
75. Hillier L. Exploring the Nature of the Therapeutic Alliance in Technology-based Interventions for Mental Health Problems. Lancaster: Lancaster University; 2018.

Abbreviations
ARM: Agnew Relationship Measure
DTA: digital therapeutic alliance
HCI: human-computer interaction
mARM: Mobile Agnew Relationship Measure
mHealth: mobile health
MOST: moderated online social therapy
PSD: persuasive systems design
SDT: Self-Determination Theory
TA: therapeutic alliance
WAI: Working Alliance Inventory

©Simon D’Alfonso, Reeva Lederman, Sandra Bucci, Katherine Berry. Originally published in JMIR Mental Health (http://mental.jmir.org), 29.12.2020. 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 form, provided the original work is properly cited.
Repeated Digitized Assessment of Risk and Symptom Profiles During Inpatient Treatment of Affective Disorder: Observational Study

Maike Frederike Richter¹, MSc; Michael Storck², PhD; Rogério Blitz², PhD; Janik Goltermann¹, MSc; Juliana Seipp¹, BSc; Udo Dannlowski¹, PhD; Bernhard T Baune¹,³,⁴, FRANZCP, MPH, MD, PhD; Martin Dugas², PhD; Nils Opel¹,⁵, MD

¹Department of Psychiatry, University of Münster, Münster, Germany
²Institute of Medical Informatics, University of Münster, Münster, Germany
³Department of Psychiatry, Melbourne Medical School, The University of Melbourne, Melbourne, Australia
⁴The Florey Institute of Neuroscience and Mental Health, The University of Melbourne Parkville, Melbourne, Australia
⁵Interdisciplinary Centre for Clinical Research Münster, University of Münster, Münster, Germany

Corresponding Author:
Nils Opel, MD
Department of Psychiatry
University of Münster
Albert-Schweitzer-Str. 11
Münster, 48149
Germany
Phone: 49 2518358160
Email: n_opel01@uni-muenster.de

Abstract

Background: Predictive models have revealed promising results for the individual prognosis of treatment response and relapse risk as well as for differential diagnosis in affective disorders. Yet, in order to translate personalized predictive modeling from research contexts to psychiatric clinical routine, standardized collection of information of sufficient detail and temporal resolution in day-to-day clinical care is needed. Digital collection of self-report measures by patients is a time- and cost-efficient approach to gain such data throughout treatment.

Objective: The objective of this study was to investigate whether patients with severe affective disorders were willing and able to participate in such efforts, whether the feasibility of such systems might vary depending on individual patient characteristics, and if digitally acquired assessments were of sufficient diagnostic validity.

Methods: We implemented a system for longitudinal digital collection of risk and symptom profiles based on repeated self-reports via tablet computers throughout inpatient treatment of affective disorders at the Department of Psychiatry at the University of Münster. Tablet-handling competency and the speed of data entry were assessed. Depression severity was additionally assessed by a clinical interviewer at baseline and before discharge.

Results: Of 364 affective disorder patients who were approached, 242 (66.5%) participated in the study; 88.8% of participants (215/242) were diagnosed with major depressive disorder, and 27 (11.2%) had bipolar disorder. During the duration of inpatient treatment, 79% of expected assessments were completed, with an average of 4 completed assessments per participant; 4 participants (4/242, 1.6%) dropped out of the study prematurely. During data entry, 89.3% of participants (216/242) did not require additional support. Needing support with tablet handling and slower data entry pace were predicted by older age, whereas depression severity at baseline did not influence these measures. Patient self-reporting of depression severity showed high agreement with standardized external assessments by a clinical interviewer.

Conclusions: Our results indicate that digital collection of self-report measures is a feasible, accessible, and valid method for longitudinal data collection in psychiatric routine, which will eventually facilitate the identification of individual risk and resilience factors for affective disorders and pave the way toward personalized psychiatric care.

(JMIR Ment Health 2020;7(12):e24066) doi:10.2196/24066
KEYWORDS
affective disorders; digital data collection; psychiatry; P4 medicine

Introduction

In what has become known as P4 or precision medicine [1,2] a major goal of medical research and applied health care is the evolution from a reactive treatment approach toward medical care that is predictive, preventative, personalized, and participatory.

This approach is particularly relevant in the field of psychiatry, as the imprecise nature of psychiatric nosology, in part due to the heterogeneity of clinical populations, complicates the identification of vulnerable groups and effective treatments [3,4]. Affective disorders such as major depressive disorder exemplify this problem. Only approximately one-third of patients with moderate to severe depression respond to the first treatment attempt with medication [5,6]. This leads to a prolonged illness duration for nonresponders, which is associated with worse overall health outcomes and significantly higher costs to the health care system [7,8]. Precision psychiatry could help alleviate this problem by predicting (P1) the occurrence of depression as well as individual disease course and preventing (P2) unfavorable outcomes such as chronification and suicide by personalizing (P3) treatment plans according to individual risk and resilience factors.

Preliminary attempts have been made to achieve the prediction of disease course and treatment outcome of major depressive disorder through the use of predictive modeling approaches [9-13]. These are a first step toward the identification of biomarkers for depression, which may ultimately inform clinicians who is at risk for relapse or a particularly severe outcome and would benefit from more invasive interventions such as electroconvulsive therapy [14]. However, previous work relies on extremely homogeneous study populations that are carefully selected according to strict inclusion and exclusion criteria. Any predictive models using such data are of little value when findings are to be generalized to clinical reality—highly diverse inpatient populations [15,16]. In contrast to the aforementioned data from homogeneous, well-characterized study samples, data that are routinely gathered in clinical practice are highly heterogeneous, unvalidated, and often not standardized or are inaccessible for predictive analysis [17].

In order to achieve the long-term aims of precision medicine in major depressive disorder treatment, the implementation of a standardized data collection routine in naturalistic environments from real clinical populations is needed. While high temporal resolution data on sleep and activity levels can be tracked with smartphone- or wearable technology–based solutions [18], differentiated data on the patients’ mood or affective state—the core features of major depressive disorder diagnosis—are needed for precise and valid models of affective disorders that accurately reflect the disease and treatment course. As this information needs to be provided by the patients themselves, much emphasis needs to be put on the participatory (P4) aspect of precision medicine in psychiatry. Patients thus need to be engaged and participate actively by contributing either self-report measures at regular intervals or by participating in clinical interviews or ratings. As external assessments are time-intensive and require clinical training, the use of self-rating scales, which can be completed by patients independent from the presence of a researcher or clinician, might be preferable. Recent evidence suggests reasonably high agreement when comparing patient self-reports with diagnostic clinical interviews [19], which supports their use in clinical practice. Previous studies also found the incorporation of self-report measures of symptom severity into routine care to foster engagement between patients and health care professionals and enhance care delivery in various fields of medicine [20]. In psychiatric populations with affective disorders however, questions can be raised as to the patients’ ability and motivation to provide such data, considering the lack of energy as well as cognitive impairments that define major depressive disorder during an incapacitating episode requiring inpatient treatment [21,22]. The difficulty of recruiting depressed patients for randomized controlled trials has been well documented [23], although some investigations revealed that patients did report positive attitudes toward research participation when they felt they were contributing meaningfully to the advancement of major depressive disorder treatments [24]. It remains unclear how the collection of standardized patient reports throughout the treatment course would compare in inclusion rate to the usually much more time-intensive and elaborate study protocol of a randomized controlled trial. It remains equally unknown whether certain patient subgroups may be systematically less willing or able to provide such data regularly, either due to their symptom severity or other disease-specific or sociodemographic factors, which may constitute exclusion criteria in randomized controlled trials.

Another point to consider when striving to make health care truly participatory is that assessments should preferably be collected in a digital format, as digitization allows for quick data analysis and, ideally, feedback for patients on their personal outcomes [15,25]. In general, digital solutions outperform paper-and-pencil questionnaires in practicality, acceptability, and completeness of data across studies in different fields [26-28]. Digital data collection with tablet computers, specifically, is well accepted among psychiatric patients [29]. However, previous investigations with patient reports in psychiatry only included a single assessment or pre–post comparisons as opposed to tracking individual symptom levels throughout the duration of their hospitalization [27]. The feasibility and acceptability of such a study protocol in psychiatric populations remains therefore hitherto unclear.

We established a system of longitudinal digital data collection that gives patients the opportunity to participate actively in providing data concerning their mood and symptom levels throughout the course of their inpatient treatment for an affective disorder via tablet computers. This study assessed whether affective disorder inpatients are willing and able to participate in repeated digital data entry throughout the treatment course. We additionally examined whether age, gender, symptom...
severity, and global functioning systematically co-vary with the feasibility and acceptability of such research efforts in clinical populations with affective disorders. We furthermore aimed to validate self-report measures of depression severity with the use of an external assessment performed by a clinical interviewer.

Methods

Sample
A total of 364 psychiatric patients who were recently admitted to the inpatient service of the Department of Psychiatry, University of Münster were approached during the assessment period from March 2019 to March 2020. Patients who were admitted to the closed ward could not be assessed. Patients who were admitted and discharged over the course of one weekend could equally not be assessed due to the study design, which required presence of study personnel to assist in tablet handling. Criteria for initial eligibility were therefore the admission to any of the open inpatient services of the hospital, a treatment duration of more than 3 days, and the diagnosis of any affective disorder (International Statistical Classification of Diseases, Tenth Revision codes F30.0 through F39.9) at the time of admission. In order to be included as an active participant, patients needed to be sufficiently mentally stable, cognitively able, and proficient in reading and writing German to fill in questionnaires. Due to the naturalistic setting of this investigation, inclusion criteria were intentionally kept as broad as possible in order to achieve the best possible representation of the true population seeking psychiatric inpatient treatment. The study was approved by the local institutional review board and written informed consent was obtained before participation. Patients did not receive compensation for their participation.

Procedure
Patients with the appropriate diagnoses were identified with a patient recruitment system [30] within the electronic health record based on the diagnosis entered by the treating clinician after an initial examination and diagnostic exploration on the day of admission to the ward. The clinical team approved research participation for all included patients and could consent. When eligible participants refused, their reasons for refusal were recorded and later classified into 4 categories: lack of interest or problems with handling the equipment. The amount of assistance that patients required with handling the tablet was rated, and the time they took for data entry was recorded immediately after each assessment.

Data were entered via Apple iPads, using the Mobile Patient Survey [31], a web-based multilanguage electronic patient-reported outcome system. The standardized data processing and the standardized data export were realized with the single-source metadata architecture transformation [32], an extension of the electronic health record system which uses Module Driven Software Development to generate standardized apps.

Completed digital assessments were exported into the electronic health record automatically and could be accessed in full detail (including patient responses to all questions or items) by the clinical team. Participants did not receive insight into their data automatically, however, clinicians could provide them with updates on individual outcomes or assessment results upon request.

Assessments and Measures
Reasons for exclusion were predefined according to the following categories: organizational reasons, severe cognitive deficits, insufficient language skills, and objective mental distress, with the last item referring to any psychological symptoms that would hinder participation or the ability to consent. When eligible participants refused, their reasons for refusal were recorded and later classified into 4 categories: lack of interest in the study, subjective mental distress, lack of general adherence, and data security concerns.

An external judgement of patients’ tablet-handling competency was made at each assessment based on a 4-point Likert-scale according to the following categories: 1, no required support: patient enters data independently; 2, little required support: patient needs few instructions before entering data; 3, some required support: patient needs instructions several times during data entry; and 4, a lot of required support: patient largely depends on the researcher for data entry. The median was calculated from all support ratings.

The researcher kept the time in minutes of each data entry. In order to achieve an individual entry pace factor, which signifies the deviation from the group mean, patients’ individual times

https://mental.jmir.org/2020/12/e24066

J. Med. Health 2020 | vol. 7 | iss. 12 | e24066 | p. 40

XSL-FO RenderX
for the baseline, interim, and discharge assessments were divided by the group mean for each assessment. A mean was calculated from these 3 assessments, resulting in a relational measure of individual data entry pace.

A digital version of the Beck Depression Inventory (BDI) [33,34] was used as a self-report measure of depressive symptoms. The Hamilton Depression Scale (HAMD) [35] and the Global Assessment of Functioning (GAF) [36] were conducted by the researcher as an objective measure of depression severity and global (ie, psychological, social, and occupational) functioning. An overview of all instruments included in the assessment battery can be found in Multimedia Appendix 1.

Statistical Analyses

Statistics were computed using SPSS software (version 26; IBM Corp). For all models, uncorrected $P$ values as well as Benjamini-Hochberg false discovery rate–corrected $P$ values were generated.

Participants

Required Support

To assess the influence of age, gender, depression severity and global functioning on the amount of required researcher support during data entry, we estimated an ordinal logistic regression model that included age, gender, and the baseline sum scores for BDI, HAMD, and GAF as predictors and required support as the dependent variable.

Data Entry Pace

A linear regression model was used to investigate the influence of these same variables on data entry pace. We estimated a linear regression model with age, gender, and the baseline sum scores for BDI, HAMD, and GAF as predictors and entry pace as the dependent variable.

Self-Report Measure Validation

In order to validate the self-report measure of depression severity with an external assessment, BDI and HAMD baseline sum scores were first correlated. To check for differences in agreement between self-reports and external assessments depending on age and gender, we additionally investigated potential interactions with age and gender based on linear regression models. The first model included BDI, age, and the interaction term age $\times$ BDI as predictors and HAMD as the dependent variable. The second model included BDI, gender, and the interaction term gender$\times$BDI as predictors and HAMD as the dependent variable.

Nonparticipants

Two-tailed independent sample $t$ tests and chi-square tests were calculated to assess whether patients who were excluded by clinicians or study personnel and patients who refused participation differed in age or gender. The same tests were used to assess age and gender differences between participants and nonparticipants, while potential differences in depression severity and global functioning between these 2 groups could not be compared, as the data were not available for nonparticipants.

Results

Participants

Overview

All 242 participants were diagnosed with an affective disorder. The majority of our sample had a diagnosis of major depressive disorder (215/242, 88.8%), and 11.2% (27/242) had bipolar disorder. Additionally, 97 out of 242 participants (40.1%) were diagnosed with at least 1 psychiatric comorbidity, such as anxiety disorders, eating disorders, or personality disorders while 40 out of 242 (16.5%) participants also had a diagnosed somatic comorbidity. On average, participants completed 4 assessments during their hospital stay with a minimum of 1 and a maximum of 15 assessments. The average length of stay was 9.46 weeks (SD 6.88). Adherence with assessments was generally high, with an average of 79% of expected assessments being completed during the duration of treatment. Two participants dropped out of the study before their scheduled discharge from inpatient treatment, stating that the regular participation was disrupting their daily schedule. Two additional participants had to be excluded after initial participation: 1 due to cognitive limitations exacerbated by electroconvulsive therapy and 1 due to the development of delirious symptoms during the treatment course.

Required Support

Of 242 participants, 216 (89.3%) participants did not require support and managed data entry independently during all assessments. Little support was needed by 16 patients (6.6%), whereas 7 patients (2.9%) required some support, and 2 patients (0.8%) struggled to enter data independently and relied largely on the researcher for assistance.

For the ordinal logistic regression, predictor variables were tested a priori to rule out violations of the assumption of multicollinearity. Model fit was given ($\chi^2=26.1, P<.001$). According to Nagelkerke $R^2$, the model explained 23.1% of the variance in required support. Age was found to be the only significant contributor to the model as can be seen in Table 1; the odds of needing support with data entry and tablet handling increased with older age (odds ratio [OR] 1.08, 95% CI 1.04-1.12; $P_{FDR}=.004$). Gender ($P_{FDR}=.94$), depressive symptom severity (BDI: $P_{FDR}=.68$; HAMD: $P_{FDR}=.52$), and global level of functioning ($P_{FDR}=.34$) did not contribute significantly to the model.
were excluded or refused to participate. The group of
out of the 364 patients who were eligible for inclusion, 122

BDI and age (\(P=0.519, t=8.90, P<.001, P_{\text{FDR}}=.004\)). There was a trend for global level

functioning to predict entry pace, although this association

was not upheld when controlling for multiple comparisons

(\(P=.184, t=-2.24, P_{\text{FDR}}=.09\)). Gender (\(P_{\text{FDR}}=.27\) and

level of depressive symptoms (BDI: \(P_{\text{FDR}}=.94\); HAMD:

\(P_{\text{FDR}}=.94\)) revealed no effect (Table 2).

The linear regression model was significant and explained 31.1%
of the variance in data entry time (\(R^2=0.311\), \(F_{5,213}=19.0\),

\(P<.001\)). Age was a significant predictor of entry pace (\(\beta=0.519,

t=8.90, P<.001, P_{\text{FDR}}=.004\)). The additional regression models confirmed these

results: The first regression model was significant, with BDI

gender explaining 47.3% of the variance in HAMD

(\(F_{1,212}=63.4, P<.001, R^2=0.473\)). BDI was a significant predictor

(\(F_{1,212}=186.3, P<.001, P_{\text{FDR}}=.004\)) while gender was not

(\(F_{1,212}=0.3, P=.58, P_{\text{FDR}}=.90\)). There was no significant

interaction between BDI and gender (\(F_{1,212}=0.1, P=.717,

P_{\text{FDR}}=.94\)). The second regression model was also significant, with BDI and age explaining 48.2% of the variance in HAMD

(\(F_{1,212}=65.9, P<.001, R^2=0.482\)). BDI was a significant predictor

(\(F_{1,212}=18.2, P<.001, P_{\text{FDR}}=.004\)) while age was not (\(F_{1,212}=0.1,

P=.81, P_{\text{FDR}}=.94\)). There was no significant interaction between BDI and age (\(F=1.3, P=.26, P_{\text{FDR}}=.52\)).

Nonparticipants

Out of the 364 patients who were eligible for inclusion, 122

(33.5%) were excluded or refused to participate. The group of

nonparticipants could be split into patients who were excluded by clinicians or study personnel (77/122; 63.1%) and patients who refused participation upon being approached for the study (45/122; 36.9%). There were no differences in age (t\(_{120}=0.207, P=.84, P_{\text{FDR}}=.94\) or gender (n=122, \(\chi^2=0.003, P=.96, P_{\text{FDR}}=.96\) between these 2 subgroups.

Half of the patients who had to be excluded from the study were

excluded due to organizational reasons such as a very short hospital stay (ie, under 1 week). Other reasons for exclusions were insufficient German language proficiency, limited cognitive ability (ie, severe attentional or memory deficits), and acute mental distress as judged by the treating clinician (ie, severe agitation, psychotic symptoms, or tendency to dissociate). Within the group of patients who refused to participate, a majority cited general noninterest in the study as their reason for refusal as they were not willing to take on the extra effort of completing regular assessments during the course of treatment. Some patients also expressed that they were not interested, as they did not feel like they would personally benefit from the study. Fewer patients expressed that they felt too incapacitated by their symptoms to participate, displayed general nonadherence with treatment and thus refused to participate in additional assessments, or expressed concerns over data security.

### Table 1. Ordinal logistic regression analysis results predicting the effects of age, gender, and symptom severity on support required during data entry.

<table>
<thead>
<tr>
<th>Variables</th>
<th>B (SE)</th>
<th>Wald</th>
<th>Odds ratio (95% CI)</th>
<th>P value</th>
<th>(P_{\text{FDR}}) value(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.08 (0.02)</td>
<td>15.96</td>
<td>1.08 (1.04-1.12)</td>
<td>&lt;.001</td>
<td>.004</td>
</tr>
<tr>
<td>Gender (male=1)</td>
<td>0.16 (0.54)</td>
<td>0.09</td>
<td>1.18 (0.41-3.42)</td>
<td>.76</td>
<td>.94</td>
</tr>
<tr>
<td>Beck Depression Inventory</td>
<td>-0.03 (0.04)</td>
<td>0.69</td>
<td>0.97 (0.90-1.05)</td>
<td>.41</td>
<td>.68</td>
</tr>
<tr>
<td>Hamilton Depression Scale</td>
<td>-0.08 (0.07)</td>
<td>1.15</td>
<td>0.93 (0.81-1.07)</td>
<td>.28</td>
<td>.52</td>
</tr>
<tr>
<td>Global Assessment of Functioning</td>
<td>-0.07 (0.05)</td>
<td>2.03</td>
<td>0.93 (0.85-1.03)</td>
<td>.16</td>
<td>.34</td>
</tr>
</tbody>
</table>

\(a\) Model Nagelkerke \(R^2=0.231\).

\(b\) Benjamini-Hochberg false discovery rate–corrected \(P\) value.

### Table 2. Linear regression results predicting the effects of age, gender, and symptom severity on required time for data entry.

<table>
<thead>
<tr>
<th>Variables</th>
<th>B (SE)</th>
<th>(\beta)</th>
<th>(t) value</th>
<th>(P) value</th>
<th>(P_{\text{FDR}}) value(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.009 (0.001)</td>
<td>0.519</td>
<td>8.90</td>
<td>&lt;.001</td>
<td>.004</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.053 (0.033)</td>
<td>-0.095</td>
<td>-1.61</td>
<td>.11</td>
<td>.27</td>
</tr>
<tr>
<td>Beck Depression Inventory</td>
<td>0.000 (0.002)</td>
<td>0.019</td>
<td>0.233</td>
<td>.82</td>
<td>.94</td>
</tr>
<tr>
<td>Hamilton Depression Scale</td>
<td>-0.001 (0.004)</td>
<td>-0.012</td>
<td>-0.13</td>
<td>.90</td>
<td>.94</td>
</tr>
<tr>
<td>Global Assessment of Functioning</td>
<td>-0.006 (0.003)</td>
<td>-0.184</td>
<td>-2.24</td>
<td>.03</td>
<td>.09</td>
</tr>
</tbody>
</table>

\(a\) Model \(R^2=0.313\).

\(b\) Benjamini-Hochberg false discovery rate–corrected \(P\) value.

### Self-Report Measure Validation

We found overall high agreement between the patient reported outcome of depression severity and the external clinical rating of depression severity as demonstrated through a strong positive correlation between BDI and HAMD sum scores (\(r_{214}=0.69, P<.001\)). The additional regression models confirmed these results: The first regression model was significant, with BDI and gender explaining 47.3% of the variance in HAMD (\(F_{1,212}=63.4, P<.001, R^2=0.473\)). BDI was a significant predictor (\(F_{1,212}=186.3, P<.001, P_{\text{FDR}}=.004\)) while gender was not (\(F_{1,212}=0.3, P=.58, P_{\text{FDR}}=.90\)). There was no significant interaction between BDI and gender (\(F_{1,212}=0.1, P=.717, P_{\text{FDR}}=.94\)). The second regression model was also significant, with BDI and age explaining 48.2% of the variance in HAMD (\(F_{1,212}=65.9, P<.001, R^2=0.482\)). BDI was a significant predictor (\(F_{1,212}=18.2, P<.001, P_{\text{FDR}}=.004\)) while age was not (\(F_{1,212}=0.1, P=.81, P_{\text{FDR}}=.94\)). There was no significant interaction between BDI and age (\(F=1.3, P=.26, P_{\text{FDR}}=.52\)).
Please refer to Figure 1 for more detailed visualization of the distribution of reasons for exclusion and refusal to participate. The nonparticipating group was significantly older ($t_{362}=3.31$, $P<.001$, $P_{FDR}=.004$), and there was a trend of this group to consist of more women than the participating group, although this association was not upheld when correcting for multiple comparisons ($n=364; \chi^2=4.34, P=.04, P_{FDR}=.11$). See Table 3 for sociodemographic and clinical characteristics.

Figure 1. Study flow chart.
### Discussion

With this study, we demonstrated the feasibility and acceptance of repeated digitized assessment of risk and symptom profiles in affective disorder inpatients. Our results indicate that participatory medicine can be achieved in patients with affective disorders, as they are willing and able to contribute self-report measures throughout the duration of their inpatient treatment. This study, therefore, provides important insight into the possibility of routinely collecting longitudinal data in real-world clinical cohorts that may guide the way toward personalized psychiatric care.

During an assessment period of 1 year, we achieved an inclusion rate of 66.5% of patients with a diagnosed affective disorder. This rate is similar to those reported from other investigations performed on the general population and nonpsychiatric patient groups, which indicates that major depressive disorder or bipolar disorder symptomology does not constitute a barrier toward participation [37,38]. Adherence to assessments was high, and the dropout rate after initial participation was very low. Exclusion by the clinician or researcher and refusal to participate by the patients themselves were largely not due to symptom severity or cognitive impairment but for organizational reasons or a general disinterest in the study, which mirrors reasons for nonparticipation in research from nonpsychiatric populations [39]. It also confirms previous research on the generally high level of acceptance of longitudinal self-reporting technology in patients with bipolar disorder [40,41].

A vast majority of patients who did participate were able to enter data independently and did not encounter technical difficulties. More importantly, we found no association between symptom severity at baseline and the amount of required support in handling the equipment or prolonged data entry times. This is in line with previous investigations on the feasibility and acceptability of digitally based assessments in psychiatric populations [29]. Moreover, we were able to demonstrate the validity of patient reports with the use of an external measure of depression severity performed by a clinical interviewer. The level of agreement between self-reported depression severity with the BDI and the external rating based on the HAMD was comparable to findings from the literature [42] and indicated high validity of the digital self-report measure. This is an especially promising result for the implementation of patient-reported data collection technologies into routine documentation as well as its use for research purposes because it indicates that little to no additional personnel resources are required in order to gain valuable longitudinal data throughout the course of treatment. Moreover, the digital implementation of such assessments will allow for more accurate data collection and immediate data storage and analysis [15]. In the future, such an infrastructure of digital data collection could be used to communicate treatment outcomes and visual representations thereof directly to patients. Similar approaches have been found to improve communication between patients and health care providers [43] and would also constitute an improvement in the participatory aspect of precision medicine.

Although our results generally support the feasibility of longitudinal digital data collection in affective disorders, a few
systematic recruitment and accessibility issues must be addressed. Despite the fact that no association between symptom severity and performance during data entry in our participating sample was detected, 6% of all potential participants were excluded beforehand due to reduced cognitive ability or clinicians’ concerns over their acute mental distress. Although this embodies only a small percentage of our sample, this result suggests that a systematic exclusion of more severe cases may not be avoidable. Nevertheless, it may be worthwhile to critically consider such cases individually, as it has been shown that carees overestimate the amount of distress patients are put under during research participation [24].

We, furthermore, found a statistical trend for women to be more likely to refuse or be excluded from the study than men. Although this finding did not survive the correction for multiple comparisons, it is nevertheless meaningful to consider potential reasons this gender disparity may occur, as it may again be due to symptom severity. Although the reasons recorded at the time of exclusion do not suggest symptom severity is the main factor, evidence does suggest that women are more likely to seek psychiatric treatment and report more severe symptoms [44,45]. As we do not have data on the symptom levels of the excluded group, this question cannot be answered with certainty.

Regardless of gender, the factor that impacted both the amount of required assistance and time during data entry was older age. Older adults found it more difficult to handle the tablet-based assessments and took longer to complete them. However, the percentage of participants who needed assistance was comparatively small, and even those who did require assistance were able to complete assessments regularly, which suggests that their difficulties with handling the equipment did not stop them from participating. Although our study did not assess subjective attitudes toward technology, previous studies [38,46] found that digital methods of data collection are well-accepted even among older adults. It can also be expected that technological literacy will rise in older populations over the years, as smartphones and tablets are becoming increasingly ubiquitous, which will alleviate the difficulties for this specific age group in the future. Studies [47,48] also show that, although older adults lag behind in digital literacy, such competencies can be acquired through social support. Nevertheless, our findings suggest that options for support of older participants should currently be offered to not systematically exclude technologically less well-versed patients from participatory care.

In addition to older age, there was also a trend for lower global functioning to be related to slower data entry pace; however, it was not associated with the amount of required assistance. Even though this association was not upheld when correcting for multiple comparisons, this finding should also be critically discussed. It suggests that patients with a generally lower level of global functioning take longer to complete assessments but are still able to do so independently. Moreover, the added time expenditure does not lead to participants dropping out of the study, indicating that the slower entry pace is tolerable and not a barrier that would keep lower functioning patients from participating in such research.

Furthermore, patients who are not proficient in the language spoken by their health care providers are a systematically disadvantaged group in psychiatric care who could not be included in this investigation. At equal or greater levels of need, persons with an immigration background are known to seek mental health treatment less often and are less likely to report favorable treatment outcomes [49,50]. This suggests that the inclusion of marginalized populations would be of great importance especially when investigating individual risk factors for affective disorders on the way to precision psychiatry. In fact, digitally assessed self-report measures present the opportunity to get detailed, standardized assessments despite language barriers, as questionnaire measures can be made available in every language. The app that we used for data collection supports the implementation of multilanguage assessments [51] and could therefore be used in future investigations in order to also reach and assess non-German speaking clinical populations. This would provide a wealth of standardized, quantifiable information about patients of diverse cultural backgrounds that could guide treatment but also assist in identifying suitable interventions for clinical populations with specific ethnic or cultural differences and risk or resilience factors.

To the best of our knowledge, this study is the first naturalistic investigation to incorporate repeated, digitally assessed patient reports throughout the course of inpatient treatment in affective disorders. Overall, the acceptability and feasibility of such study protocols within the clinical routine is high while required resources remain comparatively low. Patients are willing and able to provide data at regular intervals and are not systematically disadvantaged by the severity of their affective symptoms. Future implementations should keep gender, age, and cultural factors in mind when approaching patients and offer assistance with any technological equipment as needed.

In conclusion, this study is a first step in demonstrating that the participatory aspect of precision medicine can be achieved in psychiatry. In the future, the information gathered routinely through patient reported assessments could be combined with other potential data sources such as fitness trackers and information gained from electronic health records [25,52-54]. This may pave the way for data-driven predictive models that could, in the more distant future, be used to predict and prevent the occurrence of affective disorders, as well as facilitate the identification of individual risk profiles. Even without the use of predictive modeling, self-reports and the direct exchange of such information between patients and clinicians might improve treatment [43]. Overall, such advances in psychiatry will be invaluable as personalized treatments tailored to such individual risk factors may lead to much shorter and less frequent hospitalizations, which would equate to more cost-effective treatments and a pronounced reduction in patient suffering.
Acknowledgments

Funding was provided by the Interdisciplinary Center for Clinical Research of the medical faculty of Münster (grant SEED 11/19 to NO), as well as the Innovative Medizinische Forschung of the medical faculty of Münster (grants OP121710 to NO).

We are deeply indebted to all participants of this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary table.
[DOCX File, 34 KB - mental_v7i12e24066_app1.docx ]

References


10. Richter et al JMIR Mental Health 2020 | vol. 7 | iss. 12 | e24066 | p. 46 https://mental.jmir.org/2020/12/e24066


Abbreviations

BDI: Beck Depression Inventory
GAF: Global Assessment of Functioning
HAMD: Hamilton Depression Scale
Original Paper

Mobile-Assisted Cognitive Behavioral Therapy for Negative Symptoms: Open Single-Arm Trial With Schizophrenia Patients

Eric Granholm1, PhD; Jason Holden2, PhD; Kristen Dwyer2, PhD; Tanya Mikhael2, BA; Peter Link2, MA; Colin Depp1,2, PhD

1VA San Diego Healthcare System, San Diego, CA, United States
2Department of Psychiatry, University of California, San Diego, San Diego, CA, United States

Corresponding Author:
Eric Granholm, PhD
VA San Diego Healthcare System
9500 Gilman Drive #0737
San Diego, CA
United States
Phone: 1 858 534 2542
Email: egranholm@ucsd.edu

Abstract

Background: Negative symptoms are an important unmet treatment need for schizophrenia. This study is a preliminary, open, single-arm trial of a novel hybrid intervention called mobile-assisted cognitive behavioral therapy for negative symptoms (mCBTn).

Objective: The primary aim was to test whether mCBTn was feasible and could reduce severity of the target mechanism, defeatist performance attitudes, which are associated with experiential negative symptoms and poor functioning in schizophrenia.

Methods: Participants with schizophrenia or schizoaffective disorder (N=31) who met prospective criteria for persistent negative symptoms were enrolled. The blended intervention combines weekly in-person group therapy with a smartphone app called CBT2go. The app extended therapy group skills, including recovery goal setting, thought challenging, scheduling of pleasurable activities and social interactions, and pleasure-savoring interventions to modify defeatist attitudes and improve experiential negative symptoms.

Results: Retention was excellent (87% at 18 weeks), and severity of defeatist attitudes and experiential negative symptoms declined significantly in the mCBTn intervention with large effect sizes.

Conclusions: The findings suggest that mCBTn is a feasible and potentially effective treatment for experiential negative symptoms, if confirmed in a larger randomized controlled trial. The findings also provide support for the defeatist attitude model of experiential negative symptoms and suggest that blended technology-supported interventions such as mCBTn can strengthen and shorten intensive psychosocial interventions for schizophrenia.

Trial Registration: ClinicalTrials.gov NCT03179696; https://clinicaltrials.gov/ct2/show/NCT03179696

(JMIR Ment Health 2020;7(12):e24406) doi:10.2196/24406

KEYWORDS
motivation; persistent negative symptoms; dysfunctional attitudes; mHealth; blended intervention; mobile phone

Introduction

Negative symptoms account for much of the poor functional outcome in schizophrenia and are an unmet treatment need [1-3]. Negative symptoms can refer to reduced expressive (eg, facial affect and voice tone) or experiential (eg, avolition and asociality) symptoms, which comprise two separate factors [4-6]. Experiential negative symptoms are particularly important to treat, because they are strongly associated with functioning [5,7]. Unfortunately, available pharmacological treatments have only limited benefits for negative symptoms [8,9].

Beck and colleagues [10-12] have proposed that interventions that reduce defeatist attitudes may improve negative symptoms and functioning in schizophrenia. Several studies have found that cognitions such as defeatist performance (eg, “Why try, I always fail”) and social disinterest (eg, “I’m better off alone”) attitudes are associated with negative symptoms, and to some extent poor functioning, even after accounting for depression.
Defeatist attitudes can be targeted in cognitive behavioral therapy (CBT). Clinical trials of CBT for psychosis have found mixed results for reducing negative symptoms [20,21], but some CBT interventions that specifically targeted defeatist attitudes have found more promising results [22-27]. Social skills training (SST) has also produced significant but modest improvements in negative symptoms [28,29]. In our cognitive behavioral social skills training (CBSST) intervention, which combines CBT and SST [30], we have found significant improvements in defeatist attitudes and in experiential negative symptoms and functioning; improvements in experiential negative symptoms were mediated by improvements in defeatist attitudes [23]. Modification of defeatist attitudes, therefore, may be a mechanism of change in CBSST, whereby reduction in defeatist attitudes contributes to increased motivation and effort toward goal-directed tasks. CBSST, however, is an intensive and lengthy (ie, up to 36 sessions), high-burden, multicomponent intervention. If the intervention could be shortened and the focus on reducing defeatist attitudes strengthened by using mobile interventions, the cost and burden of CBSST implementation could be reduced. In this way, blended interventions that combine mobile interventions with in-person interventions could increase access to evidence-based interventions for schizophrenia by reducing implementation barriers.

Smartphones are widely available, affordable, and frequently used by individuals with serious mental illness [31-34]. In previous clinical trials, we have found significant improvements in social functioning and symptoms in individuals with schizophrenia and bipolar disorder using 12-week mobile interventions that incorporate CBT principles with minimal therapist contact [35,36]. A number of other mobile interventions for schizophrenia have been developed and received preliminary testing, but few have specifically targeted negative symptoms or motivation and few have blended in-person plus app interventions; for reviews, see Camacho et al [37] and Firth and Torous [38]. One recent pilot trial of a blended intervention for psychosis [39] used a brief coping-focused intervention for distressing voices—SAVVy (Smartphone-Assisted coping-focused interVention for Voices)—which blended four sessions of in-person therapy with ecological momentary assessment (EMA) or ecological momentary intervention between sessions; the trial found that the intervention was feasible, improved coping, and, at a trend level, reduced severity of voices. Another app-only trial that did target motivation in schizophrenia—PRIME (Personalized Real-time Intervention for Motivational Enhancement) [40,41]—led to improvements in self-efficacy beliefs and social interactions with peers, as well as reduced defeatist beliefs and self-reported negative symptoms in individuals with recent-onset schizophrenia. These trials support the feasibility and potential benefits of this approach.

Given the promise of in-person and mobile CBT interventions targeting defeatist attitudes and motivation in schizophrenia, we developed a blended intervention, called mobile-assisted cognitive behavioral therapy for negative symptoms (mCBTn), which combines in-person, 90-minute, weekly groups with a mobile app called CBT2go. The mCBTn intervention primarily targets defeatist attitudes to improve experiential negative symptoms in schizophrenia. The CBT components and skills-training approach of our CBSST group intervention were combined with mobile thought-challenging interventions that were based on our Mobile Assessment and Treatment of Schizophrenia (MATS) [35] and CBT2go [36] interventions. This thought-challenging algorithm incorporated EMA to sample attitudes and moods in real-world contexts and then used personalized evidence to challenge dysfunctional beliefs (eg, if a participant rates their expectation for pleasure in a planned social interaction at a clubhouse as low, they would receive the challenge, “But you said you had fun at the clubhouse last time”). In addition to thought challenging, the CBT2go app and group intervention used in mCBTn both also incorporated recovery goal setting and tracking, scheduling pleasurable activities and social interactions, and pleasure-savoring interventions.

We conducted an open preliminary trial of mCBTn in patients with schizophrenia or schizoaffective disorder with moderate to severe persistent negative symptoms, and hypothesized that defeatist attitudes and experiential negative symptoms would be significantly reduced from baseline to end of treatment. This open trial was funded as part of the National Institute of Mental Health Experimental Therapeutics Program (RFA-MH-18-704 R61/R33), which involves preliminary testing of an intervention’s impact on a target mechanism (ie, defeatist attitudes) associated with an important clinical outcome (ie, experiential negative symptoms). The primary aim of the study was target engagement; that is, we hypothesized that mCBTn would lead to a significant reduction in severity of defeatist attitudes. We also assessed participants at 12, 18, and 24 weeks of treatment to determine which dose of treatment could produce at least a medium effect size (Cohen $d=0.5$) improvement in defeatist attitudes. In the Experimental Therapeutics Program, contingent on changing the target in this single-arm open-trial phase, a larger randomized controlled trial (RCT) will be conducted with treatment at the dose identified.

**Methods**

This was a single-arm, open-trial, pre-post evaluation of the feasibility and preliminary effect of mCBTn. This trial was registered at ClinicalTrials.gov (NCT03179696).

**Sample**

The study protocol was reviewed and approved by the Institutional Review Board of the University of California, San Diego, prior to initiating research activities with participants. Participants with schizophrenia or schizoaffective disorder were selected who have moderate to severe persistent experiential negative symptoms [3], as well as moderate to severe defeatist attitudes, recognizing that an intervention targeting defeatist attitudes is not likely to be helpful for consumers who do not...
have them. Participants were required to meet all inclusion and exclusion criteria over a 2-week evaluation phase. Inclusion criteria were as follows:

1. Voluntary informed consent to participate and capacity to consent.
2. Aged 18-65 years.
3. Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-5), diagnosis of schizophrenia or schizoaffective disorder based on a Structured Clinical Interview for DSM-5 (SCID-5) interview and available medical record review.
4. Moderate to severe experiential negative symptoms in at least two of the three Clinical Assessment Interview for Negative Symptoms [5] Motivation and Pleasure (CAINS-MAP) domains (mean score of 2 [ie, moderate] or greater for items averaged within the Social, Work, or Recreational domain) at the beginning and end of the 2-week evaluation phase.
5. Moderate to severe defeatist attitudes (score of >50 on the Defeatist Performance Attitude Scale [DPAS] [10]).
6. A 6th grade or higher reading level on the Wide Range Achievement Test-4 Reading subtest.
7. Clinically stable and on stable medications (ie, no hospitalizations or medication changes in 4 months prior to enrollment).

Exclusion criteria were as follows:

1. Prior CBT in the past 2 years.
2. Greater than moderate Positive and Negative Syndrome Scale (PANSS) [42] positive symptoms (score >5 for any item: item P1 [delusions], item P2 [disorganization], item P3 [hallucinations], or item P6 [suspiciousness]).
3. Severe depression on the Calgary Depression Scale for Schizophrenia (CDS) [43] (score >8).
5. DSM-5 alcohol or substance use disorder in past 3 months based on the SCID-5.
6. Level of care required interferes with outpatient therapy (eg, hospitalized or severe medical illness).
7. Unable to adequately see or manually manipulate the mobile device.

**Blended mCBTn Intervention**

The therapy group and the mobile app integrated skills-based interventions, including recovery goal setting, thought challenging, scheduling of pleasurable activities and social interactions, and pleasure-savoring interventions to modify defeatist attitudes and improve motivation and pleasure negative symptoms. A modified 12-session version of the Cognitive Skills Module of CBSST [30] was delivered in 90-minute, weekly group therapy sessions with two masters-level therapists and approximately 6 participants per group. The 12-session module provided all the core mCBTn content within the minimum number of sessions that might be identified as the optimal dose, but the module was repeated for a total of 24 sessions to determine if more sessions were needed to change the defeatist attitude target.

The mCBTn manual included a therapist guide and a patient workbook describing the skills and homework assignments, as well as a collection of games and exercises to make learning fun and promote engagement [30]. The module began with setting a meaningful living, learning, working, or socializing recovery goal, and the goal was broken down into short-term goals and goal steps. Defeatist attitudes and avoidance behaviors that interfere with working on the goal were then modified using CBT skills. Group members were introduced to the general concepts of CBT, including the relationship between thoughts, actions, and feelings (ie, generic cognitive model); thought challenging through behavioral experiments and examining evidence for beliefs; and mistakes in thinking. The primary thought-challenging skill trained was the 3Cs: Catch It, Check It, Change It—"It" is an unhelpful defeatist thought. Group time was also spent practicing the smartphone interventions, developing content for the mobile device (eg, personalized motivational and thought-challenging statements about goals, socializing, and pleasurable activities), and reviewing data collected with the device to challenge defeatist attitudes.

An iPhone 5s or 5SE was provided to all participants with an unlimited data plan and could receive and send unlimited texts and phone calls. Mobile interactions were triggered by an app notification in the morning, with reminder prompts at midday and evening. If participants responded to the notification earlier in the day, the second and/or third notifications were not delivered. The purpose of the notifications was to prompt daily engagement with the app. Device training was provided on how to operate and charge the device, the meaning of all questions and response choices, procedures for carrying the device, responding to prompts, how to access crisis lines, and how to use various apps. This information was also provided in a written manual given to participants. Participants returned the device at the end of treatment.

The CBT2go app was used to prompt and track each group member’s goal-directed activities in the community, facilitate adherence to homework assignments involving community practice of thought-challenging skills trained in group, and prompt performance and savoring of personalized pleasurable activities and social interactions planned in group. The CBT2go app used personalized statements developed in group to challenge social disinterest and defeatist attitudes in real-time, real-world environments. After groups, therapists could enter personalized comments that participants made in group into a web-based dashboard (eg, “Having coffee with Jim is fun” and “Angie always makes you laugh”), which were used by the app to challenge low expectation ratings of motivation, anticipatory pleasure, or anticipated success for planned activities (see sample screenshots in Figure 1).

Participants carried the device and received the mobile intervention for the entire 24-week, blended group-plus-app intervention period. Each day during treatment, the CBT2go app alerted participants in the morning to make an action plan, which involved responding to a multiple-choice question about what to do that day: (1) a goal step, (2) social interaction, (3) pleasurable activity, or (4) homework assignment, or take the day off. Following the participant’s choice, the CBT2go app asked for three EMA ratings (slide bar ranging from 1 [not at
all] to 7 [very much]) of (1) motivation to do the activity, (2) anticipated success, and (3) anticipated pleasure for the action plan. High ratings were reinforced (eg, “That’s right, socializing can be fun”), and low ratings were challenged (eg, “Don’t forget, you said walking on the beach was fun” or “But your goal is to make a close friend”). Behavioral experiments were then suggested to test out beliefs (eg, “Try asking someone to go for a short walk”). If a pleasurable activity action plan was selected, the pleasurable activities component of the CBT2go app was opened, which displayed several personalized activities that the participant previously entered. For the selected activity, the time and place to do the activity was queried (ie, a reminder alert was delivered at the time planned) and anticipated pleasure ratings were queried on a slide bar ranging from 0 (low) to 10 (high). Pleasure was again rated if the activity was completed, and these ratings were saved and available on demand in the group sessions for the therapist to discuss how people often think activities will not be as fun as they turn out to be and how this low expectation can reduce the likelihood of doing the activity. The app also prompts pleasure savoring of completed activities by taking a selfie or other photo or journaling about their experiences in the app. These photos and journal entries are also available on demand for review by participants or therapists in group sessions. Therapists helped develop the plans for pleasurable activities in group. Finally, if homework was the action plan selected, participants were directed to use their homework sheets and workbooks from the group to complete their weekly homework assignment.

Finally, an on-demand recovery goal–setting component of the CBT2go app was also provided and was populated during the goal-setting sessions in group by therapists and participants, as well as between sessions for homework by participants. A long-term goal was set and short-term goals and goal steps that would facilitate achievement of the long-term goal were entered into the app. The app could be accessed on demand to remind participants of goals, and goal steps could be checked off to track and motivate goal progress.

**Assessments**

Participants were assessed at 2 weeks prior to baseline; at baseline; and at 12, 18, and 24 weeks of treatment. Dysfunctional attitudes were measured on the DPAS [10] (ie, primary target mechanism) and the Asocial Beliefs Scale (ABS) [45]. The CAINS-MAP [5] subscale was the primary negative symptom outcome measure. The CDS [43] and the PANSS [42] positive symptom subscale were used to assess secondary symptom outcome domains. Functioning was assessed on the Abbreviated Quality of Life Scale (A-QLS) [46] and the Social Functioning Scale (SFS) [47].

**Statistical Analyses**

Mixed-effects regression models, utilizing HLM (hierarchical linear modeling) v6.08 (Scientific Software International), were estimated to predict each in-lab outcome assessment and mobile CBT2go app ratings of motivation, success, and anticipated pleasure for activities using time in weeks since baseline as a level-1 predictor. Paired-sample, 2-tailed t-tests between baseline and each follow-up assessment tested whether significant change was found for each outcome at each assessment point to inform the dose of treatment that might achieve a significant improvement in DPAS and CAINS-MAP scores with at least a medium effect size (Cohen $d=0.5$).

**Results**

**Sample**

We recruited and assessed 67 participants; see the CONSORT (Consolidated Standards of Reporting Trials) diagram in Figure 2. However, 36 of the 67 participants (54%) were excluded because they did not meet persistent negative symptom criteria. For the 31 participants enrolled in treatment, excellent retention was found with 28 (90%), 27 (87%), and 25 (81%) participants assessed at the 12-, 18-, and 24-week assessment points, respectively. There was only one adverse event, which was a hospitalization for symptom exacerbation in the context of medication nonadherence. The participant remained in the study. The sample had a mean age of 48.3 (SD 9.5) years and a mean of 11.8 (SD 1.5) years of education. The sample was 65%...
(20/31) male, was 65% (20/31) White, and had moderate baseline symptom severity—the PANSS total mean score was 60.5 (SD 12.7). Out of 31 participants, 21 (68%) had a schizophrenia diagnosis and 10 (32%) had schizoaffective disorder. Table 1 shows descriptive statistics and comparisons between baseline and each assessment point for all outcome variables.

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) flow diagram of participants through the open trial. CBT: cognitive behavioral therapy; DPAS: Defeatist Performance Attitude Scale.
Table 1. Outcome variables and paired \( t \) tests at each assessment point relative to baseline.

<table>
<thead>
<tr>
<th>Outcome measure and time points</th>
<th>n (%)</th>
<th>Score, mean (SD)</th>
<th>Cohen ( d )(^a)</th>
<th>2-tailed ( t ) test ((df))</th>
<th>( P ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Defeatist Performance Attitude Scale</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>31 (100)</td>
<td>66.3 (14.4)</td>
<td>N/A(^b)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>12 Weeks</td>
<td>28 (90)</td>
<td>60.7 (15.9)</td>
<td>0.40</td>
<td>2.23 (27)</td>
<td>.03</td>
</tr>
<tr>
<td>18 Weeks</td>
<td>27 (87)</td>
<td>56.4 (15.7)</td>
<td>0.70</td>
<td>4.02 (26)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>24 Weeks</td>
<td>25 (81)</td>
<td>52.2 (17.3)</td>
<td>1.00</td>
<td>4.10 (24)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td><strong>Asocial Beliefs Scale</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>30 (100)</td>
<td>5.9 (3.1)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>12 Weeks</td>
<td>28 (93)</td>
<td>6.5 (2.9)</td>
<td>-0.20</td>
<td>0.58 (27)</td>
<td>.57</td>
</tr>
<tr>
<td>18 Weeks</td>
<td>27 (90)</td>
<td>6.2 (3.2)</td>
<td>-0.10</td>
<td>0.21 (26)</td>
<td>.84</td>
</tr>
<tr>
<td>24 Weeks</td>
<td>25 (83)</td>
<td>6.4 (3.3)</td>
<td>-0.15</td>
<td>0.19 (24)</td>
<td>.85</td>
</tr>
<tr>
<td><strong>Clinical Assessment Interview for Negative Symptoms Motivation and Pleasure</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>31 (100)</td>
<td>23.1 (3.4)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>12 Weeks</td>
<td>28 (90)</td>
<td>21.3 (6.1)</td>
<td>0.55</td>
<td>2.07 (27)</td>
<td>.048</td>
</tr>
<tr>
<td>18 Weeks</td>
<td>27 (87)</td>
<td>20.5 (6.2)</td>
<td>0.75</td>
<td>2.95 (26)</td>
<td>.007</td>
</tr>
<tr>
<td>24 Weeks</td>
<td>25 (81)</td>
<td>20.0 (6.8)</td>
<td>0.90</td>
<td>3.16 (24)</td>
<td>.004</td>
</tr>
<tr>
<td><strong>Clinical Assessment Interview for Negative Symptoms Expression</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>31 (100)</td>
<td>4.9 (3.2)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>12 Weeks</td>
<td>28 (90)</td>
<td>4.7 (3.6)</td>
<td>0.05</td>
<td>0.56 (27)</td>
<td>.58</td>
</tr>
<tr>
<td>18 Weeks</td>
<td>27 (87)</td>
<td>4.3 (3.3)</td>
<td>0.20</td>
<td>1.18 (26)</td>
<td>.25</td>
</tr>
<tr>
<td>24 Weeks</td>
<td>25 (81)</td>
<td>4.4 (3.3)</td>
<td>0.15</td>
<td>1.11 (24)</td>
<td>.28</td>
</tr>
<tr>
<td><strong>Positive and Negative Syndrome Scale positive symptoms</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>31 (100)</td>
<td>13.4 (4.5)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>12 Weeks</td>
<td>28 (90)</td>
<td>13.0 (4.3)</td>
<td>0.10</td>
<td>0.75 (27)</td>
<td>.46</td>
</tr>
<tr>
<td>18 Weeks</td>
<td>26 (84)</td>
<td>12.8 (4.8)</td>
<td>0.10</td>
<td>0.64 (25)</td>
<td>.52</td>
</tr>
<tr>
<td>24 Weeks</td>
<td>24 (77)</td>
<td>11.5 (4.0)</td>
<td>0.45</td>
<td>2.46 (23)</td>
<td>.02</td>
</tr>
<tr>
<td><strong>Calgary Depression Scale</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>31 (100)</td>
<td>3.4 (2.2)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>12 Weeks</td>
<td>28 (90)</td>
<td>3.4 (2.6)</td>
<td>0</td>
<td>0 (27)</td>
<td>&gt;.99</td>
</tr>
<tr>
<td>18 Weeks</td>
<td>27 (87)</td>
<td>3.1 (2.5)</td>
<td>0.10</td>
<td>0.85 (26)</td>
<td>.40</td>
</tr>
<tr>
<td>24 Weeks</td>
<td>25 (81)</td>
<td>2.7 (2.5)</td>
<td>0.30</td>
<td>1.99 (24)</td>
<td>.06</td>
</tr>
<tr>
<td><strong>Abbreviated Quality of Life Scale</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>31 (100)</td>
<td>23.3 (6.1)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>12 Weeks</td>
<td>28 (90)</td>
<td>25.8 (7.9)</td>
<td>0.40</td>
<td>2.17 (27)</td>
<td>.04</td>
</tr>
<tr>
<td>18 Weeks</td>
<td>27 (87)</td>
<td>24.9 (8.3)</td>
<td>0.25</td>
<td>1.78 (26)</td>
<td>.09</td>
</tr>
<tr>
<td>24 Weeks</td>
<td>25 (81)</td>
<td>24.8 (8.8)</td>
<td>0.25</td>
<td>1.77 (24)</td>
<td>.09</td>
</tr>
<tr>
<td><strong>Social Functioning Scale</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>31 (100)</td>
<td>114.7 (21.8)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>12 Weeks</td>
<td>27 (87)</td>
<td>118.5 (22.2)</td>
<td>0.15</td>
<td>2.27 (26)</td>
<td>.03</td>
</tr>
<tr>
<td>18 Weeks</td>
<td>27 (87)</td>
<td>115.9 (23.6)</td>
<td>0.05</td>
<td>1.70 (26)</td>
<td>.10</td>
</tr>
<tr>
<td>24 Weeks</td>
<td>25 (81)</td>
<td>117.3 (26.7)</td>
<td>0.10</td>
<td>1.70 (24)</td>
<td>.10</td>
</tr>
</tbody>
</table>

\(^a\)A positive Cohen \( d \) value indicates improvement on the outcome from baseline to follow-up.
Dysfunctional Attitudes
The effect of time was significant for defeatist performance attitudes (DPAS: \( \gamma = -0.59, t_{30} = -4.27, P < 0.001 \)). Change in DPAS score from baseline (see Table 1) was significant at all assessment points during treatment, with medium to large effect sizes. The minimal treatment dose needed to achieve at least a medium effect size was at 18 weeks. In contrast, the effect of time was not significant for asocial beliefs (ABS: \( \gamma = 0.01, t_{29} = 0.47, P = .64 \)).

Symptoms
Significant reduction in severity of experiential negative symptoms was found (CAINS-MAP: \( \gamma = -0.14, t_{29} = -3.12, P = 0.004 \)) with medium to large effect sizes (see Table 1). The minimal treatment dose needed to achieve at least a medium effect size was at 12 weeks. For expressive negative symptoms, assessed using the CAINS Expression (CAINS-EXP) subscale, the effect of time was not statistically significant and no significant reduction from baseline was found at any assessment point (CAINS-EXP: \( \gamma = -0.02, t_{30} = -1.06, P = .30 \)). Significant reduction in severity of positive symptoms was found by week 24, but not at earlier assessment points, and the effect of time was not significant (PANSS-positive subscale: \( \gamma = -0.06, t_{30} = -1.46, P = .16 \)). Similarly, significant reduction in severity of depressive symptoms was found by week 24 but not at earlier assessment points, and the effect of time was not significant (CDS: \( \gamma = -0.03, t_{30} = -1.66, P = .11 \)).

Mobile Symptom and Attitude Ratings
CBT2go app ratings of motivation and anticipated pleasure and success for completing planned activities increased significantly for motivation (\( \gamma = 0.007, t_{28} = 2.17, P = .04 \)) and anticipated pleasure (\( \gamma = 0.008, t_{28} = 2.53, P = .02 \)) but not for anticipated success (\( \gamma = 0.006, t_{28} = 1.62, P = .12 \)). This indicates steady gains in self-reported motivation and pleasure (0.7-0.8 points per 100 days on a 0-7 scale) over treatment. To estimate effect sizes for these changes over the course of treatment, we computed the mean ratings for the first 4 weeks and last 4 weeks of treatment for mobile CBT2go app ratings. Large increases in mean ratings were found between the first 4 weeks of treatment and the last 4 weeks for motivation (Cohen \( d = 0.65 \); first 4 weeks: mean 2.9, SD 1.8; last 4 weeks: mean 4.1, SD 1.8) and anticipated pleasure (Cohen \( d = 0.80 \); first 4 weeks: mean 3.0, SD 1.5; last 4 weeks: mean 4.2, SD 2.1). Medium increases were found for anticipated success (Cohen \( d = 0.55 \); first 4 weeks: mean 3.1, SD 1.5; last 4 weeks: mean 3.9, SD 1.9).

Functioning
Significant improvement in A-QLS scores was found between baseline and 12 weeks but not at other assessment points, and the effect of time was not significant (\( \gamma = 0.08, t_{30} = 1.49, P = .15 \)). Participants also showed significant improvement on the SFS total score between baseline and 12 weeks but not at the other two assessment points, and the effect of time was only at a trend level (\( \gamma = 0.19, t_{30} = 1.77, P = .09 \)).

App Engagement
The CBT2go app prompted participants to select an action plan each day during up to 168 days of treatment. There was a mean of 18.7 (SD 21.3) responses to 84 action plan prompts (22%) at 12 weeks and a mean of 32.3 (SD 31.5) responses to 168 action plan prompts (19.2%) at 24 weeks for participants who did not drop out of treatment by each assessment point; this indicates engagement in homework and skills practice more than once per week, with minimal fatigue effects over the course of treatment. The number of action plans completed was not significantly correlated with any symptom measure at baseline (range of \( r = -0.05 \) to 0.07).

In HLM analyses examining the association between app engagement and outcome, the number of action plans by time interaction was significant for change in CAINS-MAP scores (\( \gamma = -0.003, t_{29} = -2.11, P = .04 \)) but not for DPAS scores (\( \gamma = -0.002, t_{29} = -0.36, P = .72 \)). The change in CAINS-MAP scores relative to baseline was also marginally correlated with the number of action plans completed at 12 weeks (\( r = -0.33, P = .09 \)) and 18 weeks (\( r = -0.35, P = .07 \)) but not 24 weeks (\( r = -0.28, P = .18 \)). This did not hold true for change in DPAS scores relative to baseline at 12 weeks (\( r = -0.23, P = .24 \)), 18 weeks (\( r = -0.13, P = .51 \)), and 24 weeks (\( r = -0.07, P = .75 \)). Thus, greater engagement with the app was associated with greater reduction in motivation and pleasure negative symptoms.

Discussion
The results of this open trial of mCBTn showed significant, large, within-group improvements in defeatist attitudes and negative symptoms and defeatist attitudes by 18 weeks of treatment, which demonstrates feasibility and engagement of the defeatist attitudes target, and justifies a larger RCT. These findings also provide support for the defeatist attitude model of negative symptoms [10,12]. Some specificity was found for improvements in defeatist performance beliefs but not asocial beliefs, perhaps because the intervention focused on all living, learning, working, and socializing goals, rather than only socialization. The intensive focus of mCBTn on modifying defeatist beliefs, both in group therapy and in real-world environments using the CBT2go app, resulted in a large improvement in experiential negative symptoms within a relatively short 18-week treatment duration and medium improvement by 12 weeks, which is relatively rapid change. If blended interventions such as mCBTn can strengthen and thereby shorten intensive psychosocial interventions, then implementation barriers associated with burden and cost of lengthy psychotherapy interventions would be reduced. This would improve access to evidence-based practices for consumers with severe and persistent negative symptoms, which would have considerable public health significance. However, further testing in a larger RCT with a control condition is needed to confirm the efficacy of mCBTn and justify work on implementation.

This study adds to the growing literature on CBT-based mobile interventions for schizophrenia [37,38,48,49] and indicates that
patients with severe and persistent negative symptoms may be amenable to this approach. In the first trial to use mobile CBT-informed interventions for schizophrenia in a text-messaging platform, we [35] used a similar approach that integrated EMA of symptoms and behaviors (e.g., social isolation, medication adherence, and voices) in everyday contexts and delivered personalized just-in-time thought-challenging messages when participants reported symptom distress or defeatist attitudes; we also found improvements in auditory hallucinations, socialization, and medication adherence. In another prior trial using this same algorithm in an earlier version of the CBT2go app, we found improvements in total symptoms in a large RCT of patients with schizophrenia and bipolar disorder [36]. Ben-Zeev and colleagues [50,51] also used a similar CBT-informed algorithm with the FOCUS app and added sleep interventions and on-demand educational components; they found reductions in several symptom domains in schizophrenia. Finally, Bucci et al [52] developed the Actassist app, which uses an algorithm based on the cognitive model of psychosis, and found large improvements in symptoms relative to symptom monitoring only in participants with early psychosis.

Related to this, recent meta-analyses have suggested that SST may be a more effective treatment for negative symptoms of schizophrenia than CBT [28,29]. The mCBTn intervention in this study did not include the SST or problem-solving components included in CBSSST. Thus, while SST may be an effective treatment for negative symptoms, this trial suggests that the CBT components of CBSSST targeting defeatist attitudes can improve experiential negative symptoms without SST.

With regard to secondary outcomes, modest improvements were found in positive symptoms and depression with a longer 24-week treatment period, which was not expected, given that participants were screened for severe positive symptoms and depression. Changes in functioning were mixed, with significant improvements found early in treatment but then dissipated with only trend-level improvements found overall on the A-QLS and SFS. The 24-week follow-up period may be too brief to expect meaningful changes in functioning.

This study had a high exclusion rate during the run-in period, with 36 out of 67 (54%) participants not meeting the strict, persistent, negative symptom entry criteria. A high screen failure rate during run-in periods is common in clinical trials with persistent, negative symptom entry criteria. A high screen failure rate of 44% was found in a psychosocial trial with similar persistent negative symptom criteria. For example, a screen failure rate of 44% was found in a psychosocial trial using similar criteria, except DPAS [26], and this rate is slightly higher than in pharmaceutical trials with similar criteria [26,53]. The proportion of patients with documented persistent negative symptoms (46%) is consistent with other estimates [3] and suggests negative symptoms are a common treatment need in patients with schizophrenia. Clinical trials of participants with persistent negative symptoms are rare, so this study makes an important contribution by demonstrating improvements in experiential negative symptoms as a primary target rather than secondary improvements related to changes in positive or depressive symptoms. It is important to note that participants were also recruited specifically for persistent experiential negative symptoms, which are more strongly linked to functional impairments than expressive symptoms [5,7,54,55].

Retention rates were excellent (81%-90% across assessments), especially for this negative symptom population, suggesting the intervention is feasible. It may be important that transportation was provided to therapy groups, which likely facilitated retention and may be necessary to maintain engagement of this population, especially in a large county with limited public transportation where this study was conducted. We have found much better retention in CBSSST trials when transportation was provided [56] than when it was not [57]. The CBT2go app may have also promoted engagement, for example, through daily reminders of personalized recovery goals. The mCBTn intervention focused on recovery goal work, which can improve motivation and promote engagement in psychiatric rehabilitation [58].

Engagement with the CBT2go app was mixed. The app was designed to promote engagement in recovery activities as often as every day. On average, however, participants responded to prompts to make an action plan for the day about one and a half times per week. While this proportion of days with completed action plans may seem low, practicing skills and completing homework assignments more than once per week is greater homework adherence than would be expected with CBT group therapy alone, where participants are typically expected to complete a single homework assignment per week. Homework adherence in CBT psychosocial interventions across multiple disorders is approximately 20%-56%. Thus, completing approximately one and a half action plans per week is better community engagement in recovery activities than might be expected in CBT therapy alone with one assignment per week. Greater engagement with the app was also associated with greater improvement in motivational negative symptoms and was unrelated to baseline severity of negative symptoms, suggesting the app played an important role in strengthening the treatment’s impact on this important outcome.

This trial had several limitations. First, as described above, further app development is needed to promote engagement (e.g., simplified interface and rewards and feedback to motivate). In addition, patients with greater severity of defeatist attitudes were recruited, because this was the target mechanism, so the findings may not generalize to patients whose experiential negative symptoms may not be related to defeatist attitudes. Participants were also excluded for severe positive symptoms or depression, so findings may not generalize to these populations. Finally, and importantly, this was a preliminary open trial that did not control for the effects of time, therapist contact, trips out of the home to come to group, socialization with staff and other patients, and other nonspecific factors. The next step is to complete an RCT with a contact control condition, which we are currently conducting. If this ongoing RCT with the modified CBT2go app confirms the findings of this open trial, this would provide stronger support for the defeatist attitude model of experiential negative symptoms and suggest that blended interventions like mCBTn can strengthen and shorten intensive psychosocial interventions for negative symptoms in schizophrenia.
Acknowledgments

We thank the participants who volunteered for this study. Research reported in this publication was supported by the National Institute of Mental Health (principal investigator EG; R61MH110019). This open trial was funded as part of the National Institute of Mental Health Experimental Therapeutics Program (RF-A-MH-18-704 R61/R33). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Department of Veterans Affairs or National Institutes of Health.

Conflicts of Interest

EG has an equity interest in Granholm Consulting, Inc, and may benefit from the research results as he receives income from the company for CBSST workshops and consulting. The terms of this arrangement have been reviewed and approved by the University of California, San Diego, in accordance with its conflict of interest policies. All other authors have no conflicts of interest to declare.

References


Abbreviations

ABS: Asocial Beliefs Scale
A-QLS: Abbreviated Quality of Life Scale
CAINS-MAP: Clinical Assessment Interview for Negative Symptoms Motivation and Pleasure
CBSS: cognitive behavioral social skills training
CBT: cognitive behavioral therapy
CDS: Calgary Depression Scale for Schizophrenia
CONSORT: Consolidated Standards of Reporting Trials
DPAS: Defeatist Performance Attitude Scale
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, fifth edition
EMA: ecological momentary assessment
HLM: hierarchical linear modeling
MATS: Mobile Assessment and Treatment of Schizophrenia
mCBTn: mobile-assisted cognitive behavioral therapy for negative symptoms
PANSS: Positive and Negative Syndrome Scale
PRIME: Personalized Real-time Intervention for Motivational Enhancement
RCT: randomized controlled trial
SAVVy: Smartphone-Assisted coping-focused interVention for Voices
SCID-5: Structured Clinical Interview for DSM-5
SFS: Social Functioning Scale
SST: social skills training
Usage and Acceptability of the iBobbly App: Pilot Trial for Suicide Prevention in Aboriginal and Torres Strait Islander Youth

Joseph Tighe1,2, PhD; Fiona Shand1, PhD; Kathy McKay3,4, PhD; Taylor-Jai Mcalister5, BPsys(Hons); Andrew Mackinnon1, PhD; Helen Christensen1, PhD

1Black Dog Institute, Sydney, Australia
2Men's Outreach Service, Broome, Australia
3Tavistock and Portman NHS Foundation Trust, London, United Kingdom
4Public Health, Policy and Systems, Institute of Population Health, University of Liverpool, Liverpool, United Kingdom
5Centre for Health Informatics, Australian Institute of Health Innovation, Macquarie University, Sydney, Australia

Corresponding Author:
Joseph Tighe, PhD
Black Dog Institute
Hospital Rd
Randwick
Sydney
Australia
Phone: 61 292824530
Email: joetighe@hotmail.com

Abstract

Background: The proliferation of mental health apps purporting to target and improve psychological wellbeing is ever-growing and also concerning: Few apps have been rigorously evaluated, and, indeed, the safety of the vast majority of them has not been determined. Over 10,000 self-help apps exist but most are not used much after being downloaded. Gathering and analyzing usage data and the acceptability of apps are critical to inform consumers, researchers, and app developers.

Objective: This paper presents pilot usage and acceptability data from the iBobbly suicide prevention app, an app distributed through a randomized controlled trial.

Methods: Aboriginal and Torres Strait Islander participants from the Kimberley region of Western Australia completed a survey measuring their technology use in general (n=13), and data on their experiences with and views of the iBobbly app were also collected in semistructured interviews (n=13) and thematically analyzed. Finally, engagement with the app, such as the number of sessions completed and time spent on various acceptance-based therapeutic activities, was analyzed (n=18). Both groups were participants in the iBobbly app pilot randomized controlled trial (n=61) completed in 2015.

Results: Regression analysis indicated that app use improved psychological outcomes, although only minimally, and effects were not significant. However, results of the thematic analysis indicated that the iBobbly app was deemed effective, acceptable, and culturally appropriate by those interviewed.

Conclusions: There is a scarcity of randomized controlled trials and eHealth interventions in Indigenous communities, while extremely high rates of psychological distress and suicide persist. In this environment, studies that can add evidence from mixed-methods approaches are important. While the regression analysis in this study did not indicate a significant effect of app use on psychological wellbeing, this was predictable considering the small sample size (n=18) and typically brief app use. The results on engagement with the iBobbly app were however positive. This study showed that Indigenous youth are early and frequent users of technology in general, and they regarded the iBobbly app to be culturally safe and of therapeutic value. Qualitative analyses demonstrated that iBobbly app use was associated with self-reported improvements in psychological wellbeing, mental health literacy, and reductions in shame. Importantly, participants reported that they would recommend other similar apps if available to their peers.

(JMIR Ment Health 2020;7(12):e14296) doi:10.2196/14296

KEYWORDS
mHealth; suicide; depression; eHealth; Indigenous; Aboriginal; First Nations; mental health; suicide ideation; apps
Introduction

Aboriginal and Torres Strait Islander (herein referred to as Indigenous Australian) suicide rates continue to be recorded at extraordinarily high levels, with rates of suicide among Indigenous Australians twice that of the non-Indigenous population [1]. This paper focuses on Indigenous youth, whose suicide rates are 4 times those of other Australian youth [1]. Indigenous suicide and suicidal behavior are different from those of non-Indigenous Australians and influenced by poverty and a range of historical, sociocultural, and sociopolitical factors [2,3]. Indeed, the terms “mental health/illness” and much of the western biomedical terminology around suicide are of questionable value to Indigenous health experts and community members, who view mental health through their holistic interconnected model of Social and Emotional Wellbeing (SEWB) [4,5]. This collectivist understanding of health care takes a community-wide view and values connection to family, culture, ancestry, land, and spirituality as some of the factors in maintaining wellness [4,5]. Disconnection, dispossession, and ongoing trauma since colonization have contributed to the severe health inequities Indigenous people suffer today [4,6,7]. These social determinants of health are contributing to the disturbingly high rates of Indigenous youth suicide evident over recent decades [4,6,7].

Rates of suicide among Indigenous people are even higher in regional and remote Australia where fewer health services are available [1]. Access to SEWB services deemed culturally safe is difficult in these vast and sparsely populated regions, although Indigenous-controlled health organizations attempt to address this [7,8]. Despite the remoteness of many communities, the vast majority have effective telecommunication systems including mobile phone coverage. Varying levels of trust of non-Indigenous health services and health researchers lead to challenges in establishing purposeful collaborations with Indigenous communities [9]. Effective research with Indigenous peoples requires significant investment in relationship building, community approvals, and additional ethical clearances [9].

Unsurprisingly, co-design and effective partnerships are recommended for progress [9-11], and cultural safety can be achieved through co-design and testing of SEWB interventions with end users [12-14]. Technology solutions are worth examining, as more Indigenous people are using smartphones than other Australians (70% vs 66%) and Indigenous youth are also using social media such as Facebook 20% more frequently than other Australian youth [15].

App or mobile health (mHealth) interventions have the potential to overcome some of the common barriers to help-seeking [14,16], as they are “always on” by design and available to hand without the need to attend face-to-face appointments even if the latter are available. Anonymity can be ensured through minimal sign-up, and autonomy and empowerment are supported as users choose to use the apps or not. mHealth and eHealth can serve as adjunctive or alternative treatment to traditional face-to-face services. Indigenous people have extensively reported feelings of “shame” in accessing some health services and in speaking with professionals about mental health issues [16,17]. Apps can attempt to bypass this “shame” in addition to overcoming any embarrassment or discomfort due to poor literacy through the inclusion of voiceover features [13].

The acceptability and usability of apps are critical if they are to effect change in mental health symptoms [18]. To create apps that are widely used, it is also important for researchers to understand the motivations for and pathways to participation, the capacity of apps to maintain users’ engagement, and the potential benefits of increased app use. Culturally appropriate interventions and SEWB services can assist in tackling Indigenous youth suicide [7,10,19]. The iBobbyl app was developed as one such intervention [13]. Having previously reported on the effectiveness of iBobbyl in reducing suicidal ideation, depression, distress, and impulsivity [12,13], the authors sought to understand the extent to which the app was used and to explore cultural appropriateness, acceptability, and therapeutic value. Featuring acceptance-based therapeutic activities and based on acceptance and commitment therapy, the app was designed and developed in partnership with local Indigenous stakeholders in the Kimberley region of north Western Australia and trialed following ethical approval between September 2013 and March 2015 [12,13]. In remote areas, it was designed to operate without internet connectivity [12,13]. This development involved the input of focus groups of Indigenous youth around content and the creation of original artwork or imagery and audio by Indigenous artists working closely with the app developers. As the first trial of an app to target suicide prevention in any population and considering the sensitivities around youth suicide, it was important for stakeholders to examine participant engagement. As a small qualitative component of the randomized controlled trial (RCT) itself, researchers had a rare opportunity to conduct interviews with a smaller group of participants. It was hoped that feedback regarding possible improvements to future versions of the app or to the research processes could be gathered. This paper addresses the following research questions:

- What was the nature of general technology use among this cohort of Indigenous youth?
- To what extent did participants find the iBobbyl app acceptable, culturally appropriate, and therapeutically effective?
- Was greater time spent on the app associated with improved therapeutic outcomes?

Methods

Participants

Participants (n=13) were purposefully selected due to their close proximity to the research team in Broome, Western Australia and their willingness to engage. These participants were part of the larger cohort (N=61) who participated in the iBobbyl app RCT in the wider region [12,13]. Indigenous people from 18 to 35 years of age were recruited if they recorded moderate scores on the depression or psychological distress, with or without current suicidal ideation [12,13]. While initially targeting just those aged 18-25 years (youth), the age range was widened to aid recruitment.
Data Collection

General Internet Use Survey

Data on general internet use were collected for 13 trial participants using The Measures of Technology Use survey developed by the Young and Well Cooperative Research Centre [20] (Multimedia Appendix 1). The survey contained 10 items covering frequency and duration of internet use in general, types of devices used online, and types of online activities engaged in. There were no adoptions made to the survey for the cohort interviewed.

Semistructured Interviews

Data were collected from semistructured interviews with the same 13 participants on their experience with using the app. The interview guide (Multimedia Appendix 2) consisted of 12 questions that were designed for a brief interview. The interview questions were reviewed and approved by Indigenous mental health professionals, and all participants were asked the same fixed, yet open-ended, questions. Of the 13 participants approached to be interviewed, all 13 agreed to take part. Interviews were not recorded, as recording Indigenous youth speaking about their own suicidality with an unfamiliar non-Indigenous researcher was regarded as potentially uncomfortable for participants.

Procedures

Three researchers (JT, FS, and HC) developed the interview guide relating to whether the app was deemed effective, culturally appropriate, and acceptable and recorded referral pathways and personal motivations around participation. As part of the consent process for the trial, participants were asked if they were willing to be followed up. This allowed us to follow them up with both standardized measures and to invite them to be part of the qualitative interviews. JT approached participants and interviewed those who provided consent to complete the interview and the aforementioned survey. The interview was conducted face-to-face following the 6-week app trial. All notes were hand-written verbatim, verified with participants to ensure accuracy, revised if needed, and transcribed. There were few edits made to the original notes following revision by participants. In order to maintain confidentiality, participants are referred to by their transcription code: age, gender, and transcription number. For example, 18-F-3 refers to an 18-year-old female who was the third person to be interviewed. After all data were collected, two researchers (JT, KM) independently analyzed the transcripts using an inductive approach appropriate given the exploratory nature of the study. These transcripts were read several times, with KM clarifying cultural or experiential uncertainties with JT, as he had conducted the interviews. A thematic analysis was conducted in line with the approach set out by Braun and Clarke [21], where the 2 researchers systematically coded the transcripts and sorted these codes into themes. Any disagreement was discussed and resolved. It was acknowledged that, while the transcripts were not always long, the focus of the research was clear and narrow, with dense sample specificity [22]. KM is also a highly experienced qualitative researcher and JT highly experienced in the subject area, so they were able to gain as much nuance as possible from the transcripts [22]. In addition, a third author, TM, an Indigenous community member, ensured all authors were representing the cultural aspects of the themes as accurately as possible.

Usage Data Automatically Downloaded From the Trial

The iBobby app was developed with the challenges of remote and very remote regions in mind [12-14]. Internet connectivity was not required to use the app, thereby enabling participants living remotely without a mobile phone network to participate. A Samsung Galaxy 7 tablet with the app pre-installed was provided to participants, removing the need to use a personal mobile phone. When connected to the internet — either during or after the trial — usage data were automatically sent to and stored in a database at the Black Dog Institute in Sydney. These data were available for 40 participants that participated in the RCT, while data were not collected for the remaining 21 participants for numerous reasons such as tablets with flat batteries or connectivity issues. Data were recorded on a number of domains visited: time spent logged onto the app (derived from each individual login and logout times), time spent on each of the 3 self-assessments, time spent on each of the 3 content modules, and number of times the emergency help button was pressed, which presented emergency telephone numbers.

SPSS version 25 was used to conduct regression analyses on 3 measures: suicidal ideation, depression, and psychological distress. Suicidal ideation was measured with the Depressive Symptom Inventory – Suicidality Subscale (DSI-SS), depression with the Patient Health Questionnaire 9 (PHQ-9), and psychological distress with the Kessler Psychological Distress Scale (K10). We sought to examine if time spent on the app led to improved scores on these outcomes.

Results

Of the 13 participants interviewed and surveyed, all were of Indigenous background, and 10 identified as female; they had a mean age of 24.15 years (SD 4.7 years). The age range was 19-29 years. More than half (7/13, 54%) were either in full-time or part-time study.

Survey Data: General Internet and Technology Use

All participants spent time online every day in a typical week, with a mean of 4.0 hours per day reported (range: 30 minutes to 10 hours). Less time was reported for weekend use, with an average of 3.36 hours (range: 0 minutes to 9 hours). The most popular times for online activity were the afternoon and evening (from 3 pm onwards), with the least activity occurring before midday. All participants went online after 11 pm at least 1 day per week, and 31% (4/13) went online after 11 pm at least 3 days per week.

Smartphones were by far the device of choice for online access, with 77% (10/13) using one daily. No participants used a desktop computer, 54% (7/13) used a tablet, 23% (3/13) used a laptop, and 15% (2/13) used gaming technologies or consoles to go online. The majority reported they would miss their smartphone more than any other technology if they no longer...
had access to it (9/13, 69%), compared to 23% (3/13) who reported that they would miss their television most. Responses by 85% (11/13) indicated that they accessed the internet via their smartphone or tablet. This compared to 15% (2/13) who accessed the internet at work or in a public place such as a library or internet cafe. The most popular was the use of social networking sites or apps such as Facebook (11/13, 85%). Of the respondents, 69% (9/13) used the internet for email, 61% (8/13) for watching or downloading video clips or movies, and 46% (6/13) to access health information or for gaming.

In relation to the iBobbly app trial, 77% (10/13) of participants interviewed were informed about the trial by a fellow Indigenous community member and 23% (3/13) by a non-Indigenous person (health worker/researcher). All 13 participants stated they would recommend the app to others, and 92% (12/13) stated that they would take part in a similar trial again.

**Usage Data**

**App Usage: Total Time**

The usage data of 40 participants were analyzed. The average total time spent using the app was 73.7 minutes, with a range between 8 minutes to 5 hours 42 minutes (over a 6-week period). Of the 15 participants who used the app for more than 1 hour, 4 participants used it for more than 2 hours in total, and 1 participant used it for more than 5 hours.

**App Usage: Number of Logins**

The average number of logins (visits to the app) among the 40 participants was 12.4, with a range between 1 and 43 (over a 6-week period). The average duration of a session using the app was 5.94 minutes. Of the participants for whom there were data, 85% (34/40) completed all 6 activities on the app.

**Regression Analyses**

Regression analyses were conducted using SPSS version 25 to examine a potential relationship between time spent on the app and improved outcomes (a dose-response relationship). There were no significant relationships between usage time and any of the 3 outcomes analyzed (suicidal ideation, depression, and psychological distress) for the intervention group (n=18); however, all associations were in a positive direction.

The primary outcome, suicidal ideation, was measured with the DSI-SS: The impact of usage time on reducing scores on this measure was minimal. For every minute spent on the app, DSI-SS scores would reduce by .013 points (R²=.29). This means that, for 2 hours spent on the app, DSI-SS scores could reduce by over 1.5 points. The DSI-SS records scores on a range of 0-12, with a score of zero indicating no suicidality and a score ≥2 indicating a risk of suicide in the general population.

High levels of both depression and psychological distress were recorded for this intervention group at baseline, as reported in the trial results [12,13]. Psychological distress was measured with the K10, and for every minute spent on the app, K10 scores would reduce by .007 of a point (R²=.35). In simpler terms, it would require 2.5 hours of app use to potentially result in a reduction of just of over 1 K10 point. The K10 records scores on a range of 10-50. Depression was measured using the PHQ-9, and although still in a positive direction, for this small sample, for every minute spent on the app, PHQ-9 scores would reduce by .001 points (R²=.268). The PHQ-9 records scores on a range of 0-27.

What the analyses did show was that participants with higher needs (higher baseline and follow-up scores on the psychological measures) used the app more frequently.

**Thematic Analysis**

The researchers were investigating to what extent participants found the iBobbly app acceptable, culturally appropriate, and therapeutically effective. Two researchers (KM, JT) identified and agreed upon 3 themes from the interview transcripts: (1) acceptability — being private and acceptable, (2) cultural appropriateness and future help-seeking, and (3) helping with feelings and creating distractions. These themes are explored in detail in the next sections.

**Acceptability — Being Private and Acceptable**

Perhaps not surprisingly given that this was an important part of the interview, exploration of acceptability was a dominant theme in the narratives. Many participants gave brief responses acknowledging their belief that the app was acceptable: “good” (19-F-1, 18-F-3, 29-F-5, 28-F-1, 23-F-13), “very good” (27-M-2, 19-M-6), “great” (21-F-9), or “fantastic” (33-F-7). In comparison, other participants gave longer responses indicating the app was acceptable because it provided a service that was accessible when they needed it:

> “Accessible when you feel like shit and you need to find a way to de-stress. It’s very helpful.” [27-F-11]

Indeed, accessibility was important when participants spoke about times when they would not have been able (either physically or psychologically) to access a professional health care service, even if they were willing to do so. Some may have felt known in a small community or simply hesitant to engage a service because they felt uncomfortable. The app allowed them a choice in health care that was previously unavailable:

> “...accessibility and private... Definitely overcomes the stigma of being known as a client/opting into a service provider... Not everyone is interested in seeing service providers... Encourages you to have a conversation with yourself.” [23-F-13]

Indeed, the privacy offered by the app was valued because it was “less worrying than actually talking to someone” [27-M-2]. The ability to interact with the app privately, without anyone else needing to be present, meant that youth who may have been reluctant or afraid to speak to family members or health care professionals in a face-to-face setting could still access support:

> “...because it’s shame to talk... embarrassed to ask for help. Family support not as strong as others. Can confide in iBobbly instead of family.” [27-F-11]

> “...it’s non-judgmental. Biggest stigma and shame is being judged by a counsellor or anyone. Can sit in a room on your own and do it.” [26-F-8]

However, participants also explored ideas around the app’s acceptability in terms of potential negative impacts. 26-F-8 felt...
that, after a person initially accessed support, the app “got repetitive.” In this way, there was a sense that the app was not harmful, despite its sensitive content, but that people “might get frustrated with it but not suicidal” (26-F-8). One way to potentially mitigate such frustration was to refine iBobbly so it was “…longer, have more content” [27-M-2]. Indeed, 23-F-13 highlighted an important consideration when creating an app that is accessible to vulnerable people who may not have access to any other support — an app may simply not be enough in the face of strong emotions:

If really remote [users] may feel isolated, may feel they have no one to confide in. Risk of invoking feelings, no follow up. [23-F-13]

Other participants shared similar concerns, although they were positive about the role of the app in the broader picture of mental health care:

But might over-rely on the app and not see people. Good, but just one of the steps to getting help. [27-M-2]

After using app, you are more comfortable answering questions, which may make it easier to talk face-to-face down the track. [28-F-10]

**Cultural Appropriateness and Future Help-Seeking**

Closely related to the theme of acceptability, participants’ references to their culture spoke to their views that the app was culturally appropriate. There was a sense that the app overcame the participants’ concerns about a lack of privacy or negative judgment that could often be connected to accessing traditional face-to-face services. In line with this, one participant spoke about “shame” connected to both feeling suicidal and seeking help. This term — which may encompass shame, embarrassment, and a fear of judgement — is often used by young Indigenous people to describe how they are feeling:

Young Aboriginal people have shame factor, so don’t have to talk to people, just use the app. [Might not want to talk to someone because of] privacy and confidence or feeling horrible about something. [19-F-1]

Indeed, there was a consistent belief that the app provided support that did not judge the person using it, which created a safe space:

Gives positive and clear way. Non-judgmental education. [33-F-7]

“The look and feel” of the app and its original multimedia creations were appreciated as culturally relevant and “right.” This created a safe space for Indigenous people to explore mental health coping strategies without the shame factor:

Really like the voice overs and story sharing... [27-M-2]

... down to earth ... right wording, all comes down to that and easy, straight-forward, says the right things, images right. [29-F-5]

Further, 25-F-4 compared the app favorably to a cognitive behavioral therapy activity she had previously undertaken:

App might be more attractive than a mood & thought table. [25-F-4]

Indeed, the images, videos, and voiceovers were not only culturally appropriate but also inclusive to participants’ other needs:

Visually good – I have dyslexia. [23-F-13]

In addition, the app could also create a bridge to future help-seeking. Some participants could see the benefit of using the app as an educational tool that improved mental health literacy, thereby facilitating improved cross-cultural communication with mental health professionals:

...Sometimes you can’t find the words. Hearing other stories, makes it easier to find those words. Sometimes you say nothing to clinicians, but the app educates [on the] language to use. [27-M-2]

... might work for them [other Indigenous youth] if other things are not working and they are getting frustrated trying to explain their situation and how they are feeling. Maybe use the app when trying to talk to professionals [bring app to appointments]. [27-M-2]

Finally, the broader research process was positively viewed by 1 participant who regarded the face-to-face interactions of the trial to be useful:

Actual process too, has been mini little counselling sessions for me too. Good for me to have that opportunity. [29-F-5]

**Helping With Feelings and Creating Distractions**

Participants’ narratives also spoke to the therapeutic effectiveness of using the app:

It should help — when I was doing it, it reduced all for me. Feelings came and went, but not strong ones. Calmer feeling, wasn’t thinking about bad things. [19-M-12]

The psychoeducational features of the app were highlighted and referenced, as they allowed participants to feel distracted from their thoughts and reduce their distress:

...good way to distract yourself. Educates you on where you are at... Better at understanding their feelings and breaking it down, rather than being all over the place. [18-F-3]

Was helping me a bit. Did help me get stuff off my mind. Where you sort the fruit out (this referred to a psycho educational activity included within the app). [19-M-12]

Other participants spoke about how the app was helping them to identify their potentially harmful behaviors such as suicidality and even at times, prevent them by changing their perspectives and demonstrating other coping strategies:

It did help me at times, in those moments, dark, suicidal moments, distract me...Kept going through the activities, till suicidal thoughts are changed to better ones. I continue to use it regularly. [29-F-5]
…helped me push away my suicidal tendencies. Not completely but helped you settle down and settle the thoughts down and let it out without actually speaking to someone. When using it, I was stressed out, ex-prison stuff. I just jumped on and got some form of release. Helped me from going off — I could feel I was going to lose it — so I would jump on iBobbly. Instead of impulsively drinking or being destructive or having depressing thoughts, I went onto iBobbly. [27-F-11]

In line with the idea of providing a distraction or a brief intervention, the app seemed able to change such perspectives and behaviors by providing an interruption for users to give them space to improve their mental state and decision making:

> It gives a buffer, has numbers, gives an opportunity to ask for help. [33-F-7]

> …instead of smashing a window or something, you can pick up the iBobbly or have a scream at it or something. Definitely, the suicidal thinking can calm someone down, gives options instead of suicide. [26-F-8]

There was also a sense that the app’s activities helped to improve self-awareness and interpersonal communications, which also spoke to ideas around improving cross-cultural communication:

> …even to recognize that you are not quite right … encourages you to have a conversation with yourself. Activities to improve behavior and interactions with other people. [23-F-13]

However, 23-F-13 wondered whether some of the activities were repetitive and suggested that people should be able to access more advanced tools when needed:

> Looks simplistic, may devalue the messages. Didn’t allow advancement to a more advanced set of questions. You need to challenge. Was basic, aimed at younger, repetitive. [23-F-13]

In this way, the app’s therapeutic effectiveness may have been influenced by the participants’ therapeutic needs at the time of use.

Finally, there was also an indication that the app was therapeutically effective because it could travel with the participants; it did not matter where they were when they needed support. In this way, the app could offer support that was both consistent and appropriate, something that would not have been otherwise accessible:

> If you use the app when distressed and impulsive; then, that reduces the depression and suicidal thinking. I took it to [location away from home]; shit things happened, and it helped. [19-F-1]

This may indicate the usefulness of the app for brief interventions rather than for intensive therapeutic use.

**Discussion**

**iBobbly Usage and Therapeutic Impact**

Participants used the iBobbly app for over 73 minutes on average, with over 12 logins on average over the 6-week trial. This is an encouraging indication of this app’s capacity to engage users considering the challenges in engaging users of mental health apps, as previously reviewed [23]. The intervention protocol did not include encouragement, support, or weekly reminders from a clinician, all previously reported to improve adherence in low-intensity eHealth interventions [23]. Rather, the iBobbly app was purely self-directed and relied solely on user motivation to logon and complete activities. It is questionable how much therapeutic value could be derived from 73 minutes of self-assessments and therapeutic activities; however, this usage is comparable to the effective ultra-brief acceptance and commitment therapy sessions of one or two 20-30-minute sessions delivered in primary care settings by Strosahl et al [24]. Further, 2 female participants (19-F-1 and 27-F-1) gave examples where the app helped during times of suicidal thinking. This demonstrates that even a short amount of time on the app can be effective when the content is appropriate and easily accessible.

Participants returned to reuse the app regularly, on average over 12 times, although it is difficult to determine the therapeutic effectiveness of each session or the cumulative effect of each session. Though participants completed identical activities, these activities allowed for choice, and repeat visits to the app highlight that it managed to maintain participant engagement. The average duration of a session using the app was quite brief, thereby indicating that the intervention at the very least provided some distraction from distressing thoughts, feelings, or activities. Participants found app use helped “push away my suicidal tendencies,” and it was a “good way to distract yourself. Educates you on where you are at.” These patterns of multiple brief interventions may suggest the usefulness of the app and similar interventions [25] as an adjunctive treatment rather than a standalone measure. This was reinforced in the interviews, with participants regarding the app as “one of the steps” to getting help. The app helped give participants language to their experiences, as well as identify more critically stressful times in their lives and coping strategies that could be effective. Consequently, the app was a tool that could be used during consultations with health professionals to improve mental health literacy and communication, thereby helping overcome shame. While it is likely that non-Indigenous people with mental health issues also suffer shame, the cultural aspects of racism in the health care system for a disadvantaged minority such as Indigenous Australians may exacerbate their feelings of shame.

The regression analyses completed on 3 outcome measures, suicidal ideation, depression, and distress, showed just a minimal indication that increased app use would lead to improved outcomes. With a small sample of 18 participants and usage data included from just the intervention group, the analyses lacked power and precision. The analyses did, however, show that participants with higher needs (higher baseline and follow-up scores on the psychological measures) used the app
more frequently. Their scores post-trial were reduced but still elevated, again suggesting the app’s suitability as an adjunctive treatment particularly for users with high therapeutic needs. The regression analyses contrast with the qualitative results and the significant reductions in depression and anxiety as reported in the RCT results; however, this may be explained by the previous use of a larger sample size of 61 participants [13,26]. Therapeutic effectiveness was identified as 1 of 3 themes in the thematic analysis, with participants considering the app to have multiple and diverse therapeutic features. Distraction was referred to by participants, aligning with the brief nature of their engagements. It is also possible they were not interested in longer sessions due to the “repetitive” nature of some activities or simply due to time restrictions, thereby engaging with the app in the brief and typical way young people may engage with any app. More enduring therapeutic value was also identified, such as perspective gaining and reduced impulsivity, calmer feelings, and improved interpersonal relationships. In summary, while symptom improvement was not captured by the impact of time spent on the app, the app was viewed as therapeutic, and this aligns with the previously reported significant reductions in depression and distress [13].

Acceptability and Cultural Appropriateness

The analyses indicate that participants viewed the iBobby app as an acceptable intervention for themselves and their community members. Indeed, a notable feature of this trial was a very low attrition rate of 3%, as reported previously [13]. This remarkably low dropout rate may indicate the acceptability of the app for the majority of participants, a view supported by the findings of the current study. Of interest was how participants were informed about the trial and whether being informed by other Indigenous community members would increase the likelihood of signing up. It is often assumed that local support for an initiative will improve adoption and that word-of-mouth is an effective method for Indigenous community members to share either positive or negative views on a health intervention. This was evident in this trial, with a large majority of participants (77%, 10/13) referred by other Indigenous community members acting as key influencers [13]. This demonstrates a positive view of the app’s acceptability and cultural appropriateness, as does the willingness of all participants in this analysis (n=13) to potentially take part in another similar trial in the future. In addition, all participants surveyed (n=13) were willing to recommend similar trials to others.

App trials have shown favorable results on adherence even above desktop-based interventions [27]. In addition, interventions delivered within the structure of RCTs rather than open access also improved adherence [28]. The iBobby intervention followed this pattern, with participants describing benefit from the flexible nature of using a mobile device as well as iBobby app adherence benefitting from the person-to-person follow-ups and rigor of an RCT. The 2015 study by Povey et al [14] evaluated 2 mHealth interventions for Indigenous community members: the iBobby app as detailed in this paper and the AimHi mHealth app designed for health workers. The study by Povey et al [14] found that eHealth interventions, such as the iBobby and AimHi apps, are more likely to be effective and acceptable for Indigenous communities if they are designed in partnership with community members and use appropriate language and imagery [14]. The iBobby participants’ recommendations for improving future versions of the app demonstrated their interest in helping to create optimal interventions in the suicide prevention space for their own people. This feedback, coupled with the positive comments around acceptability, indicated that Indigenous youth (under 25 years old) viewed the app as both acceptable and culturally appropriate.

Conclusion

Both RCTs and eHealth interventions are extremely rare in Indigenous communities, particularly when targeting suicidal ideation. Consequently, an accurate understanding of participants’ engagement with such interventions is crucial. This analysis showed Indigenous youth are heavy consumers of internet and communications technologies and that their engagement with the iBobby app demonstrated an interest in utilizing mHealth for suicide prevention and social and emotional wellbeing. Participants considered the iBobby app to have therapeutic value and to be an acceptable and culturally appropriate tool for themselves and their wider communities. Participants reported improvements in psychological wellbeing, increases in mental health literacy, and reductions in shame, which they attributed to the iBobby app use. It is most likely that effective research partnerships are key components in the successful implementation of mHealth apps such as iBobby for Indigenous community members. Further research where the iBobby app is available as a standalone app available for download will demonstrate its power to engage users in real-world settings.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Measures of Technology Use survey developed by the Young and Well Cooperative Research Centre. [DOCX File, 16 KB - mental_v7i12e14296_app1.docx]

Multimedia Appendix 2

iBobby semi-structured interview. [DOCX File, 13 KB - mental_v7i12e14296_app2.docx]
References


Abbreviations

DSI-SS: Depressive Symptom Inventory – Suicidality Subscale
K10: Kessler Psychological Distress Scale
mHealth: mobile health
PHQ-9: Patient Health Questionnaire 9
RCT: randomized controlled trial
SEWB: Social and Emotional Wellbeing

©Joseph Tighe, Fiona Shand, Kathy McKay, Taylor-Jai Mcalister, Andrew Mackinnon, Helen Christensen. Originally published in JMIR Mental Health (http://mental.jmir.org), 01.12.2020. 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 Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.
Visual and Personalized Quality of Life Assessment App for People With Severe Mental Health Problems: Qualitative Evaluation

David Buitenweg\textsuperscript{1,2}, MSc; Dike van de Mheen\textsuperscript{1}, PhD; Jean-Paul Grund\textsuperscript{3}, PhD; Hans van Oers\textsuperscript{1}, PhD; Chijs van Nieuwenhuizen\textsuperscript{1,2}, PhD

\textsuperscript{1}Tranzo Scientific Center for Care and Wellbeing, Tilburg School of Social and Behavioral Sciences, Tilburg University, Tilburg, Netherlands
\textsuperscript{2}Geestelijke Gezondheids Zorg Eindhoven Institute for Mental Health Care, Eindhoven, Netherlands
\textsuperscript{3}CVO Addiction Research Centre, Utrecht, Netherlands

Corresponding Author:  
David Buitenweg, MSc  
Geestelijke Gezondheids Zorg Eindhoven Institute for Mental Health Care  
Eindhoven, Netherlands  
Phone: 31 40 297 01 70  
Email: david.buitenweg@ggze.nl

Abstract

Background: QoL-ME is a digital visual personalized quality of life assessment app for people with severe mental health problems. Research reveals that e-mental health apps frequently suffer from low engagement and fall short of expectations regarding their impact on patients’ daily lives. Studies often indicate that e-mental health apps ought to respect the needs and preferences of end users to achieve optimal user engagement.

Objective: The aim of this study was to explore the experiences of users regarding the usability and functionality of QoL-ME and whether the app is actionable and beneficial for patients.

Methods: End users (n=8) of QoL-ME contributed to semistructured interviews. An interview guide was used to direct the interviews. All interviews were audiorecorded and transcribed verbatim. Transcriptions were analyzed and coded thematically.

Results: Analysis revealed 3 main themes: (1) benefit, (2) actionability, and (3) characteristics of the QoL-ME. The first theme reveals that the QoL-ME app was beneficial for the majority of respondents, primarily by prompting them to reflect on their quality of life. The current version is not yet actionable; the actionability of the QoL-ME app may be improved by enabling users to view their scores over time and by supplying practical advice for quality of life improvements. Overall, participants had positive experiences with the usability, design, and content of the app.

Conclusions: The QoL-ME app can be beneficial to users as it provides them with insight into their quality of life and elicits reflection. Incorporating more functionalities that facilitate self-management, such as advice and strategies for improving areas that are lacking, will likely make the app actionable. Patients positively regarded the usability, design, and contents of the QoL-ME app.

(JMIR Ment Health 2020;7(12):e19593) doi:10.2196/19593

KEYWORDS

quality of life; qualitative evaluation; visual assessment; e-mental health; assessment app

Introduction

Quality of life assessment in people with severe mental health problems faces several challenges. First, respondents may not have had the opportunity to develop the abilities necessary to engage in traditional language-based quality of life assessments [1-3]. Alternatively, comorbid intellectual disabilities [3-5] or psychopathology [6-8] may compromise the validity of quality of life results. Second, in mental health, quality of life is understood as an inherently subjective concept that is shaped by individuals’ values and preferences [9-11]. Research underlines this notion [12-14], which calls for further personalization of quality of life measurements. Third, quality of life assessment instruments may promote patient empowerment by providing patients with insight into their quality of life scores, which is an important prerequisite for shared decision making [15-17]. Both patient empowerment and shared decision making have become important goals in
mental health services [18,19]. To meet these 3 challenges, an innovative personalized visual quality of life assessment app was developed called QoL-ME [20]. The QoL-ME app consists of a core version that can be supplemented with additional modules. The core version involves a mandatory set of 3 universal quality of life domains. In addition, respondents can choose from 8 additional modules. Every module involves a domain of quality of life that respondents may select if it is important for their quality of life. Respondents only answer questions on their selection of additional modules. After filling out the questions, respondents receive direct feedback from the app in the form of an overview of their answers. The QoL-ME app was developed cocreatively in close collaboration with patients, family members, and care professionals [20,21].

Both research and practice reveal that e-mental health apps frequently suffer from low engagement and fall short of expectations regarding their impact on the daily lives of patients. [22-26]. Researchers have, therefore, investigated what factors enable e-mental health apps to bridge the gap from development to high engagement and practical use by patients [27-29]. Generally, these studies [27-30] often indicate that e-mental health apps ought to respect the needs and preferences of patients to achieve optimal user engagement, and 2 specific factors are of special importance. First, it is essential that the app is actionable. An app is actionable if provides a useful base for practical action for patients [31]. Examples of practical action include patients altering their sleep schedule after using an app that has sleep tracking functionality [25] or opting not to engage in a romantic relationship based on the results of a self-management app [26]. Second, use of the app ought to be beneficial to patients. An app should effectively address an issue that patients care about so that they derive a tangible benefit from utilizing the app [31].

End users played a vital role in the development of the QoL-ME app. In the context of this development, participants rated the usability of the app as “very high [20].” It is unknown, however, whether the intensive user-involvement and positive usability rating translate to an instrument that is of use to patients in real-life settings.

In light of the discrepancy between the potential of e-mental health apps and their lack of impact on patients’ daily lives, it is crucial to investigate the experiences of patients who used the QoL-ME app. In addition, it is of special importance to examine to what degree the QoL-ME app is actionable and beneficial to its users. The aim of this study was to explore the experiences of users regarding usability and functionality and whether the app was actionable and beneficial for patients. To this end, participants who had used the QoL-ME app were interviewed.

Methods

Participants

This study included 3 specific populations of people with severe mental health problems: people with psychiatric problems, people being treated in forensic psychiatry, and people who were experiencing homelessness. Individuals experiencing homelessness were included in this study because of the high prevalence of severe mental health problems in this group [3,32,33]. These groups may have difficulties with traditional language-based quality of life assessments because of fewer educational opportunities [1-3], co-occurring intellectual disabilities [3-5], and compromising psychopathology [6,7]. A consortium consisting of 6 societal institutions was formed to facilitate this study and the broader research project. These institutions included a multimodal day treatment center for multiproblem young adults, a hospital for forensic psychiatry, a mental health institution, a day center for people who are homeless, and 2 research institutions focusing on lifestyle, homelessness, and addiction. Participants were recruited with the help of the consortium partners.

The sample consisted of individuals who had gained experience with the QoL-ME app in the context of a psychometric evaluation of the app. In this psychometric evaluation, respondents were invited to use QoL-ME monthly for a period of 6 months. A specific inclusion criterion of at least 5 uses of QoL-ME was employed. This criterion ensured that patients had sufficient experience with QoL-ME to be able to contribute valuable information. The aim was to include enough participants to reach saturation in the sample, defined as a lack of new information in the final 2 interviews [34].

Ethical approval was obtained from the Ethics Committee of the Tilburg School of Behavioural and Social Sciences at Tilburg University (EC-2015.44). Informed consent was obtained from each participant. All procedures performed in this study involving human participants were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

The QoL-ME App

A group of 59 patients contributed to the development of the QoL-ME app. The iterative development comprised 6 iterations divided over 3 stages. In the first stage, patients were invited to share their ideas regarding design and functionality. In the second stage, initial designs and wireframes were developed into a fully functioning prototype. This process was guided by the feedback and ideas of patients. The prototype was subjected to a usability evaluation in the final stage [20].

QoL-ME encompasses 2 separate core versions. The first core version targets people with psychiatric problems and people treated in forensic psychiatry and includes 3 domains of the Lancashire Quality of Life Profile (LQoLP) [11]: safety, living situation, and finances. A recent study indicates that these 3 LQoLP domains are universal [12]. The LQoLP uses a 7-point Likert scale, ranging from 1 (cannot be worse) to 7 (cannot be better). The second core version is tailored to people who are homeless and comprises the Dutch version of the Meaning in Life Questionnaire (MLQ) [35], a 10-item measure that assesses both the presence of meaning in one’s life and the search for meaning in life. Research indicates that having meaning in life is especially important for people who are homeless [36,37]. The MLQ also uses a 7-point Likert scale, ranging from 1 (completely disagree) to 7 (completely agree).
The additional modules served to ensure the personalization of the QoL-ME app. The following 8 domains of quality of life were included: (1) support and attention, (2) social contacts, (3) happiness and love, (4) relaxation and harmony, (5) leisure, (6) lifestyle, (7) finances, and (8) health and living. These domains were identified in a visual concept mapping study of the quality of life of people with severe mental health problems [21].

Domains are assessed using 2 to 4 visual items. Every visual item contains 3 pictures that together denote an aspect of quality of life. Users respond to these items using a visual analog scale with visual anchors. Figure 1 depicts how respondents select additional modules and provides 2 examples of items in the additional modules.

In the QoL-ME app, users first indicate which of the 2 core versions is appropriate for them and respond to the items of that core version. Next, they select a combination of the 8 additional modules based on their importance. Upon completing the visual items of the additional modules, users are provided with an overview of their answers.

A thorough description of the development of the QoL-ME app, including additional visual material, is provided elsewhere [21].

**Figure 1.** Four screenshots depicting the additional modules of the QoL-ME. The top-left panel displays how respondents select additional modules. Respondents are invited to drag eight icons, corresponding to the eight modules, to a circle that says ‘important’ or a circle that says ‘not important’. The top-right panel shows how respondents are asked to confirm their choice of additional modules. The two bottom panels provide examples of items of the additional modules.

**Approach**

A qualitative research approach was employed to explore the participants’ experiences with the QoL-ME app. Specifically, individual semistructured interviews were utilized as they allowed participants to elaborate on their experiences and allowed the researcher to clarify any confusing or unclear questions when necessary. In addition, the context of individual interviews enabled reference to the QoL-ME app to make questions more tangible. The use of semistructured interviews combined a guiding structure providing participants freedom to expand on their answers.

**Content of the Interview**

An interview guide was used in this study. Four sources of information were consulted to inform this interview guide (Table 1). First, insights regarding patients’ needs and preferences concerning QoL-ME gained during development were fed back into the interview guide. Second, the Health Information Technology Acceptance Model (HITAM) was consulted [38]. The HITAM describes consumers’ behavioral intentions toward the use of health technology. Third, relevant information was extracted from 2 questionnaires designed to evaluate mobile health apps—the Mobile App Rating Scale (MARS) [39], and the App Chronic Disease Checklist (ACDC) [40]. Fourth, scientific literature was examined, and information regarding patients’ needs and preferences regarding mobile mental health apps was extracted [27,31,41,42]. The 18 topics were grouped into 5 overarching themes (see Table 1), and each theme was introduced using a short primer.
Table 1. Overview of the interview guide used in this study. The guide includes the different factors queried in this study, their origin, and the questions used to explore them.

<table>
<thead>
<tr>
<th>Topic</th>
<th>Question</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Deriving value</strong></td>
<td>Did using the QoL-ME benefit you? And if so, how? If not, what changes can we make for you to derive benefit from using the QoL-ME?</td>
<td>Development, HITAM&lt;sup&gt;a&lt;/sup&gt;, [31]</td>
</tr>
<tr>
<td></td>
<td>Did your use of the QoL-ME result in actions? If yes, which actions?</td>
<td></td>
</tr>
<tr>
<td><strong>Content and results</strong></td>
<td>What do you think about the number of questions in the QoL-ME?</td>
<td>Development</td>
</tr>
<tr>
<td>Number of questions</td>
<td>To what degree did the questions of the QoL-ME match your world and experiences?</td>
<td></td>
</tr>
<tr>
<td>Match questions and respondents</td>
<td>At the end of the QoL-ME, you can review your answers. What do you think about that?</td>
<td>Development, ACDC&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Feedback</td>
<td>Would you welcome the possibility to compare your own results with others and why?</td>
<td>Development</td>
</tr>
<tr>
<td>Comparing results</td>
<td>What do you think about the possibility to stimulate the use of an app such as the QoL-ME through push messages or other mechanisms?</td>
<td>Development, ACDC</td>
</tr>
<tr>
<td>Stimulation / motivation</td>
<td>What do you think about the QoL-ME’s usability? Are there any changes we can make to improve its usability? If yes, which changes?</td>
<td>Development, HITAM, ACDC, [31,41]</td>
</tr>
<tr>
<td>Usability</td>
<td>Does the QoL-ME have a clear structure according to you? Why/why not?</td>
<td>Development, ACDC, MARS&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>General usability</td>
<td>Did you have to learn or practice before using the QoL-ME? If yes, what did you have to learn or practice?</td>
<td>ACDC, MARS</td>
</tr>
<tr>
<td>Structure</td>
<td>What do you think about the appearance of the QoL-ME?</td>
<td>MARS</td>
</tr>
<tr>
<td>Intuitive design</td>
<td>Did you run into any problems using the QoL-ME on your phone/tablet/laptop/computer? If yes, which problems?</td>
<td>ACDC, MARS</td>
</tr>
<tr>
<td>Appearance</td>
<td>What did you think about selecting your own topics in the QoL-ME?</td>
<td>Development, ACDC, MARS</td>
</tr>
<tr>
<td>Performance</td>
<td>During the development of the QoL-ME, some participants indicated a preference for customizing the appearance of the QoL-ME. What do you think about that?</td>
<td>Development, ACDC, MARS, [27]</td>
</tr>
<tr>
<td>Personalization</td>
<td>Were you unable to use the app for any reason? If yes, what reasons?</td>
<td>HITAM</td>
</tr>
<tr>
<td>Personalized content</td>
<td>What did you think about the credibility of the QoL-ME?</td>
<td>ACDC, MARS</td>
</tr>
<tr>
<td>Personalized appearance</td>
<td>Do you think that your data is safe and confidential in the QoL-ME? Why?</td>
<td>Development, ACDC, [42]</td>
</tr>
<tr>
<td>Privacy/data security</td>
<td>Do you know which parties get to see your data and what they do with them?</td>
<td>Development, [27]</td>
</tr>
<tr>
<td>Transparency</td>
<td>What do you think about the appearance of the QoL-ME?</td>
<td></td>
</tr>
<tr>
<td>Professional credibility</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>HITAM: Health Information Technology Acceptance Model.

<sup>b</sup>ACDC: App Chronic Disease Checklist.

<sup>c</sup>MARS: Mobile App Rating Scale.

Data Analysis

A deductive, or theoretical [43], analysis approach was employed, starting from a specific predefined research question. All interviews were audiorecorded. The recordings were transcribed verbatim, and transcripts were coded thematically utilizing the 6-step method outlined by Braun and Clarke [43] in order to capture user experience themes. Initial themes were continuously refined and reflected on using a deductive approach. In step 1, the researchers familiarized themselves with the data through checking and verifying the accuracy of the transcripts. Step 2 involved the selection of an initial set of codes and themes based on the first 3 interviews. Codes were used to label and organize qualitative data. Codes with similar content were clustered into overarching themes. The coding was performed using ATLAS.ti (version 8, ATLAS.ti Scientific Software Development GmbH). The 2 researchers compared their initial codes to ensure consistency throughout the coding process.
process. Once the initial set of codes was confirmed, the researchers independently coded all of the interviews using the initial set. This set was modified or added to if necessary. Once all the interviews had been coded and the researchers reached consensus regarding the coding of the transcribed interviews, step 3 involved clustering of the codes into overarching themes. Themes were identified based on recurring codes. In step 4, the researchers discussed the themes and modified them, when required, to reach consensus on content and labeling. Step 5 encompassed interpreting and naming emerging themes. The results of the 6-step analysis method were reported in step 6 [43].

Procedure
Participants who contributed to the quantitative evaluation of the QoL-ME (DC Buitenweg, et al, unpublished data, 2020) were invited to participate in the interview. Participants who met the inclusion criteria were contacted via email, via care professionals at the consortium institutions, or via telephone if possible. Participants who expressed interest in contributing were provided with additional information on the qualitative study. Once a participant agreed to contribute, the researcher (DB) and participant scheduled an appointment for an interview. Interviews were held at the institution that supported the participant, or at a neutral location such as a café. Prior to the interview, the researcher provided a detailed explanation of the study and of what was expected of the participant. Moreover, the researcher explained that there were no right or wrong answers and that it was important that participants freely shared their opinions. Next, the researcher and participant went through the QoL-ME together to ensure that all participants had a refreshed understanding of the QoL-ME app. The interview guide (Table 1) steered the interview, while the interviewer elaborated on topics when necessary. Upon completing the interview, the interviewer explained how the data would be analyzed and how this aided the study. Participants were given time to ask any further questions. The interview ended when all questions were addressed whereupon the participant received a gift voucher. The duration of the interviews varied between 17 and 42 minutes and the average duration was 31 minutes.

Results

Participants
A group of 19 patients contributed to at least 5 assessments in the psychometric evaluation of the QoL-ME app. Of these 19 patients, 10 patients initially agreed to participate in an interview. The 9 patients who declined reported a lack of time or interest as their reason for declining to participate in the interviews. Of the 10 patients who initially agreed, one patient could no longer be reached and another was too busy to schedule an appointment. Therefore, 8 individuals with severe mental health problems participated in this study. We were unable to continue including participants until saturation because the number of experienced users who agreed to participate in the interviews was relatively low. Participants’ demographic characteristics are provided in Table 2. Five participants were male, the mean age of participants was 34 (SD 12 years), and 5 of the 8 participants had a Dutch cultural background. All participants had experienced using QoL-ME by contributing to the psychometric evaluation of QoL-ME. On average, participants had filled out QoL-ME 6 times (range 5-7) over a period between 4 and 6 months. Of 8 participants, 6 reported using QoL-ME on their personal smartphone, and the remaining 2 participants used their personal computer. Participants primarily used QoL-ME at home, while some reported using QoL-ME at their care institution.

Table 2. Demographic characteristics of the participants.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Age (years)</th>
<th>Gender</th>
<th>Cultural background</th>
<th>Level of education</th>
<th>Occupational status</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18</td>
<td>Male</td>
<td>Dutch</td>
<td>Basic</td>
<td>Paid employment</td>
</tr>
<tr>
<td>2</td>
<td>41</td>
<td>Male</td>
<td>Turkish</td>
<td>Basic</td>
<td>Volunteer work</td>
</tr>
<tr>
<td>3</td>
<td>39</td>
<td>Female</td>
<td>Dutch Antilles</td>
<td>Basic</td>
<td>Education</td>
</tr>
<tr>
<td>4</td>
<td>33</td>
<td>Male</td>
<td>Dutch</td>
<td>Basic</td>
<td>Unemployed</td>
</tr>
<tr>
<td>5</td>
<td>43</td>
<td>Female</td>
<td>Dutch</td>
<td>Basic</td>
<td>Volunteer work</td>
</tr>
<tr>
<td>6</td>
<td>27</td>
<td>Female</td>
<td>Dutch</td>
<td>Intermediate</td>
<td>Unemployed</td>
</tr>
<tr>
<td>7</td>
<td>52</td>
<td>Male</td>
<td>Dutch</td>
<td>Intermediate</td>
<td>Volunteer work</td>
</tr>
<tr>
<td>8</td>
<td>19</td>
<td>Male</td>
<td>Indonesian</td>
<td>Basic</td>
<td>Unemployed</td>
</tr>
</tbody>
</table>

Main Findings
The following 3 themes were identified based on analysis of the interviews: (1) benefit, (2) actionability, and (3) characteristics of QoL-ME. An overview of the codes and themes is provided in Multimedia Appendix 1 and includes both an overview in table form and a graphical depiction of the network of codes and themes. As the first 2 themes pertain to the 2 concepts (beneficial and actionable) that were of special interest in this study, these themes are discussed in more detail.

Benefit
According to 6 of the 8 interviewees, using QoL-ME was beneficial to them. All 6 of these participants mentioned that using the app made them more aware of their level of satisfaction on the life domains incorporated in the QoL-ME app.

Well, because of the questions that are asked, you start to think about what you do and don’t have. In principle, I am actually satisfied with everything. But
you are going to look at how you are doing. In your relationships, your family and your finances. [Participant 6]

For some participants, being confronted with their dissatisfaction on some domains drove them to look for ways to improve their situation.

The questions about income and whether you were satisfied with how much money you can spend made me think. When I have a job later on, I have more room for big expenses. So I started thinking about that. Yeah, that’s it, yes. [Participant 7]

For other participants, the QoL-ME app facilitated the realization that they were happier than they thought they were.

Ehmm. I started to think more consciously about how happy I actually was. And I turned out to be happier than I actually thought. [Participant 8]

The 2 participants for whom the QoL-ME app was not beneficial mentioned already having sufficient insight into how satisfied they were with their lives as the main reason for this lack of benefit:

No, no the questions that were asked, I already had some kind of insight in them. In those areas. So no I didn’t really get anything out of it. [Participant 5]

Both participants did feel that the QoL-ME app would be more beneficial to them if they lacked this insight:

[Interviewer:] And if you hadn’t known how you were doing in life?
[Participant:] Yes, if you don’t have that then you can discuss it with someone: oh, this is not going well so maybe I should do something with that. So then it would help. [Participant 1]

Actionability

For 3 participants, the QoL-ME app proved to provide a useful base for taking actions in their daily lives. One participant mentioned that using the QoL-ME app assisted her in maintaining of social relationships.

Well, for example I had not seen someone for a long time and I thought: let me call them. I tried to make contact. And you are also busy with your own life, I know, but I did think about that. [Participant 3]

Another participant spoke of being more careful in public transportation as a consequence of filling out the “Safety” domain:

[Interviewer:] And based on that, have you done something, changed something to what you normally do? For example in the area of personal safety?
[Participant:] Yes, subconsciously I did, because if I don’t feel safe and I don’t have to leave, then I stay inside. And for example if I travel by public transport and I see something strange then I get off. You start thinking more about these things. [Participant 2]

None of the participants reported discussing their QoL-ME results with others, but 2 participants acknowledged the possibility:

Then you have it right in front of you: things are not going so well. And then you can discuss that with someone. Okay, how are we going to improve this? [Participant 1]

Five participants reported not having taken any concrete action based on their experiences with QoL-ME. Two participants indicated that incorporating the option to compare current results with previous results would improve the actionability of QoL-ME.

[Participant:] What seems interesting to me is to see if your answers change over the different measurement moments.
[Interviewer:] Why is that interesting to you?
[Participant:] To see if it changes or if I am consistent. Because every day is different.
[Interviewer:] Yes, and if you could see that change, how would that affect how the App benefits you?
[Participant:] When I see that I am very satisfied with a certain topic one day and not at all the next, then I start to think ‘hmm, what is the reason for that?’. Where does that difference come from? And then it is also easier to do something with it. [Participant 4]

Regarding the potential negative effects of confronting users of the QoL-ME app with a decrease in their quality of life scores in the absence of care professionals, none of the participants expected this to be a problem.

Yes for some people you wouldn’t want to see that of course. But I feel like ... it's how you feel at the time. The situation may still be the same, but the way you deal with it may be different. You can feel different every day. [Participant 3]

Some participants provided tips for improving the clarity of the results section, which would also improve the actionability of QoL-ME but this is discussed under the third theme. One participant recommended including advice for how to improve low QoL-ME scores to improve its actionability. He used a food diary app as an example. Users register what they eat on a daily basis and the app generates an advice based on user input.

[Participant:] Yes, okay, so it really is for you... yes maybe you can generate an advice at the end of such a test. We see from your answers that you score negative on these topics and maybe you can think about that. Something like that.
[Interviewer:] Is that also a way to get more benefit from it?
[Participant:] Sure, I think so. That is ultimately what you want, a system that thinks along with you. I have an example, a silly example maybe, but I have an App from the nutrition center. This keeps track of exactly what you eat, and there is also advice. We see that you eat too much salt and too many unhealthy products. And then you are really triggered like I have
Characteristics of QoL-ME

Overall, participants welcomed the opportunity to view their results upon completion. Three participants provided specific advice for improving the clarity of the results section to increase the actionability of QoL-ME:

[Interviewer:] And the results you get to see at the end, did you think they are clearly displayed?
[Participant:] Ehm, I think in the second part, that you could add something like a number or something, I think.
[Interviewer:] Add a number or replace something with a number?
[Participant:] Add a number. So that you can see more clearly what it is ... or a percentage or something I am not sure. At least something that reflects it more clearly. [Participant 1]

Seven participants appreciated the possibility to personalize the content of QoL-ME. The one participant who disagreed indicated that he found all domains important and therefore preferred a version in which no choices had to be made. Participants were divided regarding the option to personalize the appearance of the QoL-ME app. Four participants welcomed this functionality, but the other participants thought it added too little value.

Several participants commented on the content of the QoL-ME app. One participant thought that the items on the financial situation of respondents were too direct and advised an alternative formulation. Four participants commented on the images used in the additional modules of the QoL-ME app. One participant recommended more variety (i.e., avoiding the use of similar pictures). Three participants reported that some of the images used were unclear to them. They advised including a written description of the content of the item using a word or a short sentence for clarification.

None of the participants had trouble with the duration of filling out or the number of questions. Three participants did miss a clear ending message, and they advised including this. One participant had issues with the low contrast between foreground and background elements due to her visual impairment. Seven participants thought the QoL-ME app looked professional, primarily due to its uncluttered and simple layout.

No participants reported having insight into which persons and parties had access to their data. Still, 6 participants trusted the security of their data. The inclusion of a disclaimer containing information regarding data access and use was a welcome addition for 7 participants.

In general, all participants were very positive regarding the design and usability of the QoL-ME app. Participants appreciated the clear structure of the app and favored the navigational system.

Discussion

Principal Findings

This study explored the experiences of users regarding the usability and functionality of the QoL-ME app and whether the app was actionable and beneficial for patients. As it is important that an e-mental health tool such as the QoL-ME app is both beneficial and actionable to its users, special attention was paid to these concepts. The interviews revealed that using the QoL-ME app is beneficial to most users, primarily by pushing them to consider their satisfaction in various life domains. The QoL-ME app did not prove to be actionable for most respondents. In addition, respondents were positive about the design and usability of the QoL-ME app but also had some tangible tips and advice for improvement.

The main way in which the QoL-ME app was beneficial to users was through providing insight and facilitating reflection. Some respondents indicated that their use of the QoL-ME app made them realize that they were more satisfied with their lives than they had expected. This result echo findings by Morton and colleagues [26] in their evaluation of a quality of life self-monitoring tool for people with bipolar disorder. Respondents also indicated that they were sometimes surprised by how high their scores were, which led to the insight that “things were not so bad.” Two participants indicated that they already had sufficient insight into their own quality of life and therefore derived no extra benefit from using the QoL-ME app. This finding echoed results found by Berry and colleagues [44], who investigated views on using digital self-management tools among people with severe mental health problems; a number of participants who contributed to that qualitative interview study indicated that they were already sufficiently self-aware and expected little benefit from using digital self-management tools [44].

Participants provided 3 useful suggestions for making the QoL-ME app more actionable. First, half of the participants proposed including numerical indicators of users’ satisfaction scores for every item or domain. The results section of the current version of the QoL-ME app does not include numbers but only shows a bar that is partly filled based on underlying scores. The Personal Health Information Self-Quantification System model [45] outlines how self-quantification is of vital importance for the self-management of health. In the model, self-quantification is described as the step in which an individuals’ goal (having a good quality of life) is transformed into objectively measured units [45]. Results from Morton and colleagues [26] confirm the importance of quantification, as respondents indicated that it was the quantification of their quality of life that enabled self-management. A second important suggestion for make the QoL-ME app more actionable, from 2 participants, was to incorporate practical advice for improving users’ satisfaction on certain life domains. The tool evaluated by Morton and colleagues [26] was integrated in a larger digital self-management platform that included practical advice and strategies for self-management. The results section of the tool provided direct links to these strategies, a feature that participants were very enthusiastic about [26]. Expanding the
QoL-ME app to include similar functionality will likely make the app more actionable for users. The third suggestion pertained to enabling users to follow the development of their quality of life scores over time. Every participant saw this as a welcome addition. This finding was in accordance with those of Morton and colleagues [26] and Berry and colleagues [44]. These 3 suggestions may be used to strongly improve how beneficial and actionable an assessment tool such as the QoL-ME app is to patients.

Several participants acknowledged the possibility of discussing the results of the QoL-ME app with other individuals such as a family member or professional caregiver. The fact that none of them did so may be an indication of social isolation, which has frequently been reported in this population [1-3]. Moving toward self-management, future versions of the QoL-ME app may actively encourage users to share their results and include practical suggestions for decreasing social isolation.

Participants were unanimously positive regarding QoL-ME’s usability. They found the application easy to use, appreciated its linear structure and prized the calm and clean layout. These results confirm the findings from the usability evaluation that made up the last part of the development of QoL-ME [20] and serve as additional corroboration of the design recommendations [41,46] consulted during the apps’ development. Several respondents preferred combining the visual material used in the additional modules with a word or short sentence to denote the content of its item. Comparable pictorial assessment instruments, such as the pictorial version of the Aachen Quality of Life Interview [47] and the pictorial motivation scale in physical activity [48] also combine both visual and verbal content. Respondents had very limited insight into which persons and parties had access to their data. This did not deter them from engaging with QoL-ME. This may be because respondents used QoL-ME in the context of a scientific study or because participating did not require respondents to share any personal information.

The results draw attention to several ways in which the QoL-ME app may be modified so that it is more beneficial for patients. Future research may further investigate what images used in the QoL-ME app are unclear and identify alternative images. Moreover, the results section of the app may be updated to display the development of results over time. In addition, following the example by Morton and colleagues [26], QoL-ME may be integrated into a larger self-management platform for people with severe mental health problems.

**Strengths and Limitations**

This study provides an important contribution to the field of e-mental health app development. The qualitative methodology provided patients with the opportunity to share their opinions regarding the usability and functionality of QoL-ME and to what degree the app was beneficial and actionable to them. The results draw attention to the fact that patients require functionalities that target their needs before an app becomes beneficial to them. Specifically, patients require functionality targeting self-management. In addition, the content of the interview was partially derived from existing frameworks that have proven to be effective for evaluating health apps [49].

Still, the results do need to be interpreted in light of 3 limitations involving the sample of participants who contributed to this study. The first limitation pertains to the size of the convenience sample used in this study. The eligible research population, based on the criterion of having completed at least 5 measurements, was small. Still, the results provide important insights into user experiences and the extent to which the QoL-ME app was beneficial and actionable for users. Once a larger group of patients starts using QoL-ME, additional research will have to reveal whether these results extend to the larger population. Analyses revealed that saturation, defined as a lack of new information in the final interviews, was not attained in the sample. The final 2 interviews did contain new information, but these were not substantial insights and no changes to the codes or themes were made based on these interviews.

The context in which participants gained experience with QoL-ME formed a second limitation. Participants were aware that they had used the QoL-ME in the context of a scientific study in which the psychometric quality of the QoL-ME was evaluated. Moreover, participants were incentivized to use the QoL-ME and to participate in the interviews. Therefore, their use of the QoL-ME may not represent use in a real-life setting, and their responses in the interviews may have been biased. To counter possible bias due to the incentives, the researcher indicated that respondents were allowed to freely give their opinions before the interviews started. Future research may investigate to what degree future results are consistent this study’s results when patients’ who use QoL-ME on their own accord are interviewed.

The third limitation pertains to the absence of data on participants’ medical background, such as psychiatric diagnoses or symptom severity. Still, all participants received care from the consortium institutions, and we can therefore be certain that they were part of the target population. Future research may investigate whether individuals with specific symptoms or diagnoses have differing experiences using QoL-ME.

**Conclusions**

The QoL-ME app can be beneficial to users as it provides them with helpful insight into their quality of life. Including added functionality in support of self-management, such as advice and potential strategies for improving quality of life domains with which app users are dissatisfied will likely make the QoL-ME app more actionable. Overall, the patients who were interviewed positively regarded the usability, functionality, and contents of the QoL-ME app.

**Acknowledgments**

This work is part of the research program—Quality of Life and Health (project number 319-20-005)—which is financed by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organization for Scientific Research).

https://mental.jmir.org/2020/12/e19593
Conflicts of Interest

None declared.

Multimedia Appendix 1
Overview of codes and themes identified in the qualitative analysis.

References


29. Chandrashekar P. Do mental health mobile apps work: evidence and recommendations for designing high-efficacy mental health mobile apps. Mhealth 2018;4(6) [FREE Full text] [doi: 10.21037/mhealth.2018.03.02] [Medline: 29682510]


Abbreviations

ACDC: App Chronic Disease Checklist
HITAM: Health Information Technology Acceptance Model
LQoLP: Lancashire Quality of Life Profile
MARS: Mobile App Rating Scale
MLQ: Meaning in Life Questionnaire

©David Buitenweg, Dike van de Mheen, Jean-Paul Grund, Hans van Oers, Chijs van Nieuwenhuizen. Originally published in JMIR Mental Health (http://mental.jmir.org), 03.12.2020. 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 Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.
Evaluation of a Mobile App to Enhance Relational Awareness and Change During Cognitive Analytic Therapy: Mixed Methods Case Series

Stephen Kellett¹, BSc, MSc, D Clin; Katherine Easton², BSc, MRes, PhD; Martin Cooper¹, BA, MSc; Abigail Millings², BSc, PhD; Melanie Simmonds-Buckley², BSc, PhD; Glenys Parry⁴, BA, PhD

¹Sheffield Health & Social Care NHS Foundation Trust, University of Sheffield, Sheffield, United Kingdom
²University of Sheffield, Sheffield, United Kingdom
³Sheffield Hallam University, Sheffield, United Kingdom
⁴Catalyse, Sheffield, United Kingdom

Corresponding Author:
Stephen Kellett, BSc, MSc, D Clin
Sheffield Health & Social Care NHS Foundation Trust
University of Sheffield
1 Vicar Lane
Sheffield, S1 2LT
United Kingdom
Phone: 44 114 222 6537
Email: s.kellett@sheffield.ac.uk

Abstract

Background: There has been a lack of technological innovation regarding improving the delivery of integrative psychotherapies. This project sought to evaluate an app designed to replace previous paper-based methods supporting relational awareness and change during cognitive analytic therapy (CAT).

Objective: We aimed to assess patients’ and therapists’ experience of using the technology (ie, the “CAT-App”) and to evaluate the relationship between app usage and clinical outcome.

Methods: The design was a mixed methods case series. Patients completed the Clinical Outcomes in Routine Evaluation-Outcome Measure pre- and post-CAT. Mood data plus the frequency and effectiveness of relational awareness and change were collected via the app. Therapists and patients were interviewed about their experiences using the app.

Results: Ten patients (treated by 3 therapists) were enrolled; seven completed treatment and 4 had a reliable improvement in their mental health. App usage and mood change did not differ according to clinical outcome, but there was a statistically significant difference in app usage between completers and dropouts. The qualitative themes described by the therapists were (1) the challenge of incorporating the technology into their clinical practice and (2) the barriers and benefits of the technology. Clients’ themes were (1) data protection, (2) motivation and engagement, and (3) restrictions versus flexibility.

Conclusions: The CAT-App is capable of supporting relational awareness and change and is an upgrade on older, paper-based formats. Further clinical evaluation is required.

(JMIR Ment Health 2020;7(12):e19888) doi:10.2196/19888

KEYWORDS
cognitive analytic therapy; case series; effectiveness; outcome; eHealth; app; awareness; mHealth; innovation; therapy
Introduction

Mental health disorders are the single largest cause of health-related economic burden worldwide [1], and, globally, there is mounting pressure on health care providers to ensure rapid access to effective, evidence-based, organizationally efficient, and cost-effective psychological interventions. Digital technology bridges the demand-supply gap by offering easily accessible, flexible, and personalized support [2], thereby creating the possibility to affect large-scale and low-cost change in public mental health [3]. The World Health Authority has therefore championed development and evaluation of electronic mental health care [4]. Electronic mental health services are defined by the provision of digital interventions via mobile apps/tablets and online, web-based programs [5], with interventions either being delivered as standalone technologies or being integrated into face-to-face therapy [6]. The technology allows symptom monitoring, provides psychoeducation, and promotes ongoing self-management strategies [7]. The technology also collects in situ assessments of mental health symptoms, which are an ecologically valid source of naturalistic research data [8]. Online, web-based programs have generated a large evidence base in support of their efficacy and effectiveness [9], while the evidence in support of mental health apps is still under development [10].

Evident enthusiasm for electronic mental health innovation is somewhat tempered by evidence that certain mental health conditions (eg, depression, paranoia, or psychosis) create problems with engagement and may weaken users’ trust in the technology itself [11]. “Technology push” also occurs when the commercial concerns of digital health care companies trump the wants/needs of patients and also challenges the values of clinicians [2]. The lack of sufficient depth of clinical and academic collaboration regarding electronic mental health innovation has been highlighted as a key feature of “technology push” [12]. The speed at which electronic mental health can be developed also threatens to ignore (or be ignorant of) robust methods for treatment development and associated evaluation [13]. The content of some electronic mental health apps has also been criticized for not being grounded in sufficient theory [14] and for the fact that the outcomes achieved during clinical trials are rarely replicated in routine service delivery settings [15]. Although a plethora of electronic mental health apps are readily available for the treatment of a variety of disorders and in a variety of contexts, questions can linger concerning a lack of evidence regarding feasibility, safety, clinical effectiveness, and efficacy [12]. The potential for technological innovations to outstrip the co-development of a sufficiently robust evidence base therefore risks a loss of confidence/trust from both clinicians and patients [16].

The electronic mental health field is currently also dominated by apps reflecting the changing methods of cognitive behavioral therapy (CBT), in part because of the ease with which those changing methods can be translated into app content [12]. For example, a recent review of CBT-based apps for the management of depression found that 31 apps were available [3]. However, the acceptability of the CBT approach is not universal and so uptake and dropout rates can be very variable [17]. There are no known examples of apps that support the work of integrative and psychodynamic psychotherapies. A widely practiced integrative psychotherapy is cognitive analytic therapy (CAT) [18]. CAT is a time-limited and relationally-driven psychotherapy [18,19], which has a generally high quality evidence base [20]. A recent meta-analysis showed that CAT had moderate-to-large effects on global functioning, interpersonal difficulties, and depression in practice-based studies, while during clinical trials the pooled effect size showed a small but significant treatment effect in favor of CAT over controls [21].

CAT is an integration of personal construct and object relations theory [22]. Internalized early object relations are termed “reciprocal roles” which influence/limit relational repertoires [23]. “Target problems” are the diagnosis/presenting problem reframed in relational themes, and “target problem procedures” describe the present day “traps” (ie, vicious circles), “snags” (ie, self-sabotage), and “dilemmas” (ie, “either/or” dilemmas) maintaining relational problems [18]. The procedural sequence object relations model anchors target problem procedures in reciprocal role activation [19]. CAT has been recently summarized into a competency framework [24]. CAT uses a 3-phase approach: (1) an initial “reformulation” stage during which target problems and target problem procedures are summarized in narrative and diagrammatic reformulations, (2) a middle “recognition” stage facilitating self and relational awareness, and (3) a final “revision” stage that is focused on change. Therapeutic change during CAT is founded on effective relational awareness [25].

The CAT-App had the aim of digitizing the in-session and between-session recognition and revision tasks that constitute the middle and final phases of the therapy [26]. Previously, within each session, the CAT therapist and patient reviewed and rated recognition and revision of target problems and target problem procedures on paper-based rating sheets. Patients were also previously provided with a variety of paper-based target problem procedure recognition tools for use between-sessions. These between-session tools have been criticized for potentially creating stigma or embarrassment due to the difficulty of completing the tools surreptitiously [27]. The CAT-App was a proposed improvement on between-session methods, as the technology provides an unobtrusive method of collecting ideographic data in real time, contains dynamic feedback on recognition and revision trajectories, contains narrative and diagrammatic reformulations, and also captures users’ current moods. The potential in-session time saving would be that CAT therapists would be able to open the app during sessions and review ongoing recognition and revision rather than making single ratings within the session. The overall aims of this study were to evaluate the relationship between app usage and clinical outcome and also to attain feedback from patients and therapists on their experiences of the technology.

Methods

Design, Service, Therapists, and Ethics

The design was a mixed methods case series evaluation, and ethical approval for the study was awarded from the University
of Sheffield Ethics Committee (ethics reference number: 012217). The study setting was a private sector CAT clinic based in the United Kingdom. This is a not-for-profit service that offers CAT to people that cannot or choose not to access CAT through their local National Health Service mental health services. The service is staffed by therapists that are either accredited CAT practitioners or CAT psychotherapists. All the therapists (N=3) were in monthly clinical supervision for their CAT work. All were experienced therapists as each had completed more than 11 post-qualification years of CAT practice. Two were female CAT practitioners and one was a male CAT psychotherapist. Therapists underwent a brief (2 hour) training session on how to use the technology and were given a user manual.

### Inclusion and Exclusion Criteria

Inclusion criteria were (1) willingness to engage in CAT, (2) owning a smart phone, and (3) willingness to use the app. Eligibility was not dependent on receiving therapy. Participants were not excluded if they had previously received other forms of psychotherapy. Participants were not required to have any previous experience of using health technology. The exclusion criteria were (1) currently misusing substances to a significant degree, (2) currently frequently self-harming, (3) high on-going risk of suicide, and (4) currently posing a risk to another person to a significant degree.

### Outcome Measure

The Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM) was used to evaluate clinical outcome, as this is a valid and reliable measure of psychological distress commonly used in psychotherapy outpatient clinics [28]. The measure has good concurrent [28] and discriminant validity [29], sound internal and test-retest reliabilities [28], and is sensitive to psychotherapeutic change [29]. The CORE-OM was completed at assessment and termination of CAT. Outcomes on the CORE-OM were evaluated regarding the degree and clinical significance of change. The degree of change was assessed with the reliable change index [30]. The reliable change index tests for the degree of change required to be considered reliable, rather than that expected to occur by chance. The pre-post total CORE-OM score needs to change by 5 or more to assign a reliable improvement outcome (or reliable deterioration if the pre-post score increases by 5 or more). Clinically significant change [30] occurs when the pre-post outcome shifts in classification from “caseness” to “non-caseness.” The clinical cut-off score on the CORE-OM is 10. Simultaneous reliable and clinically significant change is a credible index of “recovery” in routine practice [31]. The pre-post effect size for the case series was computed and interpreted using Cohen’s power primer, where d≥0.20 is a “small” effect, d≥0.50 is a “medium” effect, and d≥0.80 is a “large” effect [32].

### Qualitative Interviews

Semi-structured interviews (45-60 minutes) were conducted with patients and with therapists. Interviews took place within 1 month of treatment being completed. Interviews were recorded using an encrypted digital audio recorder and transcribed verbatim. Interviews were conducted at a site convenient to the therapist or patient. Patient interviews explored the experiences of using the app, perceived benefits and barriers to usage, acceptability, perceived impact on recognition and revision of target problems and target problem procedures, and tolerability/burden of the measures embedded in the app. Therapist interviews explored experiences of incorporating the technology within the care pathway, perceived benefits and barriers to use of the CAT-App, acceptability of the technology, perceived impact on recognition and revisions efforts, and burden/tolerability of the measures. Data were analyzed using thematic analysis [33].

### Treatment

CAT is a time-limited psychotherapy delivered in 8, 16, or 24 session contracts according to diagnosis and severity [18]. Sessions were weekly and lasted for 50-60 minutes. The CAT delivered contained the 3 stages consistent with the clinical model: (1) reformulation, (2) recognition, and (3) revision. Reformulation consisted of an assessment phase enabling a narrative reformulation naming target problems and their developmental origins, how the problem is maintained (ie, target problem procedures), hypotheses about the manner in which the participant might experience the help offered by the therapist, and finally acknowledging any issues concerning the ending of the therapy. Recognition was marked by methods to enhance self-awareness of problematic states/roles/procedures, via production of a sequential diagrammatic reformulation and associated relational awareness monitoring as between-session homework. Narrative and diagrammatic reformulations were made accessible to participants on the CAT-App. As such, participants had immediate access to their personal reformulations to support relational recognition efforts (which was not possible in the previous paper-based approach). Revision focused on application of change methods (“exits,” in the language of CAT) which were bespoke to the participant, their individual reformulation, and their zone of proximal development [34]. In keeping with CAT practice, changes were visually labelled as exits on sequential diagrammatic reformulations [18] and so further change-based sequential diagrammatic reformulations that also had exits added were uploaded to the app. In the final session, both patient and therapist produced and shared “goodbye letters” [18]. The function of these letters is to reflect on the ending of the therapy and what this means to the patient, name the dominant relational patterns that occurred within the therapeutic relationship, name abandonment feelings, mark progress, and identify relapse prevention strategies [19].

### The CAT-App: Description and Usage Data

The co-design process of the CAT-App has been previously described [35]. Participants were not advised nor prompted to use the app in terms of a specified frequency but, in keeping with the model, were encouraged simply to notice and record problematic roles and procedures. The app had the ability to store and display the individual narrative reformulation, sequential diagrammatic reformulation, and goodbye letters. Patients opened the app to rate the degree to which they were recognizing (0-100 scale from ineffective recognition to effective recognition) their individual target problems and target...
If the patient was in the latter revision stages of CAT, then they would also rate the effectiveness of the degree to which they have revised (ie, changed) the associated target problem and target problem procedure. Feedback on current and all previous attempts at recognition and revision were graphed in order to visibly provide and display feedback. Patients were also allowed to write electronic notes on their recognition and revision efforts. When patients rated their current mood on 0-10 slider scales, a graph plotted current rating against previous ratings.

The number of times that the participants opened the app and the clinical details of the target problems and target problem procedures were recorded. Current mood state (rated 0-10) was collected on the following dimensions: sad-happy, anxious-calm, and excited-bored (reverse scored). These items were chosen to measure the pleasure (valence) and arousal (activation) dimensions of core affect [36], specifically, pleasure (sad-happy), pleasant deactivation (anxious-calm), and pleasant activation (bored-excited). Mood scores were grouped into observation quartiles to track mood changes over the course of treatment. Mean mood scores, variability of mood scores, and app usage were compared for patients that met criteria for reliable change versus nonreliable change.

The results are presented in two parts to meet the objectives of the study. The first section presents the quantitative results in terms of describing the sample, dropout rate, target problem and target problem procedure recognition and revision ratings, the clinical outcomes, and relationship between app usage and both clinical outcome and mood change over treatment time. The second section concerns the qualitative results of therapists’ and patients’ experience of CAT-App usage.

Ten patients initially consented to the study and 3 (30%) dropped out early during treatment (<3 sessions). Therefore, the case series (ie, the computer sample) consisted of 7 patients, 4 male and 3 female. The mean age of the patient sample was 34.71 years (SD 7.18). The presenting problems were mixed anxiety and depression (n=5), borderline personality disorder (n=1), and narcissistic personality disorder (n=1). Treatment duration was 16 sessions for the mixed anxiety and depression cases and 24 sessions for the personality disorder cases and is consistent with session allocation within the CAT model [18]. Six of the 7 completers were “cases” at assessment on the CORE-OM. Table 1 describes the 21 target problems rated in the app. In terms of usage data, the completer sample used the technology on average 119.86 times (SD 97.98) ranging from 11-239 occurrences. The dropout sample used the technology on average 5.00 times (SD 2.71) ranging from 3-9 occurrences. There was a significant difference in usage between completers and dropouts (t=2.29, P=.04). Figure 1 describes a CORE-OM Jacobson plot of the clinical outcomes for the 7 completers. There was a significant reduction in psychological distress between start (M=13.02, SD 5.41) and completion (M=4.68, SD 3.19) of treatment (z=4.50, P<.01), with a large effect size (d+=1.51). One of the patients remained a “case” on the CORE-OM at termination; 4 had a reliable reduction in distress (3 of whom also were in the community norm on the CORE-OM by end of treatment). The “recovery rate” was therefore 3/7 (42.85%).

Table 1. The target problems of the participants and associated theoretical concepts.

<table>
<thead>
<tr>
<th>Study participant ID</th>
<th>Target problem 1</th>
<th>Target problem procedure</th>
<th>Target problem 2</th>
<th>Target problem procedure</th>
<th>Target problem 3</th>
<th>Target problem procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Poor pacing</td>
<td>Dilemma</td>
<td>Grandiosity</td>
<td>Dilemma</td>
<td>Excessive drinking</td>
<td>Trap</td>
</tr>
<tr>
<td>2</td>
<td>Self-critical</td>
<td>Trap</td>
<td>Poor pacing</td>
<td>Dilemma</td>
<td>People pleasing</td>
<td>Snag</td>
</tr>
<tr>
<td>3</td>
<td>Seeing self as weak</td>
<td>Trap</td>
<td>Anxiety</td>
<td>Dilemma</td>
<td>Self-critical</td>
<td>Trap</td>
</tr>
<tr>
<td>4</td>
<td>Poor self-care</td>
<td>Trap</td>
<td>Anxiety</td>
<td>Trap</td>
<td>Emotionally cutoff</td>
<td>Dilemma</td>
</tr>
<tr>
<td>5</td>
<td>Isolating self</td>
<td>Snag</td>
<td>Social anxiety</td>
<td>Trap</td>
<td>Procrastination</td>
<td>Trap</td>
</tr>
<tr>
<td>6</td>
<td>Performance anxiety</td>
<td>Trap</td>
<td>Anxiety</td>
<td>Trap</td>
<td>Feeling judged by others</td>
<td>Trap</td>
</tr>
<tr>
<td>7</td>
<td>Boom or bust</td>
<td>Dilemma</td>
<td>Untrusting of people</td>
<td>Snag</td>
<td>Emotional suppression</td>
<td>Trap</td>
</tr>
</tbody>
</table>

For example, for the patient with narcissistic personality disorder, the target problem was “how I feel about myself keeps flipping between bigging myself up or feeling like a loser” with the associated dilemma “Either I feel superior to other people and am a bit contemptuous of them or I feel they are looking down on me and I think I’m rubbish”. Please note that the exact wording of these examples have been altered to protect patient anonymity, but without altering the clinical meaning.

In terms of usage data, the completer sample used the technology on average 119.86 times (SD 97.98) ranging from 11-239 occurrences. The dropout sample used the technology on average 5.00 times (SD 2.71) ranging from 3-9 occurrences. There was a significant difference in usage between completers and dropouts (t=2.29, P=.04). Figure 1 describes a CORE-OM Jacobson plot of the clinical outcomes for the 7 completers. There was a significant reduction in psychological distress between start (M=13.02, SD 5.41) and completion (M=4.68, SD 3.19) of treatment (z=4.50, P<.01), with a large effect size (d+=1.51). One of the patients remained a “case” on the CORE-OM at termination; 4 had a reliable reduction in distress (3 of whom also were in the community norm on the CORE-OM by end of treatment). The “recovery rate” was therefore 3/7 (42.85%).
Figure 1. Jacobson plot of clinical outcomes. CAT: cognitive analytic therapy. CORE-OM: Clinical Outcomes in Routine Evaluation-Outcome Measure.

Table 2 compares the percentage of ineffective and effective recognition and revision ratings for the clinical outcome groups. The completer sample is split between those who had no reliable change on the CORE-OM (n=3; 246 app observations, range 11-235) versus those who had a reliable pre-post reduction on the CORE-OM (n=4; 572 app observations, range 51-239).

Usage (number of observations) did not significantly differ between patients who experienced reliable change (M=143, SD 82.82) and those that did not (M=89; SD 126.54; t(5)=0.689, P=.521). Table 3 presents the mean mood scores for the completer sample and then those patients with reliable change and those that did not change.

Table 2. Recognition by clinical outcome (N=7).

<table>
<thead>
<tr>
<th>Target Problem</th>
<th>No change (n=3 patients and 246 app observations)</th>
<th>Reliably change (n=4 patients and 572 app observations)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ineffective recognition rating, %</td>
<td>Effective recognition rating, %</td>
</tr>
<tr>
<td>1</td>
<td>59.3</td>
<td>40.7</td>
</tr>
<tr>
<td>2</td>
<td>50.6</td>
<td>41.6</td>
</tr>
<tr>
<td>3</td>
<td>40.8</td>
<td>59.2</td>
</tr>
</tbody>
</table>
Table 3. Mean mood scores (SD) by clinical outcome group.

<table>
<thead>
<tr>
<th>Current mood state</th>
<th>Overall value (N=7) (SD)</th>
<th>Reliable change value (n=4) (SD)</th>
<th>Nonreliable change value (n=3) (SD)</th>
<th>t-score (P value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxious-calm</td>
<td>3.44 (1.06)</td>
<td>3.11 (0.91)</td>
<td>4.10 (1.36)</td>
<td>1.099 (.334)</td>
</tr>
<tr>
<td>Bored-excited</td>
<td>6.28 (0.67)</td>
<td>6.47 (0.65)</td>
<td>5.90 (0.72)</td>
<td>0.993 (.377)</td>
</tr>
<tr>
<td>Sad-happy</td>
<td>3.65 (0.92)</td>
<td>3.41 (0.82)</td>
<td>4.15 (1.20)</td>
<td>0.928 (.406)</td>
</tr>
</tbody>
</table>

Mood scores did not significantly differ according to clinical outcome. Figure 2 plots the observation quartile mood scores throughout treatment for the clinical outcome groups. Levene’s tests and associated t-tests for each phase failed to identify any significant differences in mood means and variances (SDs) between the clinical outcome groups (ie, P>.05 in all 4 comparisons of means and variances on each mood scale).

Of the 7 patients, 6 were interviewed (ie, a single patient was uncontactable to interview) and all 3 therapists were interviewed. Two main themes emerged from therapists: (1) incorporating the technology into clinical practice and (2) the perceived barriers and benefits to the technology use. Table 4 provides the themes, subthemes, and example quotes from the therapists and patients.

Figure 2. CAT-app mood scores grouped by observation quartiles (error bars=SD).
### Therapists' experience

#### Incorporating the app into clinical practice

<table>
<thead>
<tr>
<th>Promise of the app</th>
<th>“Some clients write a lot, others forget the paper or leave papers at school or home, especially younger clients.” (ID1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using the app</td>
<td>“I introduced the app in the first session and then we would wait until session 5 before using it because it takes time to define the target problem procedures. I just cut and pasted from the letter straight into the app.” (ID3)</td>
</tr>
<tr>
<td>Take up by patients</td>
<td>“All the clients were digitally literate. There was no hand holding and I told them to get back to me if they had any issues.” (ID3)</td>
</tr>
</tbody>
</table>

#### Perceived threats and benefits

<table>
<thead>
<tr>
<th>Non-equivalence</th>
<th>“I didn’t want the app taking any decisions from therapists. I don’t want to be a virtual therapist.” (ID2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxiety regarding the app</td>
<td>“It felt like extra work [...] I was worried in case I uploaded the wrong information to the wrong account.” (ID2)</td>
</tr>
<tr>
<td>Supporting self-reflection</td>
<td>“The app allows the client to become more of an observer of their behaviors than does the paper version.” (ID1)</td>
</tr>
<tr>
<td>Anxiety regarding data protection</td>
<td>“We need to know exactly where the client’s data is stored and what happens to that data once the client is discharged, for example.” (ID2)</td>
</tr>
</tbody>
</table>

#### Patients' experience

#### Data protection and storage

<table>
<thead>
<tr>
<th>Trust</th>
<th>“I trusted the therapist and the app” (ID4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“I’m more bothered about bank details than emotional stuff” (ID9)</td>
</tr>
<tr>
<td>Normality of data sharing</td>
<td>“What’s the worst that could happen?” (ID4)</td>
</tr>
<tr>
<td></td>
<td>“I’m happy with the Ts &amp; Cs. I don’t really read them anyway” (ID5)</td>
</tr>
<tr>
<td>Server felt safe</td>
<td>“It’s better than paper. I can’t lose it” (data) (ID6)</td>
</tr>
<tr>
<td></td>
<td>“It’s good the data isn’t on my phone. It means I can only work online but at least nothing is saved on my phone” (ID 5)</td>
</tr>
</tbody>
</table>

#### Motivation and engagement

<table>
<thead>
<tr>
<th>Initial excitement</th>
<th>“I like the idea, I set it all up straight away” (ID4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“Initially I thought it was a great idea, better than using paper” (ID 7)</td>
</tr>
<tr>
<td></td>
<td>“I was excited to try, I’ll give these things a go” (ID 8)</td>
</tr>
<tr>
<td>Reminders</td>
<td>“I liked it but I’m not driven to use it, I needed more discipline. Maybe nudges as well as reminders would help” (ID4)</td>
</tr>
<tr>
<td></td>
<td>“You can set reminders to use it but it would be good to have thought provoking messages as well, relating to CAT” (8)</td>
</tr>
<tr>
<td>Habit forming</td>
<td>“I didn’t have access for 3 weeks when I was abroad. I got used to using paper and just stuck with that” (ID7)</td>
</tr>
<tr>
<td></td>
<td>“It made me more conscious of my targets in the day. Knowing I could just not something down or reflect. I found it useful” (ID5)</td>
</tr>
</tbody>
</table>

#### Restrictions and flexibility

| Online/offline access | “It’s good that data is stored on a server but that means I can only use it if I’m online. It needs a save and sync option” (ID5) |
A benefit of the technology identified by therapists was the app was easy and unobtrusive for patients to use while prompting them to engage in relational awareness tasks. The therapists made a range of design recommendations about future iterations of the app including ability to add more diagrammatic reformulations, ability to add more than 3 target problems, modification of the recognition scale to also include “not occurred” to avoid ambiguity, nudges/incentives for change, rewards to reinforce relational change, additional journaling functions, addition of CORE-OM outcome scores, needing to be available offline, and also design improvements to color and style.

All 6 of the patient participants reported that they were frequent users of “apps” relating to social media, health and wellness, games, banking, and information (news, weather). All of the patients stated that they found the CAT-App easy to access and use, with no setup challenges reported. With respect to data protection and storage, 3 patients commented on the fact that they had no issues with agreeing to data sharing and storage, as this is such a common requirement for app usage. No other patient mentioned data storage or protection.

One patient reported that they would like to add some security to accessing the CAT-App on their phone, in addition to the personal identification number, such as adding date or birth or a fingerprint access (ID5). The limited functionality of the sequential diagrammatic reformulations hindered the use of the app for some patients. However, the easy, 24-hour access availability of the app did support some patients in practicing relational awareness unobtrusively on their phones and in the moment. A difference emerged between patients who preferred free-text writing and journaling and those who preferred tracking and quantifying change, with the former disengaging earlier from using the app. With respect to personalization, the ability to customize scales and tracking within the app was a feature that all patients mentioned they would find appealing. In terms of future development of the technology, the need for motivation to use the app was an issue, with reminders and messages being thought to be of potential use.

### Discussion

This project concerned a pragmatic, real-world clinical case series evaluation of a new electronic mental health app to support the delivery of an integrative form of psychotherapy. The use of apps to support integrative psychotherapies is an innovation in the field. The app was designed to enable relational awareness and change, making it distinct in terms of aims and content to the plethora of CBT-based apps available [3]. The CAT-App particularly sought to improve and innovate on extant in and out of session paper-based tools. The recognition phase of CAT was particularly well suited to electronic mental health adaptation, due to the emphasis on building relational awareness through regular reflection on those roles and patterns maintaining contemporary distress. The content and processes of the app were therefore theoretically mapped onto the middle
and final phases of CAT. This was because some electronic mental health has been criticized for not considering the manner in which the technology can be well integrated into routine care pathways [26]. However, the dropout rate in this study (30%) was higher than that reported in traditional delivery of the therapy (15%) [21]. The feasibility of using the app in clinical practice was supported by the brief (ie, 2 hour) training that therapists undertook.

Dynamic and real-time comparisons of recognition, revision, and mood ratings were not possible when using the extant paper-based recognition methods [18], and so patients received enhanced feedback to support their relational recognition efforts when using the app. The technology improved the old, paper-based approaches through being unobtrusive, containing more feedback features, capturing mood on validated scales, and storing the personal reformulations securely in one place. This would therefore be an example of technology being able to improve and innovate on previously widely used/accepted clinical methods. If the aim of recognition is to enable “a reflective and observing self” [19] that can then direct effective change, then the technology would appear a promising clinical tool in the support of this endeavor.

The aim of the CAT-App was not to replace the therapist and the therapy, but rather enhance the relational awareness work of the patient, and this aim was echoed in the interviews conducted with therapists. Therefore, therapy was still necessary alongside patient’s engagement with the app, with the two working symbiotically. The CAT-App is therefore an example of “blended digital treatment” [37] rather than “standalone” electronic mental health, as the technology was embedded into the CAT clinical care pathway [38]. The promise of blended approaches is that technology increases treatment efficiency by completing some tasks normally completed by therapists [39]. The app offered a time saving ability to therapists by having the ratings precollected and analyzed, thus circumventing the need to produce ratings in sessions. Additionally, as the app contained highly idiographic information (ie, target problems, target problem procedures, narrative reformulation, sequential diagrammatic reformulations, and goodbye letters), this would be an example of personalized medicine [40].

The question of how, when, and where electronic mental health is best used during psychological interventions is particularly relevant with regards to the generation of “digital phenotypes” that can predict impending relapse, and so effectively step in and support early intervention [41]. The promise of electronic mental health is particularly highlighted in terms of its potential to increase access to mental health services, but if the technology is purely standalone, then there is a risk that patients feel that they are being offered a “down-graded version” of face-to-face psychological work. Electronic mental health app users often abandon the technology after a brief intense period of usage, due to data inputting burden, loss of interest, and the lack of inclusive features [42]. This may have been the case here because those patients that dropped out of therapy did not use the app to any significant degree and also dropped out in the early stages of therapy.

In the context of a relational therapy such as CAT, it has been previously noted that electronic mental health users do use relational concepts when describing the technology, such as being open with and forming a bond with the technology [27]. There is a need for the “therapeutic relationship” between user and technology to be effective [43], and there have been calls for more intensive and dynamic measurement of the “therapeutic relationship” during ehealth interventions [44]. It has been argued that developers should pay far greater attention to developing and maintaining the relationship the patient has with the technology and that apps should involve an initial “relationship building” aspect [27]. The current study has highlighted that encouragement and prompts to continue with app usage would be helpful. On the other hand, promoting overly close or enmeshed relationships with technology can have potential unintended negative effects. For example, it is apparent that technology use has expanded and proliferated in most people’s everyday life [45], running the risk of interfering and intruding into relationships (“technoference”). The unintentional shadow cast by electronic mental health is that accessing the technology may actually disrupt or interfere with close relationships.

In terms of study limitations, while the size of the sample in the case series was appropriate for an early practice-based study, it still represents a study limitation in terms of being a small sample size. For example, the large effect size (>1.5) needs to be interpreted with due caution, due to the unreliability of Cohen d in small samples [46]. All the inferential statistics reported therefore need to be treated with prudence due to the small sample size, particularly with respect to the dropout rate reported. There was an acknowledged convenience selection bias in the recruitment of patients that could have been corrected through random sampling of referrals. Similarly, potential participants feeling less confident with the offer of an app may have excluded themselves from the study [47]. Other sources of selection bias were the need to own a smartphone and willingness to use the app, all the patients being white adults and the study being conducted in private psychotherapy practice. These issues of context and selection bias mean that the results may not generalize to other populations including specific ethnic groups, children or the elderly, or public health settings. There was no control group using the traditional paper-based recognition tools of CAT to compare app outcomes against. Clinical outcome measurement was limited to single outcome measure and could have been broader. The app is currently only available in English and as CAT is now practiced internationally [21], sister versions in other languages need to be developed quickly [48]. The promise of standalone and blended electronic mental health is pronounced in developing countries where funding for mental health provision may be piecemeal [49].

The study did not contain a follow-up period and therefore the degree to which the app was used beyond the end of CAT was not recorded. Capturing app usage during follow-up would be a goal in future studies, particularly as structured follow-up support is integral to the CAT clinical model [19]. When services are commissioned only to deliver acute-phase treatment, then no follow-up support is made available in the continuation phase [50], and so the potential role of technology supporting

http://mental.jmir.org/2020/12/e19888/
patients during follow-up appears particularly important to consider. The competence with which therapy was delivered could have been assessed using the competence in CAT measure [51]. Using the app during therapy does not change the key competencies of the CAT model [24], and this needs to be evidenced in future studies. The lack of interviewing of patients that dropped out did not enable their particular difficulties with the app to be understood.

In terms of future directions, the CAT-App needs to go through another design and content iteration assessment in order to act on the feedback received. The study was conducted in private psychotherapy practice and so evaluation in public mental health services is now indicated. Further research is needed to understand how specific and different diagnoses affect app uptake, adherence, and outcomes. For example, patients in the midst of a major depressive episode may find it difficult to use technological support due to the well-evidenced problems with attention and concentration that can interfere with people’s ability to use or interact with technology [52]. CAT principles have been applied to a 6-session guided self-help intervention for delivery at step 2 of Improving Access to Psychological Therapies services in the UK [53], and finding ways in which the app could integrate with this psychoeducational approach would be particularly useful. Similarly, the utility of the app with patients with personality disorders could be evaluated as the recognition phase of CAT is prolonged with this patient group [54]. CAT is also delivered as a one session “personal reformulation” to health professionals to help develop better professional role repertoires, and a new version of the app needs to be developed to support personal reformulations. The gamification of electronic mental health offers promise in the design of innovative delivery platforms to children and young people [55].

In conclusion, this project aimed to use a mixed methods case series design to evaluate an app designed to map onto the theoretical stages and content of a widely practiced and integrative form of psychotherapy [21]. We have learnt from this study that the recognition and revision phases of CAT can be supported using mobile technology to support patients practicing relational awareness. Such awareness has previously been illustrated to be the plinth upon which change occurs [25]. Any efforts to make relational awareness tools more accessible and less obtrusive to patients between sessions and also more time efficient for therapists within sessions are at a premium. As the app provided detailed ongoing feedback absent from the previous paper-based methods, this represents a technology-enabled advance in clinical practice. Further larger and more controlled studies are now indicated, and it may be possible to progress onto a patient preference clinical trial in which app-assisted CAT is compared against routine CAT [56]. Although the app appears to hold clinical promise, future development and associated outcome studies are also clearly indicated.

Acknowledgments
The project was funded by a CATCH research grant that was match-funded by the University of Sheffield and Catayle UK.

Authors’ Contributions
MC developed the CAT-app; GP, KE, AM and SK designed the study; KE conducted the patient and therapist interviews; MSB conducted all statistical analyses; and SK wrote the initial manuscript. All the authors reviewed the final manuscript.

Conflicts of Interest
GP is an executive director of Catalyse CAT Ltd, a nonprofit organization which cofunded this research. All other authors declare no conflicts.

References


Please cite as:
Kellett S, Easton K, Cooper M, Millings A, Simmonds-Buckley M, Parry G
Evaluation of a Mobile App to Enhance Relational Awareness and Change During Cognitive Analytic Therapy: Mixed Methods Case Series
JMIR Ment Health 2020;7(12):e19888
URL: http://mental.jmir.org/2020/12/e19888/
doi:10.2196/19888
PMID:33337342
Changes in Stress, Anxiety, and Depression Levels of Subscribers to a Daily Supportive Text Message Program (Text4Hope) During the COVID-19 Pandemic: Cross-Sectional Survey Study

Vincent Israel Ouoku Agyapong, MD, PhD; Marianne Hrabok, PhD; Wesley Vuong, MPH; Reham Shalaby, MD; Jasmine Marie Noble, MSc; April Gusnowski, BA; Kelly J Mrklas, MSc; Daniel Li, MD; Liana Urichuk, PhD; Mark Snaterse, BA; Shireen Surood, PhD; Bo Cao, PhD; Xin-Min Li, MD, PhD; Russell Greiner, PhD; Andrew James Greenshaw, PhD

1Department of Psychiatry, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada
2University of Calgary, Calgary, AB, Canada
3Alberta Health Services, Edmonton, AB, Canada
4Provincial Clinical Excellence, Alberta Health Services, Calgary, AB, Canada
5Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

Corresponding Author:
Vincent Israel Ouoku Agyapong, MD, PhD
Department of Psychiatry
Faculty of Medicine and Dentistry
University of Alberta
1E1 Walter Mackenzie Health Sciences Centre
8440 112 St NW
Edmonton, AB, T6G 2B7
Canada
Phone: 1 780 215 7771
Email: agyapong@ualberta.ca

Abstract

Background: In addition to the obvious physical medical impact of COVID-19, the disease poses evident threats to people’s mental health, psychological safety, and well-being. Provision of support for these challenges is complicated by the high number of people requiring support and the need to maintain physical distancing. Text4Hope, a daily supportive SMS text messaging program, was launched in Canada to mitigate the negative mental health impacts of the pandemic among Canadians.

Objective: This paper describes the changes in the stress, anxiety, and depression levels of subscribers to the Text4Hope program after 6 weeks of exposure to daily supportive SMS text messages.

Methods: We used self-administered, empirically supported web-based questionnaires to assess the demographic and clinical characteristics of Text4Hope subscribers. Perceived stress, anxiety, and depression were measured with the 10-Item Perceived Stress Scale (PSS-10), the Generalized Anxiety Disorder–7 (GAD-7) scale, and the Patient Health Questionnaire–9 (PHQ-9) scale at baseline and sixth week time points. Moderate or high perceived stress, likely generalized anxiety disorder, and likely major depressive disorder were assessed using cutoff scores of ≥14 for the PSS-10, ≥10 for the GAD-7, and ≥10 for the PHQ-9, respectively. At 6 weeks into the program, 766 participants had completed the questionnaires at both time points.

Results: At the 6-week time point, there were statistically significant reductions in mean scores on the PSS-10 and GAD-7 scales but not on the PHQ-9 scale. Effect sizes were small overall. There were statistically significant reductions in the prevalence rates of moderate or high stress and likely generalized anxiety disorder but not likely major depressive disorder for the group that completed both the baseline and 6-week assessments. The largest reductions in mean scores and prevalence rates were for anxiety (18.7% and 13.5%, respectively).

Conclusions: Text4Hope is a convenient, cost-effective, and accessible means of implementing a population-level psychological intervention. This service demonstrated significant reductions in anxiety and stress levels during the COVID-19 pandemic and could be used as a population-level mental health intervention during natural disasters and other emergencies.

International Registered Report Identifier (IRRID): RR2-10.2196/19292
Introduction

COVID-19, an acute respiratory disease, was first reported in December 2019 in Wuhan, China. Since the outbreak was declared a pandemic by the World Health Organization [1,2], this disease has continued to have significant, unprecedented impacts on health and patterns of human life worldwide. These impacts include school and business closures as well as the ongoing psychological and social tolls of uncertainty, vigilance, and quarantine. In addition to the obvious physical medical impact of this disease [3,4], it poses evident threats to people's mental health, psychological safety, and well-being [5-7], particularly given the risk of recurrent outbreaks [8].

In multiple global jurisdictions, a series of mental health concerns have arisen, including increased stress, anxiety, depression, fear, insomnia, and obsessive-compulsive behaviors. Population-level studies have summarized these effects [9,10]. For example, in a study in China, over half of the respondents rated the psychological impact of COVID-19 as moderate or severe, with 29% reporting significant anxiety symptoms and 17% reporting significant depressive symptoms [11].

The emergence of mental health issues during the COVID-19 pandemic was not entirely unexpected. There have been reports of increases in stress symptoms, confusion, anger, anxiety, and depression [12-15] as well as in problematic drug and alcohol use [7] related to previous pandemics. Stressors include long quarantine durations, infection fears, frustration, boredom, inadequate supplies, inadequate information, financial loss, and stigma. Quarantine, in particular, is associated with a number of negative psychological and social effects (eg, posttraumatic stress, anger, fear, financial loss, and stigma) [12].

Although research has provided a description of the psychological impact of COVID-19 [16], the literature regarding interventions or guidelines for managing the mental health impacts of the virus is limited [17]. Countries that were impacted initially by the COVID-19 pandemic identified several problems that increased the difficulty of providing psychological interventions during the pandemic, including barriers to participation, limited efficiency of outreach, and limited capacity of frontline workers to provide support due to competing demands on their time and energy [18]. Provision of psychological support during this pandemic is further complicated by the high number of people requiring support and the need to maintain physical distancing.

The COVID-19 pandemic has further reinforced the need and urgency of transforming the delivery of mental health services [19] to include telehealth, text messaging, and other digital platforms. Mobile health technologies offer a unique and innovative solution in this context. More specifically, SMS text messaging via mobile phones offers a convenient, cost-effective, and accessible means of implementing population-level interventions. In Canada, almost 90% of residents own a smartphone [20]. Additionally, SMS text messaging is embedded in 98% of mobile phones [21]. Texting is free to the majority of end users, does not require technical skill to use, and is included in most mobile plans. SMS text messages are also cost-effective for providers [22].

Previous research examining the effectiveness of supportive text messages has demonstrated positive outcomes, including reduction of depressive symptoms and high user satisfaction [23-25]. For example, evaluation of Text4Mood, a text messaging intervention administered following large-scale forest fires in Fort McMurray, Alberta, found that supportive text messages helped subscribers feel more hopeful about managing issues (82%), in charge of managing their depression and anxiety (77%), and connected to a support system (75%); moreover, subscribers stated that the intervention improved their overall mental well-being (83%) [24]. Similar findings were observed in other studies, including Text4Baby, which sought to assist women by providing supportive and informative text messaging during pregnancy, and another text intervention aimed to support the mental health of impoverished women in Bangalore. Participants in both interventions indicated that receiving these text messages gave them a sense of reassurance and made them feel supported [26].

On March 23, 2020, Alberta Health Services, along with the coauthors of this paper, initiated Text4Hope, a 3-month-long, supportive daily text messaging program using principles of cognitive behavioral therapy (CBT), as an additional mental health support for people living in Alberta during the COVID-19 pandemic [27]. The messages ultimately seek to reduce or inhibit negative thought patterns while suggesting and reinforcing the use of healthy self-coping mechanisms. This program was intended to complement existing addiction and mental health services that individuals might be accessing at the time of participation.

This paper evaluates the impact of Text4Hope on measures of stress, anxiety, and depression symptoms and provides estimates of prevalence rates 6 weeks into the program.

Methods

The Text4Hope Program

This cross-sectional comparative study sought to assess the effectiveness of community implementation of a supportive SMS text message intervention program focused on reducing symptoms of stress, anxiety, and depression during the COVID-19 pandemic. The study protocol [28] was approved by the Research and Ethics Board of the University of Alberta (Pro00086163).

In the Text4Hope program [29], individuals self-subscribe to receive daily supportive SMS text messages for three months by texting the word “COVID19HOPE” to a short code number. The messages are aligned with a cognitive behavioral...
framework, with content written by mental health professionals and coauthors of this paper (VIOA, MH). The messages were uploaded to a web-based platform, which delivered messages at 9 AM each day. The first message welcomed subscribers to the service and invited them to complete a web-based baseline survey that captured demographic and clinical information. At 6 weeks, subscribers were invited again via a text message link to complete a web-based follow-up survey. At baseline and at 6 weeks, we collected clinical information on stress, anxiety, and depression about each subject based on the 10-Item Perceived Stress Scale (PSS-10) [30], Generalized Anxiety Disorder-7 (GAD-7) scale [31], and the Patient Health Questionnaire-9 (PHQ-9) [32], respectively.

**Data Collection**

We were able to cross-reference clinical and demographic responses from individuals by asking clients to enter the mobile number they used for Text4Hope at the baseline and 6-week time points. No incentives were offered to respondents. Participation in the program was voluntary, and completing the survey was not required to receive the supportive SMS text messages. Subscribers could opt out at any time by texting “STOP” to the same sort code number used to enroll in the program. Baseline data collection occurred between March 23 and 30, 2020, and the sixth week follow-up data were collected between May 3 and 11, 2020. Figure 1 presents the subscriber flowchart, which indicates the number of subscribers who completed the web-based surveys at each time point. Furthermore, on May 11, 2020, there were 45,775 subscribers to the Text4Hope program, of which 6178 subscribers opted out of the program, giving a dropout rate of 13.5%.

**Outcome Measures**

Primary outcomes included the mean differences in scores on the PSS-10, GAD-7, and PHQ-9 scales at the sixth week versus baseline and the changes in the prevalence rates of self-reported moderate or high stress, likely generalized anxiety disorder (GAD), and likely major depressive disorder (MDD) at the sixth week from baseline.

**Hypothesis**

In a sixth-week evaluation report, 77% of subscribers to the related Text4Mood program indicated that the daily supportive text messages helped them to manage their depression and anxiety [24], which informed our decision to evaluate the Text4Hope program at the sixth week and to determine if subscribers generally had reduced anxiety and depression. Furthermore, a randomized controlled trial of daily supportive text messaging resulted in close to 25% additional improvement in mood (measured by the Beck Depression Inventory [BDI]) in the intervention group compared to the control group [23]. On this basis, we hypothesized that the Text4Hope intervention would result in >25% reduction in mean scores and prevalence rates in all 3 factors, the PSS-10, GAD-7, and PHQ-9 scales, at the sixth week versus baseline.

**Sample Size Considerations**

With a projection that daily supportive text messages would result in a 25% reduction in mean PSS-10, GAD-7, and PHQ-9 scores at the sixth week from baseline, a population variance of 5.0 for each scale mean score, a one-sided significance level $\alpha=0.05$, and an acceptable difference between sample mean and population mean score for each scale of zero ($\mu - \mu_{0} = 0$), we estimated that a sample size of 686 would be sufficient to detect mean differences between the baseline and 6-week PSS-10, GAD-7, and PHQ-9 scores with a power of 80% ($\beta=0.2$).

**Statistical Analysis**

Data analysis was undertaken using SPSS for Windows version 26 (IBM Corporation) [33]. To assess the primary outcome measures for our intervention, we used the paired $t$ test to assess the mean difference between the mean PSS-10, GAD-7, and PHQ-9 scales at baseline and the sixth week for subscribers who completed the instruments at both time points. In addition, we used the chi-square test to compare prevalence rates for perceived stress, likely GAD, and likely MDD at baseline and the sixth week for subscribers who completed the instruments at both time points. Moderate or high perceived stress, likely GAD, and likely MDD were assessed using cutoff scores of $\geq 14$ on the PSS-10 [30], $\geq 10$ on the GAD-7 [31], and $\geq 10$ on the PHQ-9 [32], respectively. There was no imputation for
missing data, and the total numbers reported represent the total responses recorded for each variable.

**Results**

**Participant Demographics**

Of the 766 individuals who completed both the baseline and 6-week surveys, 73 (9.6%) identified as male, 678 (88.7%) identified as female, and 13 (1.7%) identified as other gender. Table 1 provides the distribution of the demographic characteristics by gender of subscribers who completed both the baseline and sixth week surveys. Table 1 summarizes the demographic characteristics of the respondents who completed both the baseline and 6-week surveys as n (%). The data presented in Table 1 suggest that the majority of the respondents were aged between 26 and 60 years (601/758, 79.3%), White (656/766, 85.9%), had a postsecondary education (678/764, 88.7%), were employed (555/764, 72.6%), were married, cohabiting, or partnered (507/763, 66.4%), and owned homes (521/760, 68.6%).

Table 1. Demographic characteristics of respondents who completed both surveys by identified gender (N=766), n (%). Note that some category totals do not sum to N due to incomplete data.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Male</th>
<th>Female</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤25</td>
<td>5 (6.9)</td>
<td>53 (7.9)</td>
<td>5 (38.5)</td>
<td>63 (8.3)</td>
</tr>
<tr>
<td>26-40</td>
<td>17 (23.6)</td>
<td>207 (30.8)</td>
<td>6 (46.2)</td>
<td>230 (30.3)</td>
</tr>
<tr>
<td>41-60</td>
<td>33 (45.8)</td>
<td>337 (50.1)</td>
<td>1 (7.7)</td>
<td>371 (48.9)</td>
</tr>
<tr>
<td>60</td>
<td>17 (23.6)</td>
<td>76 (11.3)</td>
<td>1 (7.7)</td>
<td>94 (12.4)</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>55 (75.3)</td>
<td>590 (87.0)</td>
<td>11 (84.6)</td>
<td>656 (85.9)</td>
</tr>
<tr>
<td>Indigenous</td>
<td>1 (1.4)</td>
<td>16 (2.4)</td>
<td>0 (0)</td>
<td>17 (2.2)</td>
</tr>
<tr>
<td>Asian</td>
<td>4 (5.5)</td>
<td>18 (2.7)</td>
<td>0 (0)</td>
<td>22 (2.9)</td>
</tr>
<tr>
<td>Other</td>
<td>13 (17.8)</td>
<td>54 (8.0)</td>
<td>2 (15.4)</td>
<td>69 (9.0)</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school diploma</td>
<td>5 (6.8)</td>
<td>14 (2.1)</td>
<td>1 (7.7)</td>
<td>20 (2.6)</td>
</tr>
<tr>
<td>High school diploma</td>
<td>7 (9.6)</td>
<td>53 (7.8)</td>
<td>1 (7.7)</td>
<td>61 (8.0)</td>
</tr>
<tr>
<td>Postsecondary education</td>
<td>61 (83.6)</td>
<td>606 (89.4)</td>
<td>11 (84.6)</td>
<td>678 (88.7)</td>
</tr>
<tr>
<td>Other education</td>
<td>0 (0)</td>
<td>5 (0.7)</td>
<td>0 (0)</td>
<td>5 (0.7)</td>
</tr>
<tr>
<td><strong>Employment status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>52 (71.2)</td>
<td>496 (73.2)</td>
<td>7 (53.8)</td>
<td>555 (72.6)</td>
</tr>
<tr>
<td>Unemployed</td>
<td>10 (13.7)</td>
<td>79 (11.7)</td>
<td>2 (15.4)</td>
<td>91 (11.9)</td>
</tr>
<tr>
<td>Retired</td>
<td>9 (12.3)</td>
<td>58 (8.6)</td>
<td>1 (7.7)</td>
<td>68 (8.9)</td>
</tr>
<tr>
<td>Student</td>
<td>2 (2.7)</td>
<td>33 (4.9)</td>
<td>3 (23.1)</td>
<td>38 (5.0)</td>
</tr>
<tr>
<td>Other</td>
<td>0 (0)</td>
<td>12 (1.8)</td>
<td>0 (0)</td>
<td>12 (1.6)</td>
</tr>
<tr>
<td><strong>Relationship status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married, cohabiting, or partner</td>
<td>49 (67.1)</td>
<td>452 (66.8)</td>
<td>6 (46.2)</td>
<td>507 (66.4)</td>
</tr>
<tr>
<td>Separated or divorced</td>
<td>6 (8.2)</td>
<td>64 (9.5)</td>
<td>1 (7.7)</td>
<td>71 (9.3)</td>
</tr>
<tr>
<td>Widowed</td>
<td>2 (2.7)</td>
<td>16 (2.4)</td>
<td>0 (0)</td>
<td>18 (2.4)</td>
</tr>
<tr>
<td>Single</td>
<td>15 (20.5)</td>
<td>137 (20.2)</td>
<td>6 (46.2)</td>
<td>158 (20.7)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (1.4)</td>
<td>8 (1.2)</td>
<td>0 (0)</td>
<td>9 (1.2)</td>
</tr>
<tr>
<td><strong>Housing status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Own a home</td>
<td>49 (67.1)</td>
<td>465 (69.0)</td>
<td>7 (53.8)</td>
<td>521 (68.6)</td>
</tr>
<tr>
<td>Living with family</td>
<td>7 (9.6)</td>
<td>56 (8.3)</td>
<td>3 (23.1)</td>
<td>66 (8.7)</td>
</tr>
<tr>
<td>Renting</td>
<td>16 (21.9)</td>
<td>147 (21.8)</td>
<td>2 (15.4)</td>
<td>165 (21.7)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (1.4)</td>
<td>6 (0.9)</td>
<td>1 (7.7)</td>
<td>8 (1.1)</td>
</tr>
</tbody>
</table>
Outcome Measures

Table 2 presents the changes in primary outcome measures after 6 weeks from baseline for subscribers who completed both the baseline and 6-week surveys. The data displayed in Table 2 indicate that for subscribers who completed both the baseline and 6-week surveys, the mean scores on the PSS-10 and GAD-7 scales were significantly lower at 6 weeks compared to the mean scores at baseline, suggesting improvement in stress and anxiety symptoms. The effect size as measured by Cohen $d$ was small (<0.5) for both the stress and anxiety scales.

There was a reduction in the mean score on the GAD-7 scale of 19.0% at the sixth week compared to the baseline scores. The reduction in the PSS-10 scores at six weeks compared to the baseline scores, although statistically significant, was much smaller (4.1%). There was no statistically significant within-subjects difference between the baseline and sixth week PHQ-9 mean scores ($P > .05$).

Table 2. Comparison of the baseline and 6-week mean scores on the PSS-10, GAD-7, and PHQ-9 scales for subscribers who completed both the baseline and sixth week surveys (N=766).

<table>
<thead>
<tr>
<th>Measure</th>
<th>Responses, n&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Scores</th>
<th>Change from baseline, %</th>
<th>Mean difference (95% CI)</th>
<th>$P$ value</th>
<th>$t$ value</th>
<th>Effect size (Cohen d)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Baseline score, mean (SD)</td>
<td>Six-week score, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSS-10&lt;sup&gt;b&lt;/sup&gt;</td>
<td>684</td>
<td>20.35 (6.7)</td>
<td>19.51 (7.0)</td>
<td>4.1</td>
<td>$-0.83$ (0.42 to 1.24)</td>
<td>$&lt;.001$</td>
<td>3.99</td>
</tr>
<tr>
<td>PHQ-9&lt;sup&gt;c&lt;/sup&gt;</td>
<td>630</td>
<td>8.94 (6.0)</td>
<td>8.74 (5.8)</td>
<td>2.2</td>
<td>$-0.20$ (–0.17 to 0.57)</td>
<td>.28</td>
<td>1.08</td>
</tr>
<tr>
<td>GAD-7&lt;sup&gt;d&lt;/sup&gt;</td>
<td>612</td>
<td>9.62 (5.6)</td>
<td>7.82 (5.2)</td>
<td>18.7</td>
<td>$-1.80$ (1.44 to 2.16)</td>
<td>$&lt;.001$</td>
<td>9.86</td>
</tr>
</tbody>
</table>

<sup>a</sup>Not all subscribers completed all three scales; therefore, n for each scale is less than the total N.
<sup>b</sup>PSS-10: 10-Item Perceived Stress Scale.
<sup>c</sup>PHQ-9: Patient Health Questionnaire-9.
<sup>d</sup>GAD-7: Generalized Anxiety Disorder–7.

Table 3 indicates that there were statistically significant reductions in the prevalence rates of moderate or high stress and likely GAD but not of likely MDD when comparing the baseline and 6-week assessments. The largest reduction in prevalence rates was for anxiety (13.5%).

To assess the generalizability of our data, based on the mental health burden in our baseline samples, we examined the clinical parameters between people who only responded to the baseline survey versus those who responded to both surveys (baseline and sixth week) (Table 4 and Table 5). No statistical difference was elicited between the two groups (all $P > .05$), suggesting that at baseline, the mental health burden was similar between our study sample and subscribers who did not complete the 6-week survey.

Table 3. Comparison of the baseline and 6-week prevalence of moderate or high stress, likely generalized anxiety disorder, and likely major depressive disorder.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Prevalence, n/total responses (%)</th>
<th>Change in prevalence rate (sixth week from baseline), %</th>
<th>$\chi^2$ (df)</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>Sixth week</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate or high stress&lt;sup&gt;a&lt;/sup&gt;</td>
<td>642/748 (85.8)</td>
<td>582/742 (80.4)</td>
<td>$-5.4$</td>
<td>7.78 (1)</td>
</tr>
<tr>
<td>Likely major depressive disorder&lt;sup&gt;b&lt;/sup&gt;</td>
<td>288/723 (39.8)</td>
<td>262/688 (38.1)</td>
<td>$-1.7$</td>
<td>0.46 (1)</td>
</tr>
<tr>
<td>Likely generalized anxiety disorder&lt;sup&gt;c&lt;/sup&gt;</td>
<td>326/712 (45.8)</td>
<td>220/682 (32.3)</td>
<td>$-13.5$</td>
<td>26.76 (1)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Assessed using a cutoff score of $\geq 14$ on the 10-Item Perceived Stress Scale.
<sup>b</sup>Assessed using a cutoff score of $\geq 10$ on the Patient Health Questionnaire-9.
<sup>c</sup>Assessed using a cutoff score of $\geq 10$ on the Generalized Anxiety Disorder–7.

Table 4. Comparison of the prevalence rates of moderate or high stress, likely generalized anxiety disorder, and likely major depressive disorder between subscribers who only completed the baseline survey and subscribers who completed both the baseline and 6-week surveys.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Prevalence rate at baseline, n/total responses (%)</th>
<th>$\chi^2$ (df)</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subscribers who completed the baseline assessment</td>
<td>Subscribers who completed both the baseline and 6-week assessments</td>
<td></td>
</tr>
<tr>
<td>Moderate or high stress</td>
<td>4065/4798 (84.7)</td>
<td>642/748 (85.8)</td>
<td>0.62 (1)</td>
</tr>
<tr>
<td>Likely major depressive disorder</td>
<td>1848/4447 (41.6)</td>
<td>288/723 (39.8)</td>
<td>0.76 (1)</td>
</tr>
<tr>
<td>Likely generalized anxiety disorder</td>
<td>2040/4364 (46.7)</td>
<td>326/712 (45.8)</td>
<td>0.23 (1)</td>
</tr>
</tbody>
</table>

Agyapong et al

http://mental.jmir.org/2020/12/e22423/
Table 5. Comparison of the mean scores on the PSS-10, GAD-7, and PHQ-9 scales between subscribers who only completed the baseline survey and subscribers who completed both the baseline and 6-week surveys.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Score at baseline, mean (SD)</th>
<th>Independent t test</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subscribers who completed the baseline and 6-week assessment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSS-10&lt;sup&gt;a&lt;/sup&gt;</td>
<td>20.55 (6.77)</td>
<td>20.30 (6.71)</td>
<td>0.96</td>
</tr>
<tr>
<td>PHQ-9&lt;sup&gt;b&lt;/sup&gt;</td>
<td>9.03 (6.22)</td>
<td>8.94 (6.0)</td>
<td>0.35</td>
</tr>
<tr>
<td>GAD-7&lt;sup&gt;c&lt;/sup&gt;</td>
<td>9.64 (5.93)</td>
<td>9.56 (5.65)</td>
<td>0.37</td>
</tr>
</tbody>
</table>

<sup>a</sup>PSS-10: 10-Item Perceived Stress Scale.
<sup>b</sup>PHQ-9: Patient Health Questionnaire-9.
<sup>c</sup>GAD-7: Generalized Anxiety Disorder-7.

Similarly, we examined the clinical parameters between subscribers who responded to the 6-week survey only and subscribers who responded to both surveys (Table 6 and Table 7). No statistical difference was elicited in prevalence of stress, anxiety, or depression symptoms between the two groups (P>.05), suggesting that after receiving the intervention for 6 weeks, the mental health burden was similar between our study sample and subscribers who only completed the 6-week survey.

Table 6. Comparison of the prevalence rates of moderate or high stress, likely generalized anxiety disorder, and likely major depressive disorder between subscribers who completed both the baseline and 6-week surveys and subscribers who only completed the 6-week survey.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Prevalence rate at sixth week, n/total responses (%)</th>
<th>χ² (df)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subscribers who completed the baseline assessment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate or high stress</td>
<td>1217/1518 (80.2)</td>
<td>0.01 (1)</td>
<td>.91</td>
</tr>
<tr>
<td>Likely major depressive disorder</td>
<td>483/1378 (35.1)</td>
<td>1.83 (1)</td>
<td>.18</td>
</tr>
<tr>
<td>Likely generalized anxiety disorder</td>
<td>430/1361 (31.6)</td>
<td>0.09 (1)</td>
<td>.76</td>
</tr>
</tbody>
</table>

Table 7. Comparison of the mean scores on the PSS-10, GAD-7, and PHQ-9 between subscribers who completed both the baseline and 6-week surveys and subscribers who only completed the 6-week survey.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Score at sixth week, mean (SD)</th>
<th>Independent t test</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subscribers who completed the sixth-week assessments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSS-14&lt;sup&gt;a&lt;/sup&gt;</td>
<td>19.36 (7.12)</td>
<td>19.44 (7.05)</td>
<td>−0.25</td>
</tr>
<tr>
<td>PHQ-9&lt;sup&gt;b&lt;/sup&gt;</td>
<td>8.20 (5.79)</td>
<td>8.69 (5.75)</td>
<td>−1.79</td>
</tr>
<tr>
<td>GAD-7&lt;sup&gt;c&lt;/sup&gt;</td>
<td>7.55 (5.40)</td>
<td>7.71 (5.21)</td>
<td>−0.66</td>
</tr>
</tbody>
</table>

<sup>a</sup>PSS-10: 10-Item Perceived Stress Scale.
<sup>b</sup>PHQ-9: Patient Health Questionnaire-9.
<sup>c</sup>GAD-7: Generalized Anxiety Disorder-7.

**Discussion**

**Principal Findings**

The Text4Hope program was provided as an intervention tool for the general population to support the mental well-being of individuals living in the Canadian province of Alberta during the global COVID-19 pandemic. Other technology-based interfaces have been deployed during the COVID-19 pandemic to track the disease spread in populations [34], to gather data related to the general knowledge, attitudes, and behavior of the public related to the pandemic [35,36], and to offer mental health support to the public during the pandemic [37-40]. To the best of our knowledge, this is the first study to assess the impact of a text-messaging intervention on self-reported symptoms of stress, anxiety, and depression experienced during the COVID-19 pandemic. Our study yielded interesting results regarding temporal changes in the self-reported severity and rates of symptomatology related to the three psychiatric health conditions under study. After receiving daily messages for 6 weeks, we observed significant reductions in the respondents’ mean scores on the GAD-7 (18.7%) and the PSS-10 (4.0%), suggesting that the program was effective in reducing anxiety and stress symptomatology in the respondents. There was no significant reduction in the mean PHQ-9 score at 6 weeks from baseline. In terms of prevalence rates, the largest significant
reduction in prevalence rate was for likely GAD (13.5%), followed by moderate or high stress (4.1%). Again, there was no significant reduction in the prevalence rate of likely MDD at 6 weeks from baseline.

The self-reported rates of anxiety symptoms in our study at baseline were higher than those reported in other studies [41,42]. However, the greatest improvement recorded after the provision of the Text4Hope program was for anxiety, with a 13.5% reduction in symptom prevalence rate and 19% improvement in GAD-7 score. Text4Hope achieved a small Cohen d effect size (0.4) in mitigating anxiety symptoms, which is comparable to the effect sizes found for an internet CBT program aimed at reducing anxiety symptoms [43]. Typically, interventions that do not include therapist support demonstrate lower effect size outcomes compared to those including therapists [44,45]. By contrast, our intervention reached thousands of individuals during the pandemic, and it aimed to provide a general population intervention rather than individual psychotherapy. The overall change in GAD-7 scores in our study (−1.8) appears to be consistent with the magnitude of score changes recorded after providing other remote health services. For example, adding a telephone service to computerized CBT in combating anxiety yielded a reduction of 1.18 in GAD-7 scores [46]. Again, the percentage of change in GAD-7 scores in our study after providing a daily text message for 6 weeks (19.0%) is consistent with the effect of medications on anxiety symptoms; in a very large randomized controlled trial in the United Kingdom, sertraline was evidently effective in reducing GAD-7 scores by 21% after 6 weeks, and this result was described as clinically important [47].

Our findings indicate that there was a modest effect of the program on improving stress symptoms, with greater benefit than other internet-based cognitive behavioral therapy (iCBT) platforms [48]. In a Japanese study, iCBT was used to alleviate anxiety, stress, and depressive symptoms in university students [48]. The most significant effect was mainly reported for anxiety, while stress symptoms did not show a difference between case and control group members after the intervention period [48].

There was no significant change between baseline and 6-week mean scores for likely MDD. Comparing our results with other remotely delivered health services yielded variable results. Two meta-analyses found that the effectiveness of iCBT programs, including MoodGYM, in mitigating depressive symptoms showed small effect sizes, especially in short-term assessments [43,44,49]. On the other hand, a significant improvement in BDI-II score with moderate effect size was observed at 3 months in patients with depression and comorbid alcohol use disorder who received supportive SMS text messages twice daily compared to the control group, who only received a thank-you SMS text message every fortnight [23]. The Text4Hope program was primarily designed as a health promotion tool to support the general population in Alberta during the COVID-19 pandemic and to combat potential stress and anxiety symptoms that are usually associated with epidemics or global crises. Our study participants were members of the general population rather than a patient sample, which may account for the observed differences in results. Furthermore, in a previous randomized trial [23], participants received the intervention for 12 weeks compared to 6 weeks in our study, which may account for the observed differences in the effects. Another study examined the effect of sertraline, an antidepressant of the selective serotonin reuptake inhibitor class, in ameliorating depressive symptoms; this study reported only a 5% relative reduction (95% CI 7%-15%; *P* = .41) in the mean PHQ-9 score at week 6 [47], which is not vastly different from the apparent 2% improvement observed with our Text4Hope intervention.

Three months after the launch of the Text4Hope program, the dropout rate was 13.5%. A high withdrawal rate is not uncommon for a texting service provided via SMS. When Bendsten and Bendsten [50] compared an SMS texting service to services provided via email, they found that people in the SMS group opted out at significantly higher rates than those in the email group (20.1% versus 5.2%, respectively). Additionally, in a review study on behavioral changing interventions provided via SMS text message, authors reported a wide range of withdrawal rates (0%-57%) among participants. [51]. They justified this result as being due to the untailored and unilateral nature of these texting programs, which may be less engaging and therefore may result in low retention rates [51].

### Limitations of the Study

Our study has several limitations. For ethical reasons, we lacked a comparative control group that did not receive the Text4Hope intervention during the same phase of the pandemic against which the recorded changes in stress, anxiety, and depression levels could be compared. It is therefore possible that the reductions in stress, anxiety, and depression levels are not all attributable to the Text4Hope intervention. Second, we relied on self-rated scales to assess stress, anxiety, and depression symptomatology, which could potentially overestimate the levels of these mental disorders when compared with prevalence rates that would have been obtained using structured clinical interviews with the *Diagnostic and Statistical Manual of Mental Disorders 5th Edition*. Third, our results may not be generalizable to the general population and are at risk of participation bias, where individuals with pre-existing mental health conditions are characteristically more inclined to enroll in the Text4Hope program compared with individuals with no pre-existing mental health disorders. We did not ask subscribers about pre-existing mental disorders, which would have helped to distinguish new symptoms from pre-existing ones but may have resulted in limited enrolment; in our experience, subtle changes in signup processes for subscribers can result in marked decreases in participation. Finally, the sample size of subscribers who completed both the baseline and 6-week assessments was rather small, and it is possible that other subscribers had variable changes in their mental health parameters from baseline to the sixth week. This limitation notwithstanding, our sample size was larger than the projected 686 subscribers needed to detect mean differences between the baseline and 6-week PSS-10, GAD-7, and PHQ-9 scores with a power of 80%. In addition, the mental health burden in our sample at baseline and at the sixth week were not significantly different from those of subscribers who only completed the baseline survey and the subscribers who completed the 6-week survey, respectively (Tables 4 and 5). Furthermore, the demographic characteristics

http://mental.jmir.org/2020/12/e22423/
of subscribers who completed both the baseline and 6-week surveys mirror those of all 8267 subscribers who completed the baseline survey by July 12, 2020 [52]. Specifically, the proportions of the various demographic characteristics in our sample compared to the proportions of the same demographics in the larger sample of 8267 subscribers was as follows: female gender, 88.7% versus 87.1%, respectively; White, 85.9% versus 82.3%, respectively; postsecondary education, 88.7% versus 85.2%, respectively; employed, 72.6% versus 73.3%, respectively; married, cohabiting, or partnered, 66.4% versus 71.1%, respectively; and homeowner, 68.6% versus 65.9%, respectively [52]. These proportions support the generalizability of our results to all subscribers. Future studies using this style of intervention could attempt to minimize attrition by offering incentives for participation or by sending messages encouraging people to continue subscribing.

Finally, the effect sizes in our study were relatively small, which may minimize the strength of the produced results. However, interventions that do not include therapists often report low effect sizes compared to those including therapists [44,45].

Conclusion
The Text4Hope program resulted in statistically significant reductions in mean scores on the PSS-10 and GAD-7 scales but not the PHQ-9 scale at the sixth week from baseline. The program also resulted in statistically significant reductions in subscribers’ prevalence rates of moderate or high stress and likely GAD but not of likely MDD. The largest reductions in the mean scores and prevalence rates were observed for anxiety symptoms. It should be noted this paper reports data from the midpoint of the Text4Hope program implementation, and the rates of change for outcomes for stress, anxiety, and depression may differ after the program ends at 3 months.

The relatively large improvements in anxiety symptoms achieved in our sample after 6 weeks of receiving the intervention during the COVID-19 pandemic suggest that the Text4Hope program is a useful intervention that can be deployed during natural and humanitarian disasters to support individuals at the population level. Over half of Canadians have reported that their mental health needs are not fully met [53]. A commonly reported reason is the cost of services [53]. As such, free mobile-based services such as Text4Hope can help address financial barriers. In the Canadian context and in other global contexts, these services can also be delivered remotely, which helps maintain essential physical distancing requirements during pandemics but also provides a means of access for those at rural or remote locations with little or no capacity for accessing mental health support services. We did not differentiate urban from rural or remote subscribers, as the program was offered to everyone in the province; however, regardless of geographical region, the pattern of urban, rural, and remote subscribers’ locations is of interest in relation to understanding the full value of SMS text messaging in this context. This is an additional step that we are considering in evaluating this program approach.

Acknowledgments
Support for the project was received from Alberta Health Services and the University of Alberta. This study was supported by grants from the Mental Health Foundation, the Edmonton and Calgary Community Foundations, the Edmonton Civic Employees Foundation, the Calgary Health Trust, the University Hospital Foundation, the Alberta Children’s Hospital Foundation, the Royal Alexandra Hospital Foundation, and the Alberta Cancer Foundation. The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Authors’ Contributions
VIOA conceived and designed the study, including the Text4Hope program. MH and RS drafted the initial manuscript with VIOA. AG, WV, and SS participated in data collection. All authors contributed to the study design and revised and approved the final draft of the manuscript.

Conflicts of Interest
None declared.

References


Abbreviations

BDI: Beck Depression Inventory
CBT: cognitive behavioral therapy
GAD: generalized anxiety disorder
GAD-7: Generalized Anxiety Disorder–7
iCBT: internet-based cognitive behavioral therapy
MDD: major depressive disorder
PHQ-9: Patient Health Questionnaire–9
PSS-10: 10-Item Perceived Stress Scale
Associations Among Internet Addiction, Genetic Polymorphisms, Family Functioning, and Psychopathological Risk: Cross-Sectional Exploratory Study

Luca Cerniglia1*, PhD; Silvia Cimino2*, PhD; Eleonora Marzilli2, MA; Esterina Pascale2, PhD; Renata Tambelli2, Prof Dr

1International Telematic University Uninettuno, Rome, Italy
2Sapienza - University of Rome, Rome, Italy
*these authors contributed equally

Corresponding Author:
Luca Cerniglia, PhD
International Telematic University Uninettuno
Corso Vittorio Emanuele II, 39
Rome, 00186
Italy
Phone: 39 06679201
Email: l.cerniglia@uninettunouniversity.net

Abstract

Background: International research has emphasized that youths are at higher risk for the onset of internet addiction (IA), but studies investigating biological, psychological, and social factors associated with this condition are limited.

Objective: This study aims to investigate the possible association between IA and genetic polymorphisms in monoamine oxidase A (MAO-A), serotonin-transporter (5-HTTPL), dopamine receptor (DRD4), and dopamine transporter (DAT1) genes by considering the role played by the perception of young adults in their family functioning and their depression, anxiety, and avoidant personality problems.

Methods: In a sample of 104 male and female young adults aged between 19 and 23 years (mean age 21.87, SD 2.29 years) recruited from universities in the central southern part of Italy, we addressed the presence of IA using the Young criteria of the IA test. Moreover, the perception of young adults of their family functioning and their psychopathological symptoms were assessed through the Family Assessment Device (FAD) and the Adult Self-Report, respectively.

Results: We found no significant association between IA and any genetic polymorphisms, neither among males or females. Young adults with IA reported significantly higher scores in the subscale of FAD affective responsiveness (AR; P=.01) and in depressive problems (P=.02), anxiety problems (P=.009), and avoidant personality problems (P=.003) than those in the control group. Results of mediation analyses showed a mediation role played by depressive symptoms (B=.99; 95% CI 0.22 to 1.97) and avoidant personality problems (B=1.09; 95% CI 0.32 to 2.05) of young adults on the relationship between the FAD, AR, and IA. Finally, this relationship was moderated by the genotype of the 5-HTTPL (P<.001), DAT1 (P<.001), and MAO-A (P<.001) genes in young adults.

Conclusions: This exploratory study supports the recent evidence on the mutual relationship among biological, individual, and social risk factors associated with IA in young adulthood. Our findings may have important clinical implications for the development of prevention and treatment programs.

(JMIR Ment Health 2020;7(12):e17341) doi:10.2196/17341

KEYWORDS
internet addiction; mobile phones; family functioning; depression; anxiety; avoidant personality; MAO-A; 5-HTTPL; DRD4; DAT1
**Introduction**

**Background**

Over the past decade, the diffusion and use of the internet has grown rapidly, especially among adolescents and young adults [1]. From a developmental perspective, young adults may face important challenges in relationships with their family, peers, and society, the outcomes of which can be mediated and moderated by their psychological and biological characteristics [2]. On the one hand, youths can benefit from the digital revolution, which allows them to instantly search for information and communicate with others around the world [3]. In addition, there is evidence that mobile health smartphone apps can prevent diseases and improve the quality of life in young people [4]. However, on the other hand, scientific literature indicates that young adults are at higher risk of using the internet in a maladaptive way [5,6], to the point of developing internet addiction (IA) symptoms [7]. However, to date, official diagnostic criteria have not been identified [8]; although, the revision of the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) [9] has included an internet-related condition—the internet gaming disorder in Section III. This condition has been defined as a behavioral addiction that eventually leads to loss of control over internet gaming and functional impairment [10]. Some authors have defined IA as uncontrollable and obsessive-compulsive use of the internet, preoccupations regarding computer or technological devices, inability to control their use, poor time management, craving, and interpersonal problems [11-14]. Epidemiological studies have shown a prevalence of IA ranging from 6% to 35% among young adults [15,16] and from 6% to 21% among adolescents [17]. Although the highest prevalence has been reported among males [18], recent evidence has underlined an increase among females, with no significant gender difference [19]. The Developmental Psychopathology theoretical framework [20] offers a valid model to conceptualize clinical and subclinical psychological difficulties in young adulthood (such as IA) as it considers the development as a result of mutual influences between individual inherited genetic vulnerabilities and the quality of social experiences [21,22], particularly within the family context [23].

**Genetic Influences on IA**

Genes involved in dopaminergic and serotonergic systems are the candidate genes most frequently associated with IA [24]. Dopamine (DA) is an important monoamine that regulates various neural mechanisms, such as cognitive memory and emotional activity [25], and represents the final pathway of the reward system [26]. The availability of DA is regulated primarily by its transporter (DAT), which recaps dopamine at the level of nerve synapses [27]. The 3’ untranslated region of the DAT1 gene contains a variable number of tandem repetitions (VNTRs) of 40 polymorphic base pairs, and the most frequent polymorphisms are 9 or 10 repetitions [28]. Neurobiological studies have shown a dysregulated dopamine transmission associated with IA [29,30], with a decreased level of expression of DAT in the striatum [31]. Another gene crucial to the dopaminergic system is the dopamine D4 receptor gene (DRD4), which is involved in a wide range of behavioral processes related to addiction, such as attention, motivation, and emotion [32]. This gene has a polymorphic 48-base pair VNTR in exon 3, and the most frequent polymorphisms are 4 (4R), 7 (7R), and 2 (2R) repeats [33]. Previous studies have underlined significant associations between the DRD4 4R genotype with increased attention [34,35] that, in turn, has been associated with a higher risk of IA [36]. Besides the role played by dopamine, it has been evidenced that serotonin may also be implicated in IA. Serotonin (5-HT) plays an important role in feelings of well-being, happiness, anxiety, stress susceptibility [37], and depression [38]. The serotonin-transporter-linked polymorphic region (5-HTTLPR) is crucial for the fine regulation of 5-HT [39]. This functional polymorphism results in 2 common alleles, the short (S) and the long (L) [40], and studies by Sun et al [36] and Lee et al [41] have reported significant association between IA and the S/S genotype. Other studies involving allelic variants in monoamine oxidase A (MAO-A) activity have provided further support for the role played by serotonin and dopamine, given that this gene encodes an enzyme involved in the degradation of DA and 5-HT. The MAO-A gene has a 30-base pair repeat in the promoter region that affects transcriptional efficiency [42]. The high-activity variant (MAOA-H; 3.5 or 4 repeats) has shown significant association with impulsive personality traits [43] and tobacco and cannabis use [44], but other studies have reported conflicting results, with the low-activity variant (MAOA-L; 2, 3, or 5 repeats) associated with disordered gambling [45,46]. However, considering the complexity of IA, to date there is a dearth of studies that have considered possible interactions between genetic vulnerability and the risk provided by the social environment [47], although recent evidence has highlighted the genetic moderation on the impact of the environment on other disorders and behaviors related to addiction [48].

**The Role of Family Functioning**

In this field, the quality of family functioning has been suggested to be one of the main environmental risk factors for the onset and maintenance of IA among young adults [49,50]. In particular, many studies have shown that the lack of parental emotional support and connectedness, and the poor quality of the relationships with parents, have an important influence on the risk taking and addictive behaviors of their offspring [51-53], including IA [54-56]. Moreover, it has been posited that the poor quality of family functioning is associated with a wide range of psychopathological problems [57-59], which in turn may lead to a higher risk for the onset of IA [60].

**Individual Psychopathological Symptoms and IA**

International research has widely demonstrated the comorbidity between IA and other psychological difficulties, both in terms of psychopathological symptoms and personality disorder [61,62]. Although several studies have shown that this relationship may be reciprocal and bidirectional [63,64], recent evidence has suggested that IA can be considered as the result of other psychopathological problems [20], resulting in a strategy to cope with psychological discomfort [2] and physical discomfort [65]. Depression and anxiety symptoms represent the 2 psychopathological areas most frequently associated with IA, both in clinical samples [66] and in the general population.
Moreover, significant associations with cluster C personality features have been evidenced [69], especially with traits of avoidant personality disorder [70].

This Study

From a biopsychosocial perspective [71], this exploratory study aims to evaluate the role played by biological, psychological, and social factors that may contribute to the etiology of IA in a sample of 104 young adults of the general population.

We hypothesized the presence of the following: (1) significant associations between IA and genetic polymorphisms in MAO-A (ie, MAOA-H genotype), 5-HTTPR (ie, 5/5 or 5/4 genotype), DRD4 (ie, 4R/4R genotype), and DAT1 (ie, 9-repeat genotype) genes, based on previous studies that have underlined the key role of these genes in the biochemistry of addiction disorder [45,72,73], including IA [36,41,74,75]; (2) significant differences in the quality of family functioning of young adults and psychological profiles (ie, depression and anxiety problems, avoidant personality problems) between youths with IA and the control group, as suggested by previous studies that have shown significant association between IA both with a poor quality of family functioning [49,50] and with psychopathological symptoms, especially in the areas of depression and anxiety symptoms [76,77] and traits of avoidant personality disorder [78]; (3) a mediation role played by the psychological profile of young adults on the relationship between family functioning and IA, based on studies that have shown a predictive effect of the quality of family functioning on the psychopathological profile of young adults [57,58] that, in turn, has been suggested to be a significant predictor of IA [70]; and (4) the moderator role played by MAO-A, 5-HTTPR, DRD4, and DAT1 genotypes on the relationship between family functioning and IA, as suggested by recent evidence that has underlined that individual genotype play an important role in moderating individual sensitivity to environmental events [79,80] and to the development of psychopathology related to addiction [48].

Methods

Recruitment

Over a period of 1 year, 150 young adults (65/150, 43.3% boys; 86/150, 56.7% girls) aged between 19 and 23 years were recruited for this study, in the center-south of Italy. All youths signed informed consent, in which the study was illustrated in detail. To examine the youths' genotype of the MAO-A, DAT1, 5-HTTPR, and DRD4 genes, we collected biological materials from human buccal swabs (the procedure is described below). In addition, the youths who agreed to participate in this study were administered self-report questionnaires (described below). This study was approved by the Ethical Committee of the Department of Dynamic and Clinical Psychology at Sapienza University of Rome, in accordance with the Declaration of Helsinki.

Procedure for Biological Sampling

Youths were assessed through buccal swabs (Isohelix Swab Pack) by a group of psychologists, specifically trained for the purposes of the study. Subjects were aware of not having to eat (including chewing gum, candy, etc), drink (except water), smoke, and brush their teeth for at least 1 hour before sampling. Epithelial cell samples were carefully collected through buccal swabs. The biological samplings were transported, slightly chilled by Normative ice (+4°C), to the laboratories of the co-author, EP, for further processing. After buccal swabs were gathered, young adults filled self-report questionnaires (described below).

Assessment of Internet Use or Abuse Among Young Adults

The Adult Self-Report (ASR) [83] is a self-report used to elicit information regarding psychological functioning. Items are assessed on a 3-point Likert scale (0=not true, 1=sometimes true, and 2=very often true). The aim of this study is to explore the role played by depression, anxiety, and avoidant personality problems of young adults on IA. Consequently, we used the scores of the following DSM-oriented scales: depressive problems, anxiety problems, and avoidant personality problems. Research has demonstrated good reliability and validity for the scales of the ASR [83]. In this study, the ASR showed good internal coherence (Cronbach α=0.73-0.84).

Assessment of Psychological Profiles of Young Adults

The Adult Self-Report (ASR) [83] is a self-report used to elicit information regarding psychological functioning. Items are assessed on a 3-point Likert scale (0=not true, 1=sometimes true, and 2=very often true). The aim of this study is to explore the role played by depression, anxiety, and avoidant personality problems of young adults on IA. Consequently, we used the scores of the following DSM-oriented scales: depressive problems, anxiety problems, and avoidant personality problems. Research has demonstrated good reliability and validity for the scales of the ASR [83]. In this study, the ASR showed good internal coherence (Cronbach α=0.73-0.84).

Assessment of Family Functioning of Young Adults

The Family Assessment Device (FAD) [84,85] is a self-report questionnaire that was developed to measure perceptions of family functioning. It is composed of 60 items evaluated on a 4-point scale (1=strongly agree and 4=strongly disagree), and measures the 6 dimensions of the McMaster Model of Family Functioning: (1) problem solving, which refers to the family’s ability to solve problems; (2) communication, which refers to whether communication in the family is clear and direct or vague and indirect; (3) roles, which addresses the issue of how roles and responsibilities are allocated among family members; (4) affective responsiveness (AR), which refers to the ability of the family members to respond to a range of situations with appropriate quality and amount of emotion; (5) affective involvement refers to how family members experience interest in and involvement with each other; and (6) behavioral control, which assesses whether the family has norms or standards governing individual behavior and responses to emergency situations. In all dimensions, higher scores represent less satisfaction with family functioning. The validity, reliability, and internal consistency of individual scales have been thoroughly investigated [86]. The internal consistency of the 6 subscales in this study was also adequate (Cronbach α=.77-.87).
**DNA Isolation and Genotyping**

Buccal cell DNA isolations were performed using the Buccal Prep Plus DNA isolation kit (Isohelix) according to the manufacturer’s instructions. The yield of DNA is usually between 3 and 10 µg. Allelic variants of the MAO-A, DAT1, 5-HTTPR, and DRD4 genes were identified. Genotypes were grouped into dominant, recessive, and additive genetic models, with the exception of MAO-A among males. The MAO-A gene is located on the X chromosome, so males have 1 allelic variant. Consequently, for males, the MAO-A gene was grouped based on the presence of the MAOA-H activity allele (ie, 4R) and the MAOA-L activity alleles (ie, 2R, 3R, and 5R). For females, the MAO-A gene was grouped based on the presence of 2 copies of the MAO-H allele, at least one copy of the MAO-H allele or 2 copies of the MAO-L allele. For the DRD4 gene, previous studies [87, 88] have emphasized that the 4R allele is the ancestral allele and has distinct functionality with regard to other alleles. Consequently, subjects were grouped based on the presence of 2 copies of the 4R allele, at least one copy of the 4R allele, and without the 4R allele. With regard to the 5-HTTPR genotype, we considered the presence of 2 copies of the S allele (ie, S/S), at least one copy of the S allele (ie, S/L), and 2 copies of the L allele (ie, L/L). Finally, for the DAT1 gene, groups were formed based on the presence of 2 copies of the 10R allele (ie, 10/10), at least one copy of the 10R allele (ie, 9/10), and 2 copies of the 9R allele.

**Statistical Analysis**

Preliminary analyses were performed using descriptive statistics (frequencies, percentages, and mean scores). Deviations from Hardy-Weinberg equilibrium were tested using the chi-square test, which allows the comparison of the expected frequency of the specific polymorphism with the observed one [89]. The Hardy-Weinberg equilibrium of the MAO-A among males was tested using the method proposed by Graffelman and Weir [90] for biallelic variants on the X chromosome, which considers both males and females. To verify the possible association between IA (IA group vs control group) and genetic polymorphisms in MAO-A, 5-HTTPR, DRD4, and DAT1 genes, considering the role played by gender, sex-stratified logistic-regression models were conducted. For each marker, we examined the dominant, recessive, and additive genetic models. An allelic model was used to examine the possible effect of MAO-A among males. Odds ratios and 95% CIs are presented. To verify the possible differences between the 2 groups on all the FAD subscales and on the ASR scores of the DSM-oriented subscales of depressive problems, anxiety problems, and avoidant personality problems, two tailed t test for independent samples was carried out. Based on preliminary analyses, a total sample bootstrap mediation analysis for simple and multiple mediation was then conducted to explore the possible mediation effects of the emotional-behavioral functioning of young adults on the relationship between the perception of young adults of their family functioning and their scores on IAT. The scores of the independent variable and of the mediators were standardized before performing the mediation analyses. Indirect (ie, mediating) effects were evaluated with 95% bias-corrected CI based on 5000 bootstrap samples. Finally, we tested whether the relationship between FAD and IAT could be moderated by genetic polymorphisms in the MAO-A, 5-HTTPR, DRD4, and DAT1 genes in young adults. For moderation analyses, only dominant genetic models were used. We standardized the score of the independent variable before performing the moderation analyses. All analyses were performed using IBM SPSS software 25.0. Mediation and moderation analyses were performed using PROCESS for SPSS [91].

**Results**

**Sample Characteristics**

In this study, from the total sample of 150 youths, 16.6% (25/150) of youths who did not complete the assessment procedure and 12.6% (19/150) of youths for whom it was not possible to collect biological samples were excluded. The final sample consisted of 104 young adults (mean age 21.87, SD 2.29 years; 50/104, 48.1% males). All youths recruited were university students. Most of the youths recruited for the study lived in families (101/104, 97.1%). Of the parents, 88.5% (92/104) were married or cohabiting, whereas 11.5% (12/104) were separated. On the basis of the IAT cut-off for Italian validation (≥40) [82], for aims 1 and 2 of this study, the total sample was divided into 2 subgroups: (1) IA group: comprising youths who reported excessive internet use (35/104, 33.6%) and (2) control group: comprising youths in complete control of their internet use (69/104, 66.4%).

The allele and genotype frequencies of the DRD4, MAO-A, 5-HTTPR, and DAT1 genes among males (Table 1) and females (Table 2) are reported in this study.

Genotypes were distributed according to the Hardy-Weinberg equilibrium among male (DRD4: $\chi^2_j=2.1$, $P=0.14$; 5HTTPR: $\chi^2_j=0.7$, $P=0.78$; DAT1: $\chi^2_j=1.2$, $P=0.26$) and female (DRD4: $\chi^2_j=0.4$, $P=0.83$; 5HTTPR: $\chi^2_j=1$, $P=0.65$; DAT1: $\chi^2_j=4$, $P=0.48$) youths, with the exception of MAO-A (overall: $\chi^2_j=7.8$, $P=0.05$; females: $\chi^2_j=7.1$, $P=0.007$).
Table 1. Allele and genotype frequencies of the dopamine D4 receptor, monoamine oxidase A, serotonin-transporter, and dopamine active transporter 1 genes among males.

<table>
<thead>
<tr>
<th>Genes</th>
<th>DRD4, n (%)</th>
<th>MAO-A, n (%)</th>
<th>5-HTTLPR, n (%)</th>
<th>DAT1, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Allele</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>5 (5)</td>
<td>0 (0)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>3 (3)</td>
<td>16 (32)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>4</td>
<td>72 (72)</td>
<td>33 (66)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>5</td>
<td>3 (3)</td>
<td>1 (2)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>6</td>
<td>1 (1)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>7</td>
<td>5 (5)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>8</td>
<td>4 (4)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>9</td>
<td>N/A</td>
<td>N/A</td>
<td>48 (48)</td>
<td>N/A</td>
</tr>
<tr>
<td>10</td>
<td>N/A</td>
<td>N/A</td>
<td>52 (52)</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Genotype</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2R/2R</td>
<td>1 (2)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2R/4R</td>
<td>8 (16)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2R/5R</td>
<td>1 (2)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2R/7R</td>
<td>1 (2)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3R/3R</td>
<td>0 (0)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3R/4R</td>
<td>2 (4)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3R/7R</td>
<td>0 (0)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3R/8R</td>
<td>1 (2)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>4R/4R</td>
<td>28 (56)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>4R/5R</td>
<td>2 (4)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>4R/7R</td>
<td>2 (2)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>4R/8R</td>
<td>3 (6)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>5R/7R</td>
<td>0 (0)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>6R/7R</td>
<td>1 (2)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>7R/7R</td>
<td>1 (2)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>L/L</td>
<td>N/A</td>
<td>N/A</td>
<td>12 (24)</td>
<td>N/A</td>
</tr>
<tr>
<td>L/S</td>
<td>N/A</td>
<td>N/A</td>
<td>24 (48)</td>
<td>N/A</td>
</tr>
<tr>
<td>S/S</td>
<td>N/A</td>
<td>N/A</td>
<td>14 (28)</td>
<td>N/A</td>
</tr>
<tr>
<td>9/9</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>5 (12)</td>
</tr>
<tr>
<td>9/10</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>27 (42)</td>
</tr>
<tr>
<td>10/10</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>18 (36)</td>
</tr>
</tbody>
</table>

aDRD4: dopamine D4 receptor gene.
bMAO-A: monoamine oxidase A.
c5-HTTLPR: serotonin-transporter-linked polymorphic region.
dDAT1: dopamine transporter 1.
eN/A: not applicable.
### Table 2. Allele and genotype frequencies of the dopamine D4 receptor, monoamine oxidase A, serotonin-transporter, and dopamine active transporter 1 genes among females.

<table>
<thead>
<tr>
<th>Genes</th>
<th>DRD4&lt;sup&gt;a&lt;/sup&gt;, n (%)</th>
<th>MAO-A&lt;sup&gt;b&lt;/sup&gt;, n (%)</th>
<th>5-HTTLPR&lt;sup&gt;c&lt;/sup&gt;, n (%)</th>
<th>DAT1&lt;sup&gt;d&lt;/sup&gt;, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allele</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>7 (6.5)</td>
<td>2 (1.8)</td>
<td>N/A&lt;sup&gt;e&lt;/sup&gt;</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>11 (10.2)</td>
<td>33 (30.6)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>4</td>
<td>84 (77.7)</td>
<td>68 (63)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>5</td>
<td>3 (2.7)</td>
<td>5 (4.6)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>6</td>
<td>0 (0)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>7</td>
<td>3 (2.7)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>8</td>
<td>0 (0)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>L</td>
<td>N/A</td>
<td>N/A</td>
<td>62 (57.4)</td>
<td>N/A</td>
</tr>
<tr>
<td>S</td>
<td>N/A</td>
<td>N/A</td>
<td>46 (42.6)</td>
<td>N/A</td>
</tr>
<tr>
<td>9</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>42 (38.9)</td>
</tr>
<tr>
<td>10</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>66 (61.1)</td>
</tr>
<tr>
<td>Genotype</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2R/2R</td>
<td>1 (1.9)</td>
<td>1 (1.9)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2R/4R</td>
<td>5 (9.3)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2R/5R</td>
<td>0 (0)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2R/7R</td>
<td>0 (0)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3R/3R</td>
<td>3 (5.6)</td>
<td>9 (16.7)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3R/4R</td>
<td>3 (5.6)</td>
<td>15 (27.8)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3R/7R</td>
<td>2 (3.7)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3R/8R</td>
<td>0 (0)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>4R/4R</td>
<td>37 (68.5)</td>
<td>26 (48.1)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>4R/5R</td>
<td>2 (3.7)</td>
<td>1 (1.9)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>4R/7R</td>
<td>0 (0)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>4R/8R</td>
<td>0 (0)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>5R/5R</td>
<td>N/A</td>
<td>2 (3.7)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>5R/7R</td>
<td>1 (1.9)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>6R/7R</td>
<td>0 (0)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>7R/7R</td>
<td>0 (0)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>L/L</td>
<td>N/A</td>
<td>N/A</td>
<td>17 (31.5)</td>
<td>N/A</td>
</tr>
<tr>
<td>L/S</td>
<td>N/A</td>
<td>N/A</td>
<td>28 (51.9)</td>
<td>N/A</td>
</tr>
<tr>
<td>S/S</td>
<td>N/A</td>
<td>N/A</td>
<td>9 (16.7)</td>
<td>N/A</td>
</tr>
<tr>
<td>9/9</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>9 (16.6)</td>
</tr>
<tr>
<td>9/10</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>23 (42.6)</td>
</tr>
<tr>
<td>10/10</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>22 (40.8)</td>
</tr>
</tbody>
</table>

<sup>a</sup>DRD4: dopamine D4 receptor gene.
<sup>b</sup>MAO-A: monoamine oxidase A.
<sup>c</sup>5-HTTLPR: serotonin-transporter-linked polymorphic region.
<sup>d</sup>DAT1: dopamine transporter 1.
<sup>e</sup>N/A: not applicable.
Association Between IA and Genetic Polymorphisms

To verify the possible association between IA and genetic polymorphisms in MAO-A, 5-HTTPR, DRD4, and DAT1 genes among male and female youths, sex-stratified logistic-regression models were conducted. If the CI crosses 1, the difference between genotype groups can be considered not significant. As shown in Tables 3 and 4, no significant associations were found between IA and any considered genotypes, both among males and females (all CIs crossed 1).

Table 3. Association between internet addiction and genetic polymorphisms in serotonin-transporter, dopamine active transporter 1, dopamine D4 receptor, and monoamine oxidase A among male youths.

<table>
<thead>
<tr>
<th>Gene and genotype</th>
<th>Wald χ² (df)</th>
<th>P value</th>
<th>ORᵃ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>5-HTTPRᵇ</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S/S vs S/L+L/Lᶜ</td>
<td>0.5 (1)</td>
<td>.47</td>
<td>1.57 (0.45-5.45)</td>
</tr>
<tr>
<td>L/L vs S/S+S/Lᵈ</td>
<td>0.001 (1)</td>
<td>.97</td>
<td>0.98 (0.26-3.66)</td>
</tr>
<tr>
<td>S/Sᵉ</td>
<td>0.5 (2)</td>
<td>.75</td>
<td>N/Aᶠ</td>
</tr>
<tr>
<td>S/L</td>
<td>0.05 (1)</td>
<td>.67</td>
<td>1.4 (0.29-6.62)</td>
</tr>
<tr>
<td>L/L</td>
<td>0.3 (1)</td>
<td>.80</td>
<td>0.84 (0.20-3.45)</td>
</tr>
<tr>
<td><strong>DAT1ᵍ</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10/10 vs 9/10+9/9ᵉ</td>
<td>0.1 (1)</td>
<td>.73</td>
<td>0.81 (0.25-2.65)</td>
</tr>
<tr>
<td>9/9 vs 9/10+10/10ᵈ</td>
<td>0.7 (1)</td>
<td>.39</td>
<td>2.25 (0.34-14.83)</td>
</tr>
<tr>
<td>10/10ᵉ</td>
<td>0.7 (2)</td>
<td>.69</td>
<td>N/A</td>
</tr>
<tr>
<td>9/10</td>
<td>0.6 (1)</td>
<td>.40</td>
<td>0.42 (0.05-3.21)</td>
</tr>
<tr>
<td>9/9</td>
<td>0.6 (1)</td>
<td>.43</td>
<td>0.45 (0.06-3.21)</td>
</tr>
<tr>
<td><strong>DRD4ʰ</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/4 vs non 4/4ᵉ</td>
<td>0.5 (1)</td>
<td>.47</td>
<td>0.65 (0.21-2.06)</td>
</tr>
<tr>
<td><strong>MAO-Aʲ</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H vs Lʲ</td>
<td>0.007 (1)</td>
<td>.93</td>
<td>0.95 (0.29-3.11)</td>
</tr>
<tr>
<td>L vs H</td>
<td>0.007(1)</td>
<td>.93</td>
<td>1.05 (0.32-3.45)</td>
</tr>
</tbody>
</table>

ᵃOR: odds ratio.
bS-HTTLPR: serotonin-transporter-linked polymorphic region.
ᶜDominant.
ᵈRecessive.
ᵉAdditive.
ᶠN/A: not applicable.
ᵍDAT1: dopamine transporter 1.
ʰDRD4: dopamine D4 receptor gene.
ʲMAO-A: monoamine oxidase A.
ʲAllelic.
Table 4. Association between internet addiction and genetic polymorphisms in serotonin-transporter, dopamine active transporter 1, dopamine D4 receptor, and monoamine oxidase A among female youths.

<table>
<thead>
<tr>
<th>Gene and genotype</th>
<th>Wald $\chi^2$ (df)</th>
<th>P value</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>5-HTTLP</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S/S vs S/L+L/L&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.07 (1)</td>
<td>.78</td>
<td>0.78 (0.14-4.32)</td>
</tr>
<tr>
<td>L/L vs S/S+S/L&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.1 (1)</td>
<td>.69</td>
<td>1.29 (0.35-4.68)</td>
</tr>
<tr>
<td>S/S&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.1 (2)</td>
<td>.91</td>
<td>N/A&lt;sup&gt;f&lt;/sup&gt;</td>
</tr>
<tr>
<td>S/L</td>
<td>0.1 (1)</td>
<td>.69</td>
<td>0.68 (0.10-4.52)</td>
</tr>
<tr>
<td>L/L</td>
<td>0.1 (1)</td>
<td>.74</td>
<td>0.80 (0.20-3.08)</td>
</tr>
<tr>
<td><strong>DAT1</strong>&lt;sup&gt;g&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10/10 vs 9/10+9/9&lt;sup&gt;e&lt;/sup&gt;</td>
<td>1.1 (1)</td>
<td>.28</td>
<td>0.48 (0.13-1.82)</td>
</tr>
<tr>
<td>9/9 vs 9/10+10/10&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.0 (1)</td>
<td>.78</td>
<td>0.78 (0.14-4.32)</td>
</tr>
<tr>
<td>10/10&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.6 (2)</td>
<td>.43</td>
<td>N/A</td>
</tr>
<tr>
<td>9/10</td>
<td>0.06 (1)</td>
<td>.79</td>
<td>0.77 (0.11-5.24)</td>
</tr>
<tr>
<td>9/9</td>
<td>0.4 (1)</td>
<td>.49</td>
<td>1.86 (0.31-11.18)</td>
</tr>
<tr>
<td><strong>DRD4</strong>&lt;sup&gt;h&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/4 vs non 4/4&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.4 (1)</td>
<td>.48</td>
<td>1.55 (0.44-5.42)</td>
</tr>
<tr>
<td><strong>MAO-A</strong>&lt;sup&gt;i&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H/H vs H/L+L/L&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.2 (1)</td>
<td>.64</td>
<td>0.75 (0.22-2.55)</td>
</tr>
<tr>
<td>L/L vs H/H+H/L&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.6 (1)</td>
<td>.41</td>
<td>0.50 (0.09-2.62)</td>
</tr>
<tr>
<td>H/H&lt;sup&gt;e&lt;/sup&gt;</td>
<td>1.7 (2)</td>
<td>.42</td>
<td>N/A</td>
</tr>
<tr>
<td>H/L</td>
<td>0.2 (1)</td>
<td>.65</td>
<td>0.66 (0.11-3.91)</td>
</tr>
<tr>
<td>L/L</td>
<td>0.9 (1)</td>
<td>.31</td>
<td>2 (0.51-7.81)</td>
</tr>
</tbody>
</table>

<sup>a</sup>OR: odds ratio.  
<sup>b</sup>5-HTTLPR: serotonin-transporter-linked polymorphic region.  
<sup>c</sup>Dominant.  
<sup>d</sup>Recessive.  
<sup>e</sup>Additive.  
<sup>f</sup>N/A: not applicable.  
<sup>g</sup>DAT1: dopamine transporter 1.  
<sup>h</sup>DRD4: dopamine D4 receptor gene.  
<sup>i</sup>MAO-A: monoamine oxidase A.

**IA, Family Functioning, and Psychopathological Symptoms of Young Adults**

To verify possible differences between the 2 groups in family functioning and the psychological profile of young adults (ie, depression, anxiety, and avoidant personality problems), an independent samples $t$ test was conducted. Results showed significant differences between youths with IA and their peers in the control group in the scores of the FAD AR subscale scores and ASR scores in all psychopathological areas (Table 5).
Table 5. t test for family functioning and psychological profile for the 2 groups.

<table>
<thead>
<tr>
<th>Groups or study variables</th>
<th>IAG(^a) (n=35), mean (SD)</th>
<th>CG(^b) (n=69), mean (SD)</th>
<th>t test (df)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FAD(^c)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PS(^d)</td>
<td>13.60 (2.95)</td>
<td>12.98 (3.36)</td>
<td>0.91 (102)</td>
<td>.36</td>
</tr>
<tr>
<td>CM(^e)</td>
<td>20.31 (2.97)</td>
<td>19.36 (2.69)</td>
<td>1.64 (102)</td>
<td>.10</td>
</tr>
<tr>
<td>RL(^f)</td>
<td>25.97 (2.13)</td>
<td>26.27 (2.02)</td>
<td>-0.71 (102)</td>
<td>.47</td>
</tr>
<tr>
<td>AR(^g)</td>
<td>15.25 (2.47)</td>
<td>13.95 (2.62)</td>
<td>2.43 (102)</td>
<td>.01</td>
</tr>
<tr>
<td>AL(^h)</td>
<td>18.25 (2.01)</td>
<td>18.59 (1.44)</td>
<td>-0.97 (102)</td>
<td>.33</td>
</tr>
<tr>
<td>BC(^i)</td>
<td>21.51 (2.09)</td>
<td>20.73 (2.48)</td>
<td>1.58 (102)</td>
<td>.11</td>
</tr>
<tr>
<td><strong>ASR(^j)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>7.80 (5.08)</td>
<td>5.66 (4.16)</td>
<td>2.28 (102)</td>
<td>.02</td>
</tr>
<tr>
<td>Anxiety</td>
<td>7.25 (2.59)</td>
<td>5.94 (2.24)</td>
<td>2.68 (102)</td>
<td>.009</td>
</tr>
<tr>
<td>Avoidant personality</td>
<td>4.94 (2.85)</td>
<td>3.27 (1.90)</td>
<td>3.11 (102)</td>
<td>.003</td>
</tr>
</tbody>
</table>

\(^a\)IAG: internet addiction group.
\(^b\)CG: control group.
\(^c\)FAD: Family Assessment Device.
\(^d\)PS: problem solving.
\(^e\)CM: communication.
\(^f\)RL: roles.
\(^g\)AR: affective responsiveness.
\(^h\)AI: affective involvement.
\(^i\)BC: behavioral control.
\(^j\)ASR: Adult Self-Report.

The Mediation Role Played by the Psychological Profile on the Relationship Between Family Functioning and IA

On the basis of the previous analyses, we verified the possible simple and multiple mediation effects of depressive and anxiety symptoms and avoidant personality traits on the relationship between FAD AR and IA. Mediation analyses were performed using the SPSS Macro PROCESS [91], which provides coefficient estimates for the total, direct, and indirect effects of variables using ordinary least squares regression. In particular, Model 4 [91] was used, which allows us to estimate the indirect effects within a 95% CI. The indirect effect can be considered statistically significant if the CI does not include zero. In our first simple mediation model, we tested whether youth’s depressive symptoms mediated the effect of FAD AR on youth’s IA. Results showed that FAD AR was significantly associated with youth depressive symptoms (\(B=0.29; P=.002\)), which, in turn, was positively related to the youth’s scores on IAT (\(B=3.33; P<.001\)). Moreover, considering the relationship between FAD AR and IAT, both direct (\(B=1.90; P=.04\)) and total effect (\(B=2.90; P=.002\)) were significant (Figure 1).
Figure 1. Simple mediation model for depressive problems mediating the association between affective responsiveness and internet addiction. (a) Direct effect of affective responsiveness on depression; (b) direct effect of depression on internet addiction test (IAT) score; (c') direct effect of affective responsiveness on IAT score; (c) total effect (a*b+c') of affective responsiveness on IAT score. ASR: Adult Self-Report; FAD: Family Assessment Device; IAT: internet addiction test.

With regard to the indirect effect, the bootstrap CI showed that the indirect path via depression was statistically significant (Table 6).

We then tested whether the anxiety symptoms of young adults mediated the relationships between the FAD AR and IA. The results showed that the indirect effect of FAD AR through anxiety symptoms of young adults was not significant (Table 6). However, both direct ($B=2.50$; $P=.008$) and total effect ($B=2.90$; $P=.002$) of FAD AR on IAT were significant. Moreover, anxiety problems of young adults were significantly associated with their scores on IAT ($B=2.61$; $P=.005$), but the relationship between FAD AR and depression in young adults was not significant ($P=.12$; Figure 2).

In relation to the possible mediation role played by avoidant personality traits of young adults on the relationship between FAD AR and IA, results showed that FAD AR was significantly related to their avoidant personality traits ($B=0.32$; $P<.001$) that, in turn, was significantly associated with the IA of young adults ($B=3.38$; $P<.001$; Figure 3).
Table 6. Simple mediator models showing total, direct, and indirect effects of affective responsiveness on internet addiction.

<table>
<thead>
<tr>
<th>Psychopathological areas and paths</th>
<th>Path effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B (SE)</td>
</tr>
<tr>
<td><strong>Depression</strong></td>
<td></td>
</tr>
<tr>
<td>Direct path</td>
<td></td>
</tr>
<tr>
<td>AfRes$^a$→DEP$^b$</td>
<td>0.29 (0.09)</td>
</tr>
<tr>
<td>DEP→IAT$^c$</td>
<td>3.33 (0.94)</td>
</tr>
<tr>
<td>AfRes→IAT</td>
<td>1.90 (0.94)</td>
</tr>
<tr>
<td>Indirect path</td>
<td></td>
</tr>
<tr>
<td>AfRes→DEP→IAT</td>
<td>0.99 (0.44$^d$)</td>
</tr>
<tr>
<td>Total$^f$</td>
<td></td>
</tr>
<tr>
<td>AfRes→IAT</td>
<td>2.90 (0.94)</td>
</tr>
<tr>
<td><strong>Anxiety</strong></td>
<td></td>
</tr>
<tr>
<td>Direct path</td>
<td></td>
</tr>
<tr>
<td>AfRes→ANX$^g$</td>
<td>0.15 (0.09)</td>
</tr>
<tr>
<td>ANX→IAT</td>
<td>2.61 (0.92)</td>
</tr>
<tr>
<td>AfRes→IAT</td>
<td>2.50 (0.92)</td>
</tr>
<tr>
<td>Indirect path</td>
<td></td>
</tr>
<tr>
<td>AfRes→ANX→IAT</td>
<td>0.39 (0.35$^d$)</td>
</tr>
<tr>
<td>Total$^h$</td>
<td></td>
</tr>
<tr>
<td>AfRes→IAT</td>
<td>2.90 (0.94)</td>
</tr>
<tr>
<td><strong>Avoidant personality</strong></td>
<td></td>
</tr>
<tr>
<td>Direct path</td>
<td></td>
</tr>
<tr>
<td>AfRes→AvP$^i$</td>
<td>0.32 (0.09)</td>
</tr>
<tr>
<td>AvP→IAT</td>
<td>3.38 (0.94)</td>
</tr>
<tr>
<td>AfRes→IAT</td>
<td>1.80 (0.94)</td>
</tr>
<tr>
<td>Indirect path</td>
<td></td>
</tr>
<tr>
<td>AfRes→AvP→IAT</td>
<td>1.09 (0.45$^d$)</td>
</tr>
<tr>
<td>Total$^j$</td>
<td></td>
</tr>
<tr>
<td>AfRes→IAT</td>
<td>2.90 (0.94)</td>
</tr>
</tbody>
</table>

$^a$AfRes: affective responsiveness.
$^b$DEP: Depression.
$^c$IAT: internet addiction test.
$^d$Bootstrapped SE values.
$^e$Bootstrap CI values.
$^f$Proportion mediated by indirect effect is 0.34%.
$^g$ANX: Anxiety.
$^h$Proportion mediated by indirect effect is 0.13%.
$^i$AvP: avoidant personality.
$^j$Proportion mediated by indirect effect is 0.37%. 
The indirect path via avoidant personality traits of young adults was also statistically significant (Table 6).

Finally, results of the multiple mediation analyses showed a significant total indirect effect ($B=1.24; 95\%$ bootstrapped CI 0.28-2.38), implying that the changes in ASR depression, anxiety, and avoidant personality collectively mediated the relationship between FAD AR and IA. However, the specific indirect effects of the 3 mediators were not significant (CI included zero; Table 7).
Table 7. Multiple mediator model showing total, direct, and indirect effects of affective responsiveness on internet addiction.

<table>
<thead>
<tr>
<th>Paths</th>
<th>Path effect B (SE)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct path</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(A₁) AfRes² → DEPᵇ</td>
<td>0.29 (0.09)</td>
<td>0.10 to 0.48</td>
</tr>
<tr>
<td>(A₂) AfRes → ANXᶜ</td>
<td>0.15 (0.09)</td>
<td>-0.04 to 0.34</td>
</tr>
<tr>
<td>(A₃) AfRes → AvPᵈ</td>
<td>0.32 (0.09)</td>
<td>0.13 to 0.50</td>
</tr>
<tr>
<td>(B₁) DEP → IATₑ</td>
<td>1.86 (1.23)</td>
<td>-0.59 to 4.32</td>
</tr>
<tr>
<td>(B₂) ANX → IAT</td>
<td>0.60 (1.16)</td>
<td>-1.69 to 2.91</td>
</tr>
<tr>
<td>(B₃) AvP → IAT</td>
<td>1.84 (1.29)</td>
<td>-0.73 to 4.42</td>
</tr>
<tr>
<td>(C') AfRes → IAT</td>
<td>1.65 (0.95)</td>
<td>-0.23 to 3.55</td>
</tr>
<tr>
<td>Indirect path (A*B)ᶠ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1.24 (0.53)</td>
<td>0.28 to 2.38</td>
</tr>
<tr>
<td>AfRes → DEP → IAT</td>
<td>0.55 (0.42)</td>
<td>-0.17 to 1.48</td>
</tr>
<tr>
<td>AfRes → ANX → IAT</td>
<td>0.09 (0.26)</td>
<td>-0.33 to 0.76</td>
</tr>
<tr>
<td>AfRes → AvP → IAT</td>
<td>0.59 (0.47)</td>
<td>-0.27 to 1.60</td>
</tr>
<tr>
<td>Total (C=A*B+C’)ᵍ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AfRes → IAT</td>
<td>2.90 (0.94)</td>
<td>1.01 to 4.78</td>
</tr>
</tbody>
</table>

ᵃAfRes: affective responsiveness. ᵇDEP: depression. ᶜANX: anxiety. ᵈAvP: avoidant personality. ᵉIAT: internet addiction test. ᶠBootstrapped SE values and CI values. ᵍProportion mediated by specific indirect effect of AfRes → DEP → IAT is 0.24%, AfRes → ANX → IAT is 0.05%, and AfRes → AvP → IAT is 0.26%.

The Moderator Role Played by Polymorphisms in Young Adults on the Relationship Between Family Functioning and IA

Finally, we evaluated the possible moderating role played by the genetic polymorphisms in young adults in MAO-A, 5HTTPR, DRD4, and DAT1 genes on the relationship between FAD AR and their scores on IAT. Moderation analyses were conducted using the method outlined by Hayes [91] (Model 1). The results showed that the relationship between FAD AR and IA was moderated by the genotype of the 5-HTTLPR, MAO-A, and DAT1 genes in young adults. The genotype of the DRD4 gene in young adults did not moderate this relationship (P=.41). In particular, high scores on FAD AR were positively associated with high levels of IA, but only in the presence of the L/x genotype of the 5-HTTLPR gene (B=.37; t₁₀₀=3.54, P<.001; S/S: P=.85; Figure 4).

In addition, the 9/x genotype of the DAT1 was significantly related with high scores on IAT (B=0.42; t₁₀₀=3.47, P<.001; 10/10: P=.56; Figure 5).

Finally, the results showed a significant interactive effect of the MAOA-H polymorphisms (B=0.44; t₁₀₀=3.55, P<.001; MAOA-L: P=.50; Figure 6).
**Figure 4.** Moderation of youth’s 5-HTTLPR genotype on the relationship between the youth’s family functioning and IAT score. 5-HTTLPR: Serotonin-transporter linked polymorphic region; FAD_AR: Family Assessment Device Affective Responsiveness; IAT: internet addiction test.

**Figure 5.** Moderation of youth’s DAT1 genotype on the relationship between the youth’s family functioning and IAT score. DAT: dopamine active transporter; FAD_AR: Family Assessment Device Affective Responsiveness; IAT: internet addiction test.
Discussion

Principal Findings

This exploratory study aims to investigate the possible biological, psychological, and environmental risk factors associated with IA in a young adult population. On the basis of a biopsychosocial model [71] that considered IA as a result of a mutual influence between individual genotype, psychological profile, and social environment, this study aimed to verify the possible influence of genetic polymorphisms in MAO-A, 5HTTPLR, DRD4, and DAT1 genes, considering the role played by the perception of young adults of their family functioning and their psychopathological difficulties. We considered the dopaminergic system (DAT and DRD4 genes), the serotonergic system (5-HTTPLR gene), and the monoamine metabolism pathway (MAO-A gene) for their central role in impulsive behaviors [92-94]. The neurotransmitter systems encoded by these genes are reportedly associated with alcoholism or other addictions [45,72,73,75,95-98]. Moreover, several studies have shown that these genetic polymorphisms contribute to the susceptibility of an individual to environmental influences [99-101] and can act as moderators in associations between the quality of the emotional environment provided by parents and the evolutionary outcomes of young adults [102].

Our first hypothesis was that the genotypes of the 5-HTTPLR, DAT1, DRD4, and MAO-A genes differed significantly between the IA and non-IA groups. However, we found no significant association between IA and any genetic polymorphisms, neither males nor females. Specifically, we found no association between the 5-HTTPLR polymorphism and IA. Previous studies [36,41] have reported that the S/S genotype HTTLPR increased the risk of IA. Similarly, our study indicated a higher percentage of the S/S genotype in the IA group than in the control group, but there was no significant difference. Moreover, we found no association between the DAT1 genotype and IA. Previous studies have shown that the DAT1 polymorphism is associated with addiction behavior, such as alcoholism [103,104], substance use/abuse [105], and pathological gambling [106]. Some studies [72,107] have suggested that the DAT1 genotype 9/9 is associated with a higher risk for addictive disorders. However, the findings of our study are in line with the findings from studies by other researchers [74,75] that showed no significant association with internet-related addiction. Given these inconsistent findings, further research is needed to clarify whether the DAT1 gene may have an effect on IA. In addition, we found no differences in the DRD4 genotype between the IA and non-IA groups. The DRD4 4/4 polymorphism was shown to be a risk genotype for substance-related disorders [108]. Moreover, it was found to be correlated with decreased risk for inattention [34,35] and, consequently, with increased duration of internet use and a higher risk of IA [36]. However, other studies have underlined a higher risk for addictive disorders associated with the 7-repeat allele of DRD4 [109,110]. In our sample, we found a rarity of the 7-repeat allele (11/104, 10.5%) and, given that previous studies have shown that there was a statistically significant difference in allele distribution of the DRD4 gene between ethnicities [111], further evaluation is necessary for a better understanding of the possible influence of the DRD4 genotype on IA in specific ethnic groups. Finally,
Our second hypothesis was that the quality of family functioning perceived by youths and their psychological profiles differed significantly between the IA and non-IA groups. As expected, we found that youths in the IA group reported significantly higher scores (representative of a poorer family functioning) on the subscale of FAD AR and on all psychopathological areas considered (ie, depressive and anxiety symptoms and avoidant personality problems) than their peers in the control group. These findings are consistent with previous studies that have underlined that poor quality of family functioning is one of the main social-environmental risk factors for the onset and maintenance of addictive behaviors among young adult populations [51,52], including IA [49,50,114,115]. Families with poor AR are characterized by difficulties in showing their emotions [84]. In this field, some authors have suggested that, given that feelings are not expressed or tolerated, youths may be more susceptible to cope with their internal stress by themselves [116] and may use the internet excessively as a strategy to cope with negative emotions resulting from the interpersonal relationships with parents [117,118]. Moreover, it has been suggested that young adults who perceive a lack of supportive and intimate relationships with their parents are more likely to search for social support from virtual interactive experiences [119].

With regard to the possible association between IA and psychological difficulties of young adults, our results confirmed that youths of the IA groups showed higher scores on depressive, anxiety, and avoidant personality problems. These findings are in line with previous studies that have shown that young adults affected by IA also reported the presence of more severe psychopathological symptoms than their peers without IA [15], especially in the areas of depression [76,77], anxiety [120], and traits of avoidant personality disorder [70]. The nature of the association between IA and other symptomatic psychopathological areas is controversial. International scientific literature has suggested that psychopathological symptoms may cause or contribute to the onset of IA or, conversely, IA may cause or further exacerbate the course of other psychopathological symptoms [76,121,122]. However, recent longitudinal studies [123,124] have produced new evidence on the predictive role of psychological suffering on IA. In this field, some authors have suggested that youths may be at higher risk of developing IA in an attempt to cope with their psychological problems [2]. For example, a youth affected by depression could use the internet as a strategy to escape and cope with feelings of sadness and low self-esteem [122]; excessive use of the internet could be used by young people in an attempt to manage anticipatory anxiety associated with stressful situations and life events [66]; people with avoidant personalities are often socially inhibited and have difficulty communicating with others in face-to-face situations [125], and, thus, they may use web-based communication to seek interpersonal interactions in a way safer and easier for them [126].

Our third hypothesis was that the psychopathological symptoms of young adults might mediate the relationship between FAD AR and IA. International research has shown that poor quality of family functioning is a significant predictor of the psychopathological symptoms of young adults [2,127], which, in turn, has been reported as a significant predictor of IA [123,124], suggesting a possible mediation role. Our results showed that depressive symptoms and avoidant personality problems of young adults mediated the relationship between FAD AR and IA. Moreover, our results have confirmed previous studies on the significant positive association between the perception of a poor AR of families of young adults to their depressive problems [127,128] and avoidant personality traits [129]. However, anxiety symptoms in youths did not mediate the relationship between FAD AR and IA, although the direct effect on IA was confirmed and is in line with studies by Sepehrian and Lotf [130] and Razieh et al [131]. The results of multiple mediation analyses showed only a significant total indirect effect, indicating that depression and anxiety problems, and avoidant personality traits together significantly explained the association between youth’s family AR and IA. However, the specific effects of the 3 mediators were not significant. As evidenced by Hayes [91], these results are only apparently contradictory, and could be as a result of the presence of a correlation between the mediators and the small size of the sample. Although previous research has underlined that psychopathological symptoms of young adults mediate the relationships between environmental risk factors (eg, parental psychopathological symptoms, stressful life events) and IA [132,133], this is the first study to explore the possible role played by family functioning in these processes.

Finally, our fourth hypothesis was that genetic polymorphisms in MAO-A, 5HTTLPR, DRD4, and DAT1 genes might moderate the relationship between family functioning and IA. Our results showed that the relationship between FAD AR and IA was moderated by the genotype of the 5-HTTLPR, DAT1, and MAO-A genes in young adults. In particular, poor AR perceived by youths in their family functioning was associated with a high level of IA, but only in the presence of the L/x genotype of the 5-HTTLPR gene, the MAOA-H genotype, and the 9/x genotype of the DAT1. These findings are in accordance with recent evidence from geneXenvironment (GxE) interaction studies, which have indicated that not all individuals are susceptible to environmental influences to the same extent [134]. In this regard, the individual genotype (genetic variations) is considered to play a crucial role in moderating individual sensitivity to environmental exposure [79] and to the development of psychopathology related to addiction [48]. In particular, although studies by Lee et al [41] and Sun et al [36] have reported a significant association between IA and the presence of 2 copies of the short allele (S/S) of the 5-HTTLPR, our results are in line with previous studies that have reported an interaction effect between the L allele with stressful life events and a poor quality of family relationship on the onset of psychopathological problems [109,135], including difficulties...
in the area of addiction [136,137]. Moreover, the presence of the MAOA-H genotype gene has been shown to be a significant moderator in the relationship between family environmental exposure and addiction-related problems [138,139], and individuals carrying the 9-repeat allele of the DAT1 gene are at higher risk of alcoholism [140] and other addictive problems [72,107]. However, to the best of our knowledge, this is the first study to explore the possible influence of GxE on IA.

Possible Limitations, Strength, and Implications
This study has some limitations. First, the small sample size and the resulting limited statistical power that should be taken with caution in our preliminary findings, which should be confirmed by further studies with larger samples. Moreover, the cross-sectional nature of the study did not allow testing of causal links between the variables taken into account, which should be explored in subsequent longitudinal studies. Moreover, we used a self-report tool for the assessment of IA of young adults, their family functioning, and psychopathological difficulties. Although these tools have proven their worth in numerous studies, future studies should evaluate these variables through more robust methodologies, such as observational procedures or clinical interviews. Despite these limitations, this exploratory study is the first to focus on the possible moderation of genetic polymorphisms in MAO-A, 5-HTTPR, DRD4, and DAT1 genes on the relationship between the quality of family functioning and IA of young adults, reporting significant GxE effects. Overall, our findings add new evidence with regard to the biological, individual, and social risk factors associated with IA in the young adult population, which can be evaluated for the development of more targeted and effective intervention programs.

Conflicts of Interest
None declared.

References

http://mental.jmir.org/2020/12/e17341/


47. Cerniglia et al. JMIR MENTAL HEALTH 2016.12.024


Abbreviations

5-HT: Serotonin
5-HTTLPR: Serotonin-transporter linked polymorphic region
5-HTTPR: Serotonin-transporter
AR: affective responsiveness
ASR: Adult Self-Report
DA: dopamine
DAT: dopamine active transporter
DRD4: dopamine D4 receptor
FAD: Family Assessment Device
GxE: geneXenvironment
IA: internet addiction
IAT: internet addiction test
MAO-A: monoamine oxidase A
MAOA-H: MAO-A high-activity variant
MAOA-L: MAO-A low-activity variant
VNTR: variable number of tandem repetitions
Health Care Providers’ Perceptions of Quality, Acceptance, and Satisfaction With Telebehavioral Health Services During the COVID-19 Pandemic: Survey-Based Study

Jesse Wright¹, MSc; Shyam Dewan²; Donald Hilty³, MD; Naakesh A Dewan⁴, MD

¹BayCare Health System, Clearwater, FL, United States
²Emory University, Atlanta, GA, United States
³University of Southern California, Los Angeles, CA, United States
⁴Florida Blue, Jacksonville, FL, United States

Corresponding Author:
Jesse Wright, MSc
BayCare Health System
2985 Drew St
Clearwater, FL
United States
Phone: 1 727 519 1285
Email: Jesse.Wright@baycare.org

Abstract

Background: Due to rapidly increasing rates of COVID-19 across the country, system-wide changes were needed to protect the health and safety of health care providers and consumers alike. Technology-based care has received buy-in from all participants, and the need for technological assistance has been prioritized.

Objective: The objective of this study was to determine the initial perceptions and experiences of interprofessional behavioral health providers about shifting from traditional face-to-face care to virtual technologies (telephonic and televideo) during the COVID-19 pandemic.

Methods: A survey-based study was performed at a large, integrated medical health care system in West-Central Florida that rapidly implemented primary care provision via telephone and televideo as of March 18, 2020. A 23-item anonymous survey based on a 7-point Likert scale was developed to determine health care providers’ perceptions about telephonic and televideo care. The survey took 10 minutes to complete and was administered to 280 professionals between April 27 and May 11, 2020.

Results: In all, 170 respondents completed the survey in entirety, among which 78.8% (134/170) of the respondents were female and primarily aged 36-55 years (89/170, 52.4%). A majority of the respondents were outpatient-based providers (159/170, 93.5%), including psychiatrists, therapists, counselors, and advanced practice nurses. Most of them (144/170, 84.7%) had used televideo for less than 1 year; they felt comfortable and satisfied with either telephonic or televideo mode and that they were able to meet the patients’ needs.

Conclusions: Our survey findings suggest that health care providers valued televideo visits equally or preferred them more than telephonic visits in the domains of quality of care, technology performance, satisfaction of technology, and user acceptance.

(JMIR Ment Health 2020;7(12):e23245) doi:10.2196/23245

KEYWORDS

telepsychiatry; COVID-19; telehealth; perception; quality; acceptability; satisfaction; behavior; mental health; health care provider

Introduction

Since the first reported case of COVID-19 in the United States in January 2020, social distancing—a misnomer for physical distancing—has been a public health priority [1]. Declarations of public health emergencies and stay-at-home orders have made the use of telephonic and televideo care services a necessity rather than a choice, while also limiting the number of outdoor interactions among the public since March 17, 2020 [2]. During this period, many health care providers worked from home to avoid the risk of exposure to the virus. In general, telehealth options are not always readily available for health care systems
due to regulatory, reimbursement, and liability concerns [3]. Nevertheless, owing to patient requests, health care systems have been moving toward the integration of in-person and video services to offer more points-of-service by highly skilled health care providers and to efficiently use available resources [4]. However, changes involving technology are time consuming, compete with other demands, and require considerable investment.

Owing to concerns raised by both health care providers and consumers, and an ongoing state of emergency, the decision to transition to technology-based care has not only pushed but also enabled health care systems to make the change. Buy-in has been received from all participants of care (including health care providers and patients), and technological assistance has been rapidly prioritized. The effort to flatten the curve of COVID-19 spread is an opportunity to “accelerate and bend the curve” of digital health [5]. Fortunately, with the relaxation of HIPAA (Health Insurance Portability and Accountability Act) compliance guidelines, an increasing number of telehealth options are now available and are ready for implementation [6]. Some of the technology options available to the health care system include BlueJeans by Verizon (Verizon Communications), Microsoft Teams (Microsoft Corporation), Skype for Business (Microsoft Corporation), AmWell (American Well Corporation), and Doxy.me (Doxy.me, LLC).

The COVID-19 pandemic has challenged health care systems across the world in ways that both resemble as well as differ from challenges posed by other disasters and mass casualty incidents, including natural phenomena (eg, hurricanes, tornadoes, and wildfires), accidents (eg, plane crashes), or human-made crises (eg, terrorism) [7,8]. These events often cause an acute surge of patients that overwhelms hospital, community, and other resources and personnel. COVID-19 poses infrastructural, communication, and other challenges on a broad scale to responders at various levels—local, state, regional, and federal.

The objective of this study was to determine the initial perceptions of health care providers about the rapid shift in health care delivery—from in-person care to televideo and telephonic care, at a large health care system in Florida that serves acute pediatric, adult, and geriatric populations. The procedures and lessons learned from the implementation of virtual care in this health care system could offer a blueprint for other health care systems. More specifically, this study highlights how health care providers feel about the change from the traditional face-to-face visits to the new patient care approach of telephonic and televideo visits and discusses the advantages and limitations thereof.

Methods

Study Overview and Context

This study was conducted at a large, integrated medical health care system in West-Central Florida. The Behavioral Health Division of this system comprised over 750,000 annual outpatient visits and 13,000 inpatient discharges across 5 counties. In anticipation of an emergency order by the Governor of the State, starting March 18, 2020, this health institution decided to proactively shift to a health delivery model offering all care via telephone or televideo. The rationale for this rapid implementation was that the Chief Medical Officer of the Behavioral Health Division felt the need to prioritize patient and staff safety over other factors, such as reimbursement, due to the fear of being overwhelmed by the spread of COVID-19 as in the State of New York [9]. The Behavioral Health Division employs over 1,200 staff, providers, and administrators. These employees had utilized telemedicine to some extent prior to the emergency declaration by the state. However, none of the employed providers had experience using telephonic and televideo solutions in the outpatient setting. Televideo services were rarely used in the inpatient setting, only in cases wherein patient care was absolutely essential.

Study Design and Outcome

The objective of this survey-based study was to determine health care providers’ initial perceptions of telephonic versus televideo care across 5 domains: (1) quality of care (eg, alliance), (2) technology performance (eg, ability to hear), (3) user experience with technology, (4) satisfaction of technology, and (5) user acceptance. These broad domains are consistent with other behavioral health surveys [10].

An anonymous survey regarding the perceptions of telephonic and televideo care services was sent only to the providers within the Behavioral Health division of this hospital system. Overall, the approach used aligned with the checklist for reporting the results of online surveys [11], in terms of description of the purpose, time taken to complete the survey, voluntary participation, and data protection (ie, anonymous); however, no pretesting was done because of the rapid implementation of the new health care delivery model. Moreover, as this was a quality improvement process, it was approved by the institutional leadership rather than an external or academic institutional review board or human subjects committee. The survey results would offer valuable insights into health care providers’ views on telebehavioral visits for purposes of future enhancements and operations.

Participants and Procedures

Communication With and Inputs From Health Care Providers

Staff and health care providers (survey participants) met with administrators to discuss how to reschedule the upcoming intake of new patients and follow-up appointments, with new patient intake shifted to televideo unless there was no means to do that. Workflows were adjusted based on inputs received from the participants.

Technological Adjustments

Coordination was needed among nurses, physicians, therapists, and staff to use information technology. A needs assessment was performed to identify the necessary hardware, software, and other components required to add or supplement existing resources. Various options for televideo (ie, desktop, laptop, cellular phone, and other) were evaluated, but to avoid equipment preparation delays, telephone services were initially
chosen. Primary hardware considerations made for health care providers included either a laptop with a built-in web camera or a desktop with an external camera; however, they could also use a mobile smart phone or a tablet device. The software systems selected were Blue Jeans and Microsoft Teams, both of which can be run on any Windows, Android, or Mac device that the provider or patient may have. Technology planning was classified for inpatient and outpatient divisions, as each division required different software and hardware considerations. All outpatient services and programs used the same software and hardware based on its availability at the time.

**Health Care Providers’ Readiness**

Meetings were held with health care providers to assess their needs and provide education to use technology, as well as for basic clinical skills for televideo with a follow-up one-on-one consultation, if requested, particularly for those providers who had not used televideo before. Another challenge was the training and implementation of Microsoft Teams. An operations coordinator provided one-on-one functionality training to each inpatient provider to enable them to perform basic functions using the software. Furthermore, before an upcoming shift, the operations coordinator tested the Microsoft Teams application with each provider to ensure it was downloaded correctly and that audio and video calling features were fully functional.

**Collaborative Approach**

In order to quickly obtain access requirements for part-time health care providers, a collaborative approach was required across multiple departments, including privileging and credentialing, identity access management, data security, and information services.

**Clinical Rollout**

As of March 18, 2020, outpatient providers started making telephonic contact with patients, as opposed to in-person visits, to mitigate the risk of COVID-19 spread. Specifically, a provider would reach out to a patient at a scheduled time and conduct an interview from their home (or office while maintaining safe physical distance from other individuals); the patient would also connect from his or her home. Once televideo was known to be a feasible solution, it was chosen to be used across the system for the ambulatory team beginning April 7, 2020, and it was implemented as the primary means to conduct virtual patient visits as of April 22, 2020. The scheduling department sends an email invite to the patient. Then, at a designated time, the patient would need to click on the link in the email to connect with their provider. A similar process was used by therapy providers, whereas field-based providers were required to send a passcode to patients to enable them to join specific meeting rooms at a designated time.

Inpatient providers used 2 software applications that were already in place (ie, Microsoft Teams and Skype for Business) to start providing telehealth services immediately from March 18, 2020. For Microsoft Teams, a clinical operations coordinator created a meeting room using the application so the provider and whoever was using the laptop or device at the inpatient facility could initiate a televideo conference. Once the televideo conference began, the team member in the inpatient unit could take the patient to a private room and help conduct the interview with the provider. In the case of Skype for Business, the provider would have a designated individual at the inpatient facility with whom they would initiate a televideo call. Thereafter, the same process as used for Microsoft Teams was followed with regard to seeing the patients.

**Survey Design and Rollout**

The 23-item anonymous survey comprised questions addressing health care providers’ perceptions about telephonic and televideo care. Responses ranged from “Not at all” to “Perfect” on the 1-to 7-point Likert scale, respectively. Participation in the survey was voluntary with no incentives used. In all, 6 of the 23 questions were based on demographics, whereas the remaining 16 questions directly asked respondents the same questions about their perceptions of telephone and televideo care, separately. One of the questions inquired how well the provider saw the patient, which was applicable only to televideo care. The survey took about 10 minutes to complete and was conducted between April 27 and May 11, 2020; that is, it was rolled out 5 days after televideo was chosen as the primary means for care provision in order to understand the providers’ initial perceptions about this mode of virtual care. Our survey was modified based on a survey previously used in a randomized trial [10], but it was not reassessed for validity or internal consistency. Instead of a generalized view on telehealth, our survey was modified to seek opinions to compare the 2 methods used, namely, telephone and televideo. One of the authors is a pioneer in the use of internet-based surveys and evaluating user acceptance and satisfaction of behavioral health technology [12,13].

**Survey Respondents**

This closed survey was sent to a total of 280 professionals in the behavioral health division, including psychiatrists, therapists, counselors, and advanced practice nurses (APRNs). The health care providers received the survey via email, with a clickable link directing them to a SurveyMonkey (SVMK Inc.) webpage. Emails were sent to a limited number of inpatient providers who were using either telephonic or televideo solutions and all outpatient providers across the 5 counties serviced by the health care system, including practices in both urban and rural locations. We decided to enroll only a limited number of inpatient providers because not all inpatient providers were using telebehavioral solutions at that time.

We received a total of 209 survey responses, of which 170 were fully completed. Due to the software design, we were not able to account for duplicate entries or questions skipped by the respondents. However, at the time of distribution of the survey link, participants were notified to complete the survey only once and in its entirety.

**Statistical Analysis**

To determine initial differences of perceptions among users regarding the use of telephone and televideo services, analyses were performed using paired t tests.
Results

Responses on all questions were collected at a completion rate of 60.7% (170/280). Surveys that were not fully completed were not included in this analysis. In all, 78.8% (134/170) of the respondents were female, stratified into the following age ranges: 20-35 years (48/170, 28.2%), 36-55 years (89/170, 52.4%), and ≥56 years (33/170, 19.4%). With regard to the location of care, 93.5% (159/170) of respondents were outpatient providers and 6.5% (11/170) were inpatient providers. The professional backgrounds of the respondents also varied considerably, as follows: physicians (22/170, 12.9%), licensed MH therapists (54/170, 31.8%), social workers (22/170, 12.9%), APRNs (9/170, 5.3%), Bachelor’s degree holders (27/170, 15.9%), unlicensed Master’s degree holders (31/170, 18.2%), and psychiatric support (5/170, 2.9%). The majority of respondents had less than 1 year of experience with telephone (138/170, 81.2%) and televideo care (144/170, 84.7%).

The survey results suggest that health care providers valued televideo mode equally or preferred it more than telephonic mode in the domains of quality of care, technology performance, satisfaction of technology, and user acceptance. Table 1 shows differences in the scores for these domains in terms of developing patient-clinician alliance, meeting the patients’ needs, and evaluating the experience relative to face-to-face care. The following 2 aspects were of particular importance: (1) the ability to provide care as well as in a face-to-face visit and (2) how well the providers met the patients’ needs via telephonic and televideo visits, with corresponding scores of 3.92 versus 4.48 ($t=3.51, P<.001$) and 4.65 versus 5.12 ($t=3.41, P<.001$), for telephonic and televideo visits, respectively. This finding indicates that televideo was preferred for both these aspects. The biggest difference between telephonic and televideo perceptions was observed in response to the question about the patient-clinician alliance (3.98 versus 4.89, $t=7.17, P<.001$), again indicating a preference for televideo.

Table 1. Survey responses from health care providers.

<table>
<thead>
<tr>
<th>Domain and survey question</th>
<th>Score</th>
<th>Telephonic visit, mean score</th>
<th>Televideo visit, mean score</th>
<th>$P$ value</th>
<th>$t$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quality of care</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How well did you develop the patient-clinician alliance?</td>
<td>3.98</td>
<td>4.89</td>
<td>&lt;.001</td>
<td>7.17</td>
<td></td>
</tr>
<tr>
<td>How well did you meet the patient’s needs?</td>
<td>4.65</td>
<td>5.12</td>
<td>&lt;.001</td>
<td>3.41</td>
<td></td>
</tr>
<tr>
<td>Was the care as good as a face-to-face?</td>
<td>3.92</td>
<td>4.48</td>
<td>&lt;.001</td>
<td>3.51</td>
<td></td>
</tr>
<tr>
<td>How freely were you able to talk about patient issues?</td>
<td>4.70</td>
<td>4.88</td>
<td>.20</td>
<td>1.29</td>
<td></td>
</tr>
<tr>
<td><strong>Technology performance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How well were you able to hear the patient?</td>
<td>5.03</td>
<td>4.85</td>
<td>.14</td>
<td>1.48</td>
<td></td>
</tr>
<tr>
<td><strong>Satisfaction of technology</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How satisfied were you with the experience?</td>
<td>4.35</td>
<td>4.86</td>
<td>.0033</td>
<td>2.96</td>
<td></td>
</tr>
<tr>
<td><strong>User acceptance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How would you rate your sophistication?</td>
<td>5.02</td>
<td>5.26</td>
<td>.094</td>
<td>1.68</td>
<td></td>
</tr>
<tr>
<td>How comfortable are you using this technology?</td>
<td>5.36</td>
<td>5.62</td>
<td>.11</td>
<td>1.62</td>
<td></td>
</tr>
</tbody>
</table>

Discussion

We believe this is the first study to survey health care providers across disciplines and settings regarding their initial perceptions of telephonic and televideo visits during a rapid transition from face-to-face visits during the COVID-19 pandemic. Preference for televideo visits over telephonic visits was slight but significant. As health care systems, providers, technology companies, and payers struggle with how, when, and to what extent should technology be utilized [14], the findings of this survey-based study suggest that providers are ready and capable of using various means for interacting with patients. The approach used to introduce technologies is consistent with other health care systems, which have emphasized the need for obtaining inputs from staff and clinicians, notifying patients, and committing to a plan of action and trial periods in order to improve outcomes [15].

However, our study has some limitations. First, the questionnaire was not tested for validity and reliability. Second, it was a brief survey, comprising only a few questions for each domain. Third, although questions for each domain were adapted from existing questionnaires that are probably reliable and valid, the survey could have been retested as well as checked for internal consistency. Fourth, self-reporting methods allow scope for improvement, without further validation. Fifth, ideally, we would compare results from in-person, telephonic, and televideo visit groups. While our methods could have been improved, we think the findings from this study provide valuable insights in terms of ongoing operations and strategic planning. Future efforts are needed to validate this survey and provide a metric for further evaluations of health care providers’ perceptions of
telephonic and televideo visits. Finally, the results may not be
generalizable to other health care systems, in terms of the
providers surveyed, the dimensions of the health care system,
and other aspects. Therefore, further studies should evaluate
patient satisfaction and acceptance, long-term quality, and
ramifications of therapeutic alliances. These studies should also
explore concerns among the health care providers, such as
anxiety about care via televideo, telephone, and other virtual
technologies, as well as other dimensions of care provision
during the COVID-19 pandemic.

The US health care system will need to make considerable
efforts to adapt to new, post-pandemic norms, and individual
health care systems will need to evaluate and review the role
of technology in providing patient care. As a result, institutional
leadership is needed to support technological interventions so
that clinical, technological, and administrative operations can
ensure the wellbeing of providers and the health of patients as
well as the community at large [16]. In addition, to ensure
quality of care, health care systems could benefit from more
directly assessing providers’ skills, prioritizing specific ones,
and aligning the implementation of technology training with
competencies, such as those for video [17,18], social media
[19], mobile health [20], and asynchronous health care [21].
Telehealth may also afford opportunities to reach new
populations, help underserved ones, and build relationships with
community partners [21].

Conflicts of Interest
None declared.

References
[accessed 2020-11-25]
4. Hilty D, Yellowlees PM, Parrish MB, Chan S. Telepsychiatry: effective, evidence-based, and at a tipping point in health
5. 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] [Medline: 32213476]
[Medline: 17602028]
[Medline: 23830660]
10. Hilty D, Marks S, Wegelin J, Callahan E, Nesbitt T. A randomized, controlled trial of disease management modules,
including telepsychiatric care, for depression in rural primary care. Psychiatry (Edgmont) 2007 Feb;4(2):58-65 [FREE Full
text] [doi: 10.1016/j.psc.2007.07.027] [Medline: 18418262]
12. Dewan, Lorenzi NM. Behavioral health information systems. Evaluating readiness and user acceptance. MD Comput
clinic to a 100% virtual telepsychiatry clinic in response to COVID-19. Psychiatr Serv 2020 Jul 01;71(7):749-752. [doi:
10.1176/appi.ps.202000230] [Medline: 32460682]
16. Metzl JM, Maybank A, De Maio F. Responding to the COVID-19 pandemic: the need for a structurally competent health


Abbreviations

APRNs: advanced practice nurses
HIPAA: Health Insurance Portability and Accountability Act

©Jesse Wright, Shyam Dewan, Donald Hilty, Naakesh A Dewan. Originally published in JMIR Mental Health (http://mental.jmir.org), 04.12.2020. 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 Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.
Web-Based Relaxation Intervention for Stress During Social Isolation: Randomized Controlled Trial

Orignial Paper

Silvia Francesca Maria Pizzoli1,2, MSc; Chiara Marzorati2, PhD; Davide Mazzoni1, PhD; Gabriella Pravettoni1,2, PhD

1University of Milan, Department of Oncology and Hematology-Oncology, Milan, Italy
2Applied Research Division for Cognitive and Psychological Science, European Institute of Oncology IEO, IRCCS, Milan, Italy

Corresponding Author:
Silvia Francesca Maria Pizzoli, MSc
University of Milan
Department of Oncology and Hematology-Oncology
Via Festa del Perdono 7
Milan, 20122
Italy
Phone: 39 0294372099
Email: silviafrancescamaria.pizzoli@ieo.it

Abstract

Background: Relaxation practices might be helpful exercises for coping with anxiety and stressful sensations. They may be of particular utility when used in web-based interventions during periods of social isolation.

Objective: This randomized study aimed to test whether web-based relaxation practices like natural sounds, deep respiration, and body scans can promote relaxation and a positive emotional state, and reduce psychomotor activation and preoccupation related to the COVID-19 pandemic.

Methods: Participants were randomly assigned to one of three experimental conditions. Each condition was characterized by a single online session of a guided square breathing exercise, a guided body scan exercise, or natural sounds. The participants listened to one of the fully automated audio clips for 7 minutes and pre-post completed self-assessed scales on perceived relaxation, psychomotor activation, level of preoccupation associated with COVID-19, and emotional state. At the end of the session, qualitative reports on subjective experience were also collected.

Results: Overall, 294 participants completed 75% of the survey and 240 completed the entire survey as well as one of three randomly assigned interventions. Perceived relaxation, psychomotor activation/stress, and preoccupation related to COVID-19 showed a positive improvement after participants listened to the audio clips. The same pattern was observed for the valence and perceived dominance of the emotional state. The square breathing and body scan exercises yielded superior results compared to natural sounds in lowering perceived stress.

Conclusions: This study provides a novel insight that can guide the development of future low-cost web-based interventions to reduce preoccupation and stress in the general population.

International Registered Report Identifier (IRRID): RR2-10.2196/19236

Introduction

Background

To limit the spread of the SARS-CoV-2 virus, from mid-March to May 2020, Italy faced a strict lockdown [1]. People were forced to form a bubble with only members of their household, if they had any. This social isolation and resulting loneliness had an impact on health and mental well-being, thus affecting vital functions (eg, sleep quality), social connectedness, perceived support, and psychological status [2-4]. A recent rapid review of the effects of quarantine highlighted that many people were facing several negative cognitive and emotional problems, like confusion, poor concentration, irritability, insomnia, distress, frustration, and anger [5]. People were worried about the quarantine duration, insufficient information provision, economic problems, and stigma. The associated negative effects
had an impact on biopsychosocial functioning, sometimes even leading to depressive or posttraumatic stress symptoms [5,6].

The lack of a vaccine, the high chance of contagion, and the severity of symptoms—sometimes resulting in death—increased risk perception. In all high-risk situations, cognitive and rational thinking interacts with emotional appraisals, thus affecting people's state of mind: individuals feel vulnerable and may experience fear for themselves and for their loved ones [7,8]. Both physical and psychological dimensions are affected by the sense of uncertainty and the threat of contracting the virus. In a similar emergency condition, it is not uncommon that people may also experience a psychophysiological hyperactivation, thus paying excessive attention to bodily sensations and enhancing their perceptions [9]. Overall, COVID-19 and social isolation have led to negative side effects, causing widespread concern and psychophysiological reactions [5].

Starting from these premises, it appeared of primary importance to develop efficacious interventions aimed at reducing the possible preoccupations about COVID-19 and the associated psychophysiological activation.

In this regard, previous studies demonstrated that interventions based on natural sounds, respiration, and meditation helped individuals to alleviate the effects of stress by reducing physiologic arousal and restoring autonomic balance [10-13].

In fact, listening to natural sounds significantly reduces human stress processes [12-14]. For example, the stress recovery theory [15] posits that physiological (autonomic) and psychological stress are reduced within naturalistic environmental contexts because of human evolutionary adaptation to naturalistic stimuli. On the other hand, guided relaxation techniques represent widely used practices to produce a deep state of relaxation and enhance physical and emotional well-being. Deep breathing exercises and focusing attention on body perception (body scan) are two of the main techniques to reduce hyperarousal and achieve a more relaxed condition. The former may also be defined as “an efficient integrative body-mind training for dealing with stress, anxiety and psychosomatic conditions” [16]; it may help people slow their breathing, take in more oxygen, and reduce the use of shoulder, neck, and upper chest muscles, thus achieving better emotional balance and social adaptation [17]. On the other hand, body scans aim to focus attention on different parts of the body and encourage awareness of the body's sensations, such as pain, tension, warmth, or relaxation [18,19]. Applying these interventions to people who are forced into mandatory social isolation may help people become more aware of their mind-body condition and reduce negative effects.

We must add that, due to COVID-19 restrictions, face-to-face interventions were not possible, while web-based interventions represented an important opportunity. The usefulness of web-based interventions is supported by studies showing that online relaxation interventions led to significant results equal to in-person interventions [20,21]. However, a comparison of meditation techniques (eg, natural sounds, respiration, and body scan) in web-based interventions is still missing in the literature.

We chose to employ simple audio clips remotely delivered with auditory natural stimuli (water sounds), which already proved to be effective in facilitating the relieving of psychophysiological activation linked to a psychological stressor [14], along with two exercises with guiding instructions targeting body awareness and breath frequency control.

**Objective**

In this study, we aimed to test and compare the efficacy of three web-based interventions (based on natural sounds, breathing regulation, and body scan, respectively) to assess which intervention was the most effective for the target population. For this purpose, we tested the differences in the stress-reducing efficacy of three audio clips corresponding to the three relaxation practices (square breathing exercise, guided body scan exercise, and natural sounds). Specifically, we expected to find the following: (1) a decrease in the levels of psychomotor activation/stress and of preoccupation about COVID-19, as well as enhanced levels of relaxation and emotional state after exposure to all audio clips, and (2) that guided techniques (square breathing and body scan) would have a greater effect than natural sounds on the abovementioned dimensions.

**Methods**

**Participants and Procedure**

During the first week of May 2020, the invitation to take part in the study was published on Italian social media webpages (specifically, organic posts on WhatsApp, Facebook, LinkedIn, and Instagram). The readers were informed about the general aim of the study and that the study was conducted by researchers from the University of Milan. Potential participants were encouraged both to take part in the study and to share the invitation with their acquaintances. The invitation contained a link to the Qualtrics platform, where a more detailed description was available.

The eligibility criteria to take part in the study were the following: (1) being older than 18 years old, (2) being a proficient Italian speaker, (3) not having any impairment of auditory abilities, and (4) having an appropriate familiarity with computer literacy. Before taking part in the study, participants were asked to read and complete an online consent form.

Participation in the study consisted of three main parts: (1) a short questionnaire containing sociodemographic questions, a baseline anxiety evaluation (trait anxiety, anxiety for physical sensations, body vigilance), and a preintervention evaluation, (2) listening to a 7-minute audio clip, and (3) the postintervention evaluation. The estimated time for participating in the study (completing the three parts) ranged from 12-17 minutes.

Participation in the study was voluntary and participants were informed that they could withdraw from the study at any point in time. The research protocol followed the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials) V1.6 Guidelines [22] and the principles stated in the Declaration of Helsinki (59th WMA General Assembly, Seoul, 2008). The research protocol was approved by the Ethics Committee of the first author’s university on April 30, 2020, and registered with the following International Register Report Identifier (IRRID): PRRI-10.2196/19236 [23].
Overall, 294 participants completed at least 75% of the survey. The initial sample was mainly composed of female participants (women: n=216, 73.5%; men: n=78, 26.5%) and participants had a mean age of 39 years (SD 14.6, range 18-78). Overall, 87 (29.6%) of the participants who completed at least 75% of the survey had a chronic disease condition, while 207 (70.4%) had no chronic health issues. Participants were also asked to specify which chronic disease they have. The specific chronic conditions were organized into categories and are reported in Multimedia Appendix 1. The three initial groups were homogeneous in sociodemographic variables like gender ($\chi^2=91, P=.63$), age ($F[2,291]=1.59, P=.21$), educational level ($\chi^2=5.4, P=.86$), marital status ($\chi^2=1.14, P=.98$), and employment status ($\chi^2=3.5, P=.73$).

**Measures and Design**

**Baseline Questionnaire**

After completing the sociodemographic form, participants were asked to report if they had a chronic disease and how much the disease impacted their (perceived) vulnerability to COVID-19. Participants were asked to report their working situation, recent changes in occupational status due to COVID-19 restrictions, and if they had prior experience with relaxation techniques.

Participants were then asked to complete 3 self-assessed questionnaires aimed at measuring their current level of anxiety (trait anxiety), tendency to worry about physical signals and sensations (anxiety for physical sensations), and degree of attention paid to bodily feelings (body vigilance). For the assessment of these aspects, the following self-reported scales were used: the State-Trait Anxiety Inventory form-trait subscale (STAI-Y) [24,25], the Physical Concerns subscale of the Anxiety Sensitivity Index-3 (ASI-3) [26,27], and the Body Vigilance Scale (BVS) [28].

The STAI-Y is a questionnaire of 20 items scored on a 4-point Likert scale (from 1=not at all to 4=very much) [24,25], and assesses trait anxiety. The total score is calculated as the sum of all the items, after having computed the reverse scores of specific items. A higher total score indicates higher anxiety. The STAI-Y has a good internal consistency and it constitutes a reliable and valid tool for assessing anxiety symptoms in samples of healthy subjects [29]. In our sample, the Cronbach $\alpha$ was .84, indicating a good level of internal consistency.

The Physical Concerns subscale of the ASI-3 is a 6-item subscale and participants use a 5-point Likert scale (from 0=very little to 4=very much) to indicate worry related to specific physical sensations. The total score is the sum of the single items [30]. In our study, the Physical Concerns subscale had a high Cronbach $\alpha$ of .88, showing a good level of internal consistency.

Finally, the BVS is 4-item questionnaire (rated on a scale from 0=not at all like me to 10=completely like me) that asks how much attention one usually pays to body sensations. In the fourth item, participants had to rate their attention to 15 body sensations that are the core physical symptoms for panic attacks [31]. In this study, the internal consistency was .86.

**Randomization Procedure**

After completing questionnaires, participants were randomly assigned to one of the three experimental groups via the randomization procedure within Qualtrics. The randomization option was set to enroll the same number of subjects for each condition.

In each experimental condition, participants received a 7-minute audio clip aimed at promoting a state of awareness and relaxation. In the first experimental condition (square breathing), participants heard a recorded voice guiding the regulation of breathing frequency, with the aim of making every breath cycle (inhalation, hold breath, exhalation, hold breath) the same length (4 seconds). In the second experimental condition (body scan), participants listened to an audio clip with a voice that guided participant attention through every part of the body and gently requested that the listener notice and let go of tensions and unpleasant feelings. Both tracks were recorded by a trained mindfulness and yoga expert in collaboration with a psychotherapist and were pretested on 4 subjects to assess how easy the exercise is and its perceived effectiveness. In the final condition (natural sounds), participants were presented with a prerecorded audio clip of natural sounds (rain, water sounds). The scientific literature indicates that exposure to natural sounds may be an important stress reliever, thus promoting a state of relaxation [12,14].

All audio clips were preceded by instructions regarding the recommended location and body position for the exercises. More precisely, participants were invited to find a quiet room and to sit or lie down on a comfortable chair/sofa. After the instructions, participants were invited to click on the “play” key to start the audio clip.

**Pre-Post Evaluation**

As pre-post measures, before and after the audio stimuli, subjects were asked to self-rate their perceived relaxation level, perceived stress, and psychomotor activation degree (ie, motor/physical activity that is secondary to or dependent upon a psychic component and is mostly non–goal-directed [32]), how much they felt concerned about COVID-19, and to rate 3 specific features of their emotional state. Specifically, participants were requested to rate on 3 Visual Analogue Scales (VAS; 0=not at all, 10=completely) how relaxed they felt, how psychomotor-activated they felt, and how much thoughts related to COVID-19 scared them; furthermore, they completed the Self-Assessment Manikin (SAM) [33] for emotional states, which is a 3-item visual and nonverbal scale. Valence (from “unpleasant” to “pleasurable”), intensity/arousal (from “calm” to “excited”), and dominance (from “not in control” to “completely in control” of emotional state) are commonly used to quantify properties of the felt overall emotional state on 1-5 scales of images. To check if participants really listened to the audio clips, they were asked immediately after the recording if they heard the entire clip, a part of it, or nothing.

Finally, all the participants were asked to describe their personal experience and to provide suggestions for future changes. Specifically, participants were requested to write a short paragraph answering two open-ended questions about their experience and to provide suggestions for future changes.
personal experience with the exercise (ie, what they liked and what they would have changed).

**Data Analysis**

A detailed report of the data analysis approach can be found in [23]. Compared to the original protocol, we had a slightly lower sample size, which yielded a statistical power of .94 instead of the planned .95, for a medium effect size (ES). To reduce the chance of type I error, we choose to test and compare the efficacy of the audio clips with the mixed analysis of variance (ANOVA) and to report ES in the form of $\eta^2$; we also performed post hoc comparisons with the Tukey HSD test, instead of performing multiple comparisons with the $t$ test.

To test the difference in efficacy between audio clips, a one-way ANOVA on gain relaxation scores, with 3 groups and fixed effects, and no interaction, was performed. Subjects who stopped before the randomized exposure to the audio clips were excluded from this analysis. Furthermore, to assess group and time effects and their interaction, a $2 \times 3$ mixed-model ANOVA was performed for perceived relaxation scores, perceived stress/activation, and preoccupations related to COVID-19. We also performed nonparametric analysis on the items of SAM, as statistical assumptions for parametrical analysis were violated. In this case, we reported the ES in the form of Hodges-Lehmann ES.

Explorative analyses on the possible role of trait anxiety, anxiety for physical sensations, and body vigilance in moderating the effect of the audio clips were also carried out. All quantitative analyses were performed in SPSS (Version 26.0; IBM Corp).

Qualitative reports on subjective experiences were organized into different categories, to systematize the suggestions and preferences. Participants’ preferences on the web-based relaxation interventions were organized into three different categories (audio features, relaxing feeling, and awareness) to systematize participants’ experience. Finally, the reported suggestions for improving the quality of the interventions were grouped into three main themes: audio features, clarity of instructions, and length of the intervention.

**Results**

**Sample Description**

Overall, 328 participants registered on the survey website, and 294 gave written informed consent and completed more than 75% of the survey (meaning that they complete the initial questionnaire and were randomized into the three conditions), but they did not answer the postexposure questionnaire. Further, 12 of the 294 participants also stated that they did not listen to the audio clips. In total, 240 participants completed the entire survey and stated that they listened to the audio clips: 77 in square breathing group, 76 in the body scan group, and 87 in the natural sounds group. Among those who completed the entire survey (240), the mean age was 39.8 (SD 14.7, range 18-78) years, while 173 (72.1%) were female and 67 (27.9%) were male.

In regard to the differences in baseline variables (age, sex) between those who completed the entire survey and those who did not, no differences were found to be significant. Specifically, no differences emerged in gender ($\chi^2=1.28, P=.26$) and age ($F[1,292]=1.1, P=.87$).

In total, 70 (29.2%) participants had a chronic disease condition, while 170 (70.8%) had no chronic health issues. In addition, 15 (6.3%) reported to be highly limited in daily activities because of a disease, while 26 (10.8%) were partially influenced by health issues and 97 (40.4%) were not affected in daily activities by health issues. With regard to the differences between participants who completed the entire survey and those who did not, no differences emerged with the presence of a chronic disease ($\chi^2=1.11, P=.74$).

In regard to risk perception related to COVID-19, we found that participants perceived a small amount of risk of contracting COVID-19 (median 2=a little bit) and of having serious side effects of COVID-19 (median 2=a little bit). Participants reported being moderately worried about their employment status, health, and personal economic stability (median 3=quite a lot), while they were more preoccupied for their own family (median 4=a lot preoccupied). In regard to previous experience with relaxation practices, 101 participants (42.1%) had previous experience with relaxation techniques, while 139 (57.9%) had no previous experience with relaxation practices. Descriptive statistics of the final sample are reported in Table 1.

---

Table 1: Descriptive statistics of the final sample.
Table 1. Descriptive statistics of the sample.

<table>
<thead>
<tr>
<th>Sociodemographic variables</th>
<th>Frequency, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Educational level (N=240)</strong></td>
<td></td>
</tr>
<tr>
<td>Middle school diploma</td>
<td>12 (5.0)</td>
</tr>
<tr>
<td>High school diploma</td>
<td>87 (36.3)</td>
</tr>
<tr>
<td>Bachelor’s degree</td>
<td>33 (13.8)</td>
</tr>
<tr>
<td>Master’s degree</td>
<td>78 (32.5)</td>
</tr>
<tr>
<td>Postgraduate</td>
<td>30 (12.5)</td>
</tr>
<tr>
<td><strong>Marital status (N=240)</strong></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>68 (28.3)</td>
</tr>
<tr>
<td>In a relationship</td>
<td>62 (25.8)</td>
</tr>
<tr>
<td>Married or cohabitating</td>
<td>107 (44.6)</td>
</tr>
<tr>
<td>Widowed</td>
<td>3 (1.3)</td>
</tr>
<tr>
<td><strong>Employment status (N=240)</strong></td>
<td></td>
</tr>
<tr>
<td>I work</td>
<td>184 (76.7)</td>
</tr>
<tr>
<td>I do not work and I am not seeking work</td>
<td>42 (17.5)</td>
</tr>
<tr>
<td>I do not work but I am seeking work</td>
<td>14 (5.8)</td>
</tr>
<tr>
<td><strong>Work contract (N=228)</strong></td>
<td></td>
</tr>
<tr>
<td>Temporary</td>
<td>38 (16.7)</td>
</tr>
<tr>
<td>Permanent</td>
<td>110 (48.2)</td>
</tr>
<tr>
<td>Freelance</td>
<td>59 (25.9)</td>
</tr>
<tr>
<td>Other</td>
<td>21 (9.2)</td>
</tr>
<tr>
<td><strong>Work activity in the last week (N=228)</strong></td>
<td></td>
</tr>
<tr>
<td>Regular job activity</td>
<td>82 (36.0)</td>
</tr>
<tr>
<td>Less than regular job activity</td>
<td>50 (21.9)</td>
</tr>
<tr>
<td>No activity</td>
<td>75 (32.9)</td>
</tr>
<tr>
<td>Other</td>
<td>21 (9.2)</td>
</tr>
</tbody>
</table>

**Trait Anxiety, Anxiety for Physical Sensations, and Body Vigilance**

Descriptive statistics and reciprocal correlations of trait anxiety, anxiety for physical sensations, and body vigilance are provided in Tables 2 and 3.

Table 2. Descriptive statistics of the total scores of the questionnaire scores.

<table>
<thead>
<tr>
<th>Questionnaire</th>
<th>Mean (SD)</th>
<th>Median</th>
<th>Minimum-maximum</th>
<th>Correlation between total scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>STAI-Y&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>STAI-Y&lt;sup&gt;a&lt;/sup&gt;</td>
<td>44.5 (8.2)</td>
<td>43.5</td>
<td>28-68</td>
<td>1</td>
</tr>
<tr>
<td>ASI-3&lt;sup&gt;b&lt;/sup&gt;</td>
<td>18.3 (7.9)</td>
<td>18.5</td>
<td>0.3-39.7</td>
<td>N/A&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>BVS&lt;sup&gt;c&lt;/sup&gt;</td>
<td>11.4 (4.4)</td>
<td>10.5</td>
<td>6-30</td>
<td>N/A&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup>STAI-Y: State-Trait Anxiety Inventory form-trait subscale.
<sup>b</sup>ASI-3: Anxiety Sensitivity Index-3.
<sup>c</sup>BVS: Body Vigilance Scale.
<sup>d</sup>P<0.001.
<sup>e</sup>N/A: not applicable.

The three scores correlated moderately with each other, consistent with what has been found in previous literature [29,30,34]. Furthermore, there were small to moderate positive correlations with postexposure psychomotor activation/stress and preoccupation related to COVID-19. Finally, small negative correlations were found between trait anxiety and anxiety for physical sensations, and postexposure perceived relaxation.
Table 3. Descriptive statistics of the correlations between total scores.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Correlation with postexposure measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>STAI-Y&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Relaxation</td>
<td>−0.23&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Psychomotor activation/stress</td>
<td>0.48&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Fear related to COVID-19</td>
<td>0.19&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup>STAI-Y: State-Trait Anxiety Inventory form-trait subscale.

<sup>b</sup><i>P</i> < .001.

<sup>c</sup><i>P</i> < .05.

For all three scales, higher scores indicated a higher presence of anxiety; overall, we found participants to have higher scores compared to the reference values.

Specifically, for both men and women (men: mean 43.1, SD 7.9; women: mean 45.8, SD 8.3), the total score on the STAI-Y in our sample was slightly higher than the means of the reference values (men: mean 36, SD 9.7; women: mean 39.93, SD 11) [35].

In terms of the ASI-3 Physical Concerns subscale in our sample (men: mean 10.1, SD 7.9; women: mean 18.7, SD 7.8), we found participants had higher levels of anxiety for physical sensations compared to the reference values (men: mean 4.99, SD 4.28; women: mean 5.91, SD 4.78) [30]. Finally, for the BVS (men: mean 17.4, SD 7.9; women: mean 18.7, SD 7.8), participants reported higher body vigilance levels compared to the normative sample (men: mean 14.86, SD 6.92; women: mean 15.95, SD 9.71) [34].

The mean scores of the questionnaires were homogeneous between those with a chronic disease and those with no health issues, except for the ASI-3 scores (t[238]=−2.6, <i>P</i>=.01); specifically, the participants with a chronic condition had significantly higher anxiety for physical sensations (mean 12.5, SD 5) compared to those who did not have a chronic condition (mean 10.9, SD 3.9).

Furthermore, based on the total scores for each experimental group, we found that the three groups did not significantly differ in the initial scores of trait anxiety (F[2,237]=.02, <i>P</i>=.98), anxiety for physical sensations (F[2,237]=.09, <i>P</i>=.92), and body vigilance levels (F[2,237]=.17, <i>P</i>=.85).

Regarding differences between the initial sample and the final sample in psychological variables, we found no significant differences in the levels of trait anxiety (F[2,292]=1.2, <i>P</i>=.28), anxiety for physical sensations (F[2,292]=.83, <i>P</i>=.36), and body vigilance levels (F[2,284]=.25, <i>P</i>=.62).

Perceived Relaxation, Psychomotor Activation/Stress, and Thoughts Related to COVID-19

Descriptive statistics of all the assessed pre-post variables for each group are depicted in Table 4.

Table 4. Descriptive statistics of assessed pre-post variables for each group.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Square breathing (N=77)</th>
<th>Body scan (N=76)</th>
<th>Natural sounds (N=87)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
<td>Pre</td>
</tr>
<tr>
<td>Perceived relaxation, mean (SD)</td>
<td>47.2 (23.7)</td>
<td>66.4 (20.4)</td>
<td>47.4 (23.1)</td>
</tr>
<tr>
<td>Psychomotor activation/stress, mean (SD)</td>
<td>46.0 (25.3)</td>
<td>49.3 (22.7)</td>
<td>48.6 (26.2)</td>
</tr>
<tr>
<td>Thoughts related to COVID-19, mean (SD)</td>
<td>62.6 (26.6)</td>
<td>64.6 (27.8)</td>
<td>64.6 (26.9)</td>
</tr>
<tr>
<td>Valence, median (IQR)</td>
<td>5 (2)</td>
<td>5 (2)</td>
<td>5 (4)</td>
</tr>
<tr>
<td>Arousal, median (IQR)</td>
<td>4 (2)</td>
<td>5 (4)</td>
<td>5 (3)</td>
</tr>
<tr>
<td>Dominance, median (IQR)</td>
<td>6 (3)</td>
<td>7 (2)</td>
<td>5 (3)</td>
</tr>
</tbody>
</table>

Age was slightly positively correlated with fears associated with COVID-19 (<i>r</i>=.21, <i>P</i>=.001) and with ASI-3 scores (<i>r</i>=.25, <i>P</i>=.001).

As a primary analysis to compare efficacy between the audio clips, we performed a one-way ANOVA on postexposure relaxation scores. The analysis yielded nonsignificant differences between groups in perceived relaxation after exposure to the relaxing audio clips (F[2,237]=1.6, <i>P</i>=.21).

To test and compare the efficacy between the audio clips, a mixed ANOVA testing for group and time effects and for their interaction was performed for perceived relaxation scores, perceived stress/activation, and preoccupations related to COVID-19. All effects reported below are stated as significant at <i>P</i>&lt;0.001 unless otherwise specified.

A significant moderate main effect for time was found (F[1,237]=121.5, partial <i>η</i><sup>2</sup>=.34). Therefore, relaxation scores after the exposures were significantly higher than before listening to the audio clips (Figure 1). There was a nonsignificant effect of the group, indicating that the ratings from all three groups were similar (F[2,237]=1.15, <i>P</i>=.32, partial <i>η</i><sup>2</sup>=.01).
Thus, there was no overall difference in the scores of perceived relaxation between the square breathing, body scan, and natural sounds groups. Finally, results revealed a nonsignificant time × group interaction ($F[2,237]=.32$, $P=.73$, partial $\eta^2=.003$). A significant main effect of time was found ($F[1,237]=153.5$, partial $\eta^2=.39$). Hence, perceived psychomotor activation/stress scores after the exposures were significantly lower than before participants listened to the audio clips and the effect was moderate (Figure 2). Results also showed a nonsignificant time × group interaction ($F[2,237]=.06$, $P=.95$, partial $\eta^2=.0001$), while a small significant effect of the groups was found ($F[2,237]=3.6$, $P=.03$, partial $\eta^2=.03$). Contrasts between groups and post hoc tests revealed that the guided audio clips (ie, square breathing, body scan) yielded a significantly lower level of perceived stress compared to the natural sounds audio clip.

**Figure 1.** Perceived relaxation.

![Perceived relaxation chart](image1.png)

**Figure 2.** Psychomotor activation/stress.

![Psychomotor activation/stress chart](image2.png)

Specifically, those who were in the square breathing condition were significantly ($P=.04$) less stressed than the participants in the natural sounds group and those who were in the body scan condition rated their perceived level of stress as lower than the participants in the natural sounds group ($P=.03$).

A significant main effect for time was found ($F[1,237]=103.4$, partial $\eta^2=.3$). Therefore, the degree of fear of thoughts related to COVID-19 after the exposures was significantly lower than before exposure to the audio clips (Figure 3). There was a nonsignificant effect of the group, indicating that the ratings...
from all three groups were similar \( (F[2,237]=.95, P=.38, \text{partial } \eta^2=.01) \). Thus, there was no overall difference in the scores of preoccupations related to COVID-19 between the square breathing, body scan, and natural sounds groups. Finally, results showed a nonsignificant time \times \text{ group interaction} \( (F[2,237]=.37, P=.69, \text{partial } \eta^2=.003) \).

Figure 3. Thoughts related to COVID-19.

Emotional State (Valence, Arousal, and Dominance)

We performed a nonparametric analysis on the items of the SAM. To test if there was a difference in pre-post scores of the perceived emotional state, we ran, for each experimental group, the Wilcoxon signed-rank test on variables assessed before and after exposure to the audio clips and reported the ES in the form of Hodges-Lehmann ES.

For all groups, a significant positive improvement was observed on the dimensions of the pleasantness of the emotional state (valence) and perceived control over it (dominance), while a nonsignificant trend resulted for arousal.

With regard to the square breathing group, results showed a significant positive improvement in valence \( (T=640, z=3.16, P<.002; \text{ES}=0.5, 95\% \text{ CI 0-1}) \) and perceived dominance \( (T=477, z=3.12, P<.002; \text{ES}=0.5, 95\% \text{ CI 0-1}) \), while a nonsignificant trend was found for perceived arousal \( (T=707.5, z=1.84, P=.06; \text{ES}=0.5, 95\% \text{ CI 0-1}) \). For the body scan group, valence and dominance improved \( (T=1045, z=5.18, P<.001; \text{ES}=1, 95\% \text{ CI 1-1.5}) \) and \( T=393.5, z=2.88, P=.004; \text{ES}=0.5, 95\% \text{ CI 0-1} \), respectively, while arousal did not \( (T=516, z=2.25, P=.08; \text{ES}=0, 95\% \text{ CI 0-5}) \). Finally, the effects for the natural sounds group were \( T=943.5, z=3.70, P<.001; \text{ES}=0.5, 95\% \text{ CI 0-1} \) for valence, \( T=516.5, z=2.18, P=.03; \text{ES}=0, 95\% \text{ CI 0-0.5} \) for dominance, and \( T=457, z=.35, P=.728; \text{ES}=0, 95\% \text{ CI 0-1} \) for arousal.

Overall, all three variables showed a positive improvement following the audio clips, while no significant differences between the groups emerged for perceived relaxation and for COVID-19-related preoccupation, while significant differences emerged between groups in perceived psychomotor activation/stress, where both the guided audio clips yielded a decrease in perceived stress compared to natural sounds.

To assess between-groups differences, we performed the Kruskal-Wallis test as a global test on all three groups. None of the comparisons between the groups yielded significant differences; thus, we did not proceed with paired comparisons.

Specifically, valence yielded \( \chi^2=4.6, P=.1 \) pre-exposure and \( \chi^2=1.17, P=.56 \) postexposure, while arousal scores gave \( \chi^2=2.56, P=.28 \) pre-exposure and \( \chi^2=2.17, P=.34 \) postexposure, and analysis of dominance showed \( \chi^2=0.18, P=.91 \) pre-exposure and \( \chi^2=1.25, P=.54 \) postexposure.

Role of Baseline Anxiety on Intervention Efficacy

To test if baseline levels of anxiety moderated the efficacy of the audio clips, we performed mixed ANOVA testing for moderation effects of the questionnaire scores for STAI-Y, ASI-3, and BVS with group and time. Correlations between questionnaires and postexposure scores are reported in Table 4.

Overall, there was a significant moderation effect between STAI-Y scores and time \( (F[1,234]=7.22, P=.008, \text{partial } \eta^2=.03) \) on perceived relaxation and the same pattern was observed for the interaction with time \( (F[1,234]=4.4, P=.04, \text{partial } \eta^2=.02) \) on psychomotor activation/stress, suggesting a higher efficacy for those participants who were less anxious at baseline.

In regard to the ASI-3, we found that there was a significant moderation effect between ASI-3 levels and time \( (F[1,234]=4.3, P=.04) \).
In such a situation, delivering helpful interventions to manage stress and anxiety becomes increasingly difficult. However, web-based interventions have been increasingly adopted in clinical practice for facilitating psychological assessment and enabling the delivery of the same treatment to many subjects concurrently [36].

This randomized study compared the effects of remotely delivered interventions using natural sounds, deep breathing, and meditation on perceived relaxation, psychomotor activation, level of preoccupation with COVID-19, and emotional state, as well as on people’s experiences with these techniques. Indeed, a comparison of natural sounds, respiration, and body scan meditation techniques delivered as web-based interventions was missing in the literature.

In accordance with our first prediction, results showed that all three techniques produced positive effects on perceived relaxation, stress, and preoccupation related to COVID-19. Specifically, we found that perceived relaxation levels, psychomotor activation/stress, and disturbing thoughts related to COVID-19 significantly improved after exposure to the three audio clips, with a moderate effect. Starting from this evidence, we concluded that the audio clips were effective in inducing a calmer psychological state. Our findings are consistent with the results of another study that aimed to reduce anxiety and depression in patients with COVID-19, which was a web-based intervention containing breath relaxation training, a mindfulness body scan, and behavioral techniques that significantly improved mood disturbance symptoms [37].

Based on between-groups differences, results showed significant differences only on perceived psychomotor activation/stress, where guided exercises (ie, square breathing and body scan) were more effective than the natural sounds audio clip. No other differences in efficacy between groups were observed. These results partially confirmed our second hypothesis, as we obtained enhanced efficacy only on the dimension of stress and not for perceived relaxation.

We concluded that for this brief web-based relaxation intervention, the three audio clips were effective at improving psychological adjustment. We speculated that, even if the audio clips involved different psychocognitive activities (ie, regulating breath frequencies, bringing awareness to body sensations, and simply listening to natural sounds), for brief relaxation experiences, these different processes did not have significant variances in efficacy. Thus, for the purpose of simple and brief relaxation interventions, all three can be effectively applied. Consistently, Jain and colleagues [38] conducted a randomized controlled trial on the differences among mindfulness meditation, relaxation training, and a waitlist control group in reducing distress; no statistical differences were found between the two treatment groups, even if both interventions significantly improved positive emotional state and reduced distress levels.

Furthermore, regarding differences between groups, we posit that the guided exercises yielded a greater effect than natural sounds for perceived stress because they ask users to bring their attention to body sensations and breathing frequency, which promotes a calm state. A systematic review of the importance of guidance during web-based interventions showed guided interventions are more effective than nonguided ones [39]. Such instructions may not change the degree of muscle contraction or body perception, but rather they can encourage a state of pleasant awareness, which is in line with the purpose of meditation practices. In fact, answering the final open-ended questions, participants reported that the audio clips allowed them to have a break and think about their breathing rhythm, thus regulating it. Others had the opportunity to observe their “inner world” becoming calmer. Moreover, audio features may vary the pleasantness of the audio clips; most of the participants reported positive feelings and a sense of relaxation in response to the guided body scan and square breathing interventions. Nevertheless, other participants did not like the tone of the voice and became bored, hyperactive, or upset.

In regard to participants’ emotional state, assessed through the SAM, we found that all the audio clips were effective in improving the valence (pleasure) of the emotional experience.
and the perceived dominance over it, while no significant results were found on the dimension of degree of arousal related to the emotional state. On the basis of these latter results, we speculated that the relaxing audio clips were more effective at enhancing the pleasantness and sense of control of one’s emotions, rather than the physical activation linked to such an emotional state.

Limitations and Future Directions
This study has some limitations. First, it was based on single-session guided interventions. Thus, we cannot assess differences in the efficacy of more prolonged exposure to relaxation sessions and correct for the effect of training or habituation. Future studies might adjust the length of the audio clips in light of the participants’ reports and/or propose repeated exposure to the audio clips.

Second, in this study, we did not include a neutral control group and we were not able to assess the efficacy of the techniques employed in the audio clips compared to no treatment. We could only assess the differences between the three active groups. In addition, the lack of a neutral control group meant we could not compare the effect of simply remaining in a passive resting state for 7 minutes with the effect of the interventions we delivered.

Third, being a remote intervention, we could not control for participants listening to the audio clips. We inserted a self-report check at the end of the audio clips that asked if participants really listened to the audio clips, but we did not have objective measures of the behavior and degree of attention of participants while listening. Similarly, in this study we employed only self-reported measures; as our intervention was delivered in the context of social distancing, other approaches were not feasible to implement. Future studies with more objective measures and psychophysiological variables (such as skin conductance) would strengthen the evidence of the efficacy of these techniques.

Fourth, we did not perform intention-to-treat analysis on our sample because we could not analyze participants who stopped before the audio clips or during the audio clips, since they had completed only the preintervention questionnaire. Therefore, analyses were performed only on participants who also completed the postintervention questionnaire. However, we checked for the presence of differences in the main psychosocial variables (age, gender, the presence of a chronic disease, anxiety scores) between the initial sample and the final sample, and we found no significant differences.

Finally, considering the possibility of applying such interventions to a clinical population, future studies might further explore the relationship between trait anxiety and the efficacy of the techniques. In this study, we could not perform a complex moderation model, nor could we draw conclusions on how different levels and different types of anxiety might shape the efficacy of the techniques. Preliminary evidence points to the potential moderating effect of baseline anxiety on the ES of relaxation interventions.

Conclusions
Despite these limitations, our study demonstrated that even a very brief online intervention that is based on this approach can contribute to a significant stress reduction. This study provides a novel insight that can orient the development of future low-cost web-based interventions to reduce preoccupation and anxiety in the general population. Future studies might also assess and compare the efficacy of these approaches in clinical protocols for patients with anxiety and hyperarousal.

Acknowledgments
SFMP is a PhD student in Medical Humanities in the European School of Molecular Medicine (SEMM). The authors acknowledge the Fondazione IEO-CCM. Authors thank Riccardo Pisati for the recording of the audio clips.

Conflicts of Interest
None declared.

Editorial Notice
This randomized study was not registered. The authors explained that their study “involved only volunteers and not patients.” The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials because of the publication of the protocol before enrolment. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness.

Multimedia Appendix 1
Chronic diseases reported by the participants.
[DOCX File, 14 KB - mental_v7i12e22757_app1.docx ]

Multimedia Appendix 2
CONSORT-eHEALTH V1.6.2 checklist.
[PDF File (Adobe PDF File), 103 KB - mental_v7i12e22757_app2.pdf ]

References
1. World Health Organization. Coronavirus. URL: https://www.who.int/health-topics/coronavirus#tab=tab_1 [accessed 2020-11-29]


Abbreviations

ANOVA: analysis of variance  
ASI-3: Anxiety Sensitivity Index-3  
BVS: Body Vigilance Scale  
CONSORT: Consolidated Standards of Reporting Trials  
ES: effect size  
SAM: Self-Assessment Manikin  
STAI-Y: State-Trait Anxiety Inventory form-trait subscale  
VAS: Visual Analogue Scale
Measuring COVID-19 Related Anxiety in Parents: Psychometric Comparison of Four Different Inventories

Christian Kubb1, MSc; Heather M Foran1, PhD
Health Psychology Unit, Institute of Psychology, Universität Klagenfurt, Klagenfurt, Austria

Corresponding Author:
Christian Kubb, MSc
Health Psychology Unit
Institute of Psychology
Universität Klagenfurt
Universitätsstr 65-67
Klagenfurt, 9020
Austria
Phone: 43 463 2700 1631
Email: christian.kubb@aau.at

Abstract

Background: The COVID-19 outbreak and the measures to contain the global pandemic can have an impact on the well-being and mental health status of individuals. Parents of young children are particularly at risk for high levels of parental stress due to the current public health crisis, which can impact parenting behaviors and children’s well-being. Although different initial scales have been developed to measure COVID-19–related anxiety, they have not yet been tested sufficiently in parent samples. A brief measure of COVID-19–related anxiety is necessary for both quick assessment in practice and in larger epidemiological studies of parents.

Objective: The purpose of this study is to compare the distributions, validities, and reliabilities of four different COVID-19 anxiety scales: Fear of COVID-19 Scale, Coronavirus Anxiety Scale, Pandemic Anxiety Scale, and one subscale of the COVID Stress Scales. Based on the psychometric properties of these scales, we aim to provide recommendations for a brief unidimensional inventory to assess COVID-19–related anxiety among parents.

Methods: A cross-sectional web-based survey of 515 German-speaking parents (465 mothers, 90.3%) with at least one child aged 0-6 years was conducted during a 6-week period (June 29 to August 9, 2020). Half of the parents were recruited via Facebook parenting groups, while the other half were recruited through childcare centers. We psychometrically tested 25 items on COVID-19–related anxiety using the framework of classical test theory, including item analysis, correlational analysis of family variables, and exploratory factor analysis. Moreover, an item response theory approach was applied to estimate item discriminations and item difficulties.

Results: Based on the psychometric properties, three items of the Pandemic Anxiety Scale were identified as a single unidimensional factor. The adapted scale demonstrated acceptable internal consistency (α=.79), moderate to high item discrimination, strong positive intercorrelation with two other COVID-19 anxiety scales, and a small positive association with parenting stress. Mothers and fathers did not differ in total scores (t_{513}=−0.79, P=.42).

Conclusions: Factor analysis suggests that existing COVID-19–related anxiety scales measure different latent constructs of anxiety. Furthermore, all scales showed only small to moderate correlations with trait health anxiety, suggesting that COVID-19–related anxiety is distinct from general health anxiety. The adapted “disease anxiety” subscale of the Pandemic Anxiety Scale is an economical measure for assessing COVID-19–related anxiety in parents. Directions for future research are outlined.

(JMIR Ment Health 2020;7(12):e24507) doi:10.2196/24507

KEYWORDS
COVID-19; coronavirus; anxiety; parents; parenting; scale; inventory; well-being; mental health; stress
Introduction

Background

In December 2019, patients with unusual cases of pneumonia in Wuhan City, China, were reported [1]; later, in January 2020, this pneumonia was identified as being caused by the pathogen SARS-CoV-2 [2]. Confirmed individual cases and clusters were subsequently observed in almost all countries worldwide [3,4]. On March 11, 2020, the World Health Organization [5] declared the rapidly spreading COVID-19 to be a global pandemic. Shortly afterward, many countries worldwide went into lockdown, with measures of closed borders, social distancing, and quarantine orders, to curb the spread of the novel coronavirus. To date (October 14, 2020), over 38 million confirmed cases of COVID-19 and 1,080,000 deaths have been reported by Johns Hopkins University [4].

Despite the large discrepancies in infection and death rates within Europe, the socioeconomic impact of the adopted measures for disease control has been more or less equal in all countries. Since then, notable effects on physical and mental health have been widely described, including symptoms of anxiety, stress, and depression [6-8]. The danger of a global mental health crisis was discussed early in the pandemic due to manifold restrictions and changes in daily living: confrontation with death, fear of contracting the virus or transmitting it to others, isolation during quarantine, financial strain, job loss, anger against measures, closed public facilities, and visitation bans to relatives [9,10]. These severe life events can exacerbate existing mental disorders and increase the risk of new incidences of stress-related disorders [11,12].

In particular, the pandemic outbreak has confronted families with unprecedented and immediate challenges in their daily routines. Psychosocial and economic changes have challenged family life. In the United States alone, the number of employed people decreased by 20 million between February and May 2020 [13]. Increased levels of stress and anxiety were found in parents [14-19], with substantial strains in compatibility with intensive child care responsibilities and employment while facilities were closed [7]. As a result, additional problems such as homeschooling and crowded households have emerged and placed additional strain on parents [7,20,21]. During the COVID-19 crisis, parents are at increased risk for suffering from parental burnout [22], with growing evidence that mothers, children, and immigrants from households with low socioeconomic status are most affected by negative mental and physical health sequelae [23-27].

Perceived danger and fear of health consequences can be expected during pandemic outbreaks. Studies on the 2009 swine influenza pandemic showed a significant increase in health-related fears [28-31]. Generally, health anxiety is characterized by excessive fears and worries about having, contracting, or developing a serious disease. Bleichhard and Hiller [32] found that a point prevalence of about 6% of the German population suffers from health anxiety disorders, which is in line with results from other population samples [33,34]. Parents may play a crucial role in transmitting health beliefs and related behavior to their children [35]. Contemporary research has provided some evidence of the interconnected anxiety of parents and their children during pandemics [17,36], which could prompt a parent-based intervention approach to simultaneously address fears and burden in families.

In this context, less is known about the relationship between health anxiety and family variables (eg, functioning or intimate partner relationship distress). Although research on general anxiety provides evidence of associations with relational variables [37-40], they are rarely assessed together with health anxiety inventories. Anxiety-related psychopathologies often but not necessarily include worries about health [41]. Although a few studies have examined the connection between health anxiety of children and family functioning [42,43], less research has focused on health anxiety in parents, particularly in the general population. Understanding the contribution of family and couple functioning in relation to health anxiety in the general population is particularly relevant in the context of the current pandemic.

The current body of literature for measuring parental distress, parental burden, or the quality of the caregiver-child relationship is extensive, with a long history of research [44,45]. However, there is a need to assess anxiety related to the COVID-19 pandemic among parents. To our knowledge, only one scale was developed to assess parenting during a pandemic [46]; however, it has not yet been validated in a parent sample. In terms of assessing COVID-19–related stress and anxiety, Ransing and colleagues [47] provided an overview of recently published scales. They identified five different scales in the literature up to May 15, 2020: the Fear of COVID-19 Scale (FCV-19S) [48], the Coronavirus Anxiety Scale (CAS) [49], the Obsession with COVID-19 Scale [50], the COVID Stress Scales (CSS) [51], and the Perception of Threat from COVID-19 [52]. Validation of these scales among parent samples has yet to be conducted as far as we are aware.

Stressors and needs within families with young children may differ from those of other household types, where childcare obligations or homeschooling can be serious challenges that are unique to parents [27,53]. Calls for initial measures for enhancing family-based interventions and sheltering vulnerable groups have been noted [21,54,55]. From a public health perspective, valid measurements for early detection of COVID-19–related anxiety among parents at risk may be useful for epidemiological studies as well as to identify parents in need of early intervention support through health care services or other social services [56,57].

Objectives

The objectives of the current study are twofold: (1) to compare the distribution, validity, and reliability of four different COVID-19 anxiety and distress scales, namely the FCV-19S, CAS, PAS, and the Danger subscale of the CSS; (2) to perform an exploratory factor analysis (EFA) of all four scales to identify the most promising brief unidimensional scale that can be used efficiently for research and practice among parent samples. The included items should have sufficient variance in the sample, ability to detect symptoms in both mothers and fathers, moderate to high item discrimination, and associations with related constructs. We hypothesized that the COVID-19 anxiety
measures would moderately correlate with trait health anxiety, show small to moderate correlations with other measures of family functioning (ie, parental stress and general family functioning), and show weak associations with intimate partner relationship satisfaction in a cross-sectional sample based on past literature on anxiety and families reviewed above.

Methods

Participants
A total of 1526 individuals started the web-based survey, resulting in a final sample of 515 parents after data cleaning (see below). Participants were predominantly mothers (465/515, 90.3%) with a university degree (307/515, 59.6%). Most of the 515 participants had German (312, 60.6%) or Austrian (177, 34.4%) citizenship. A share of 19.8% participants (102/515) came from Carinthia. The parents were aged 18-58 years (mean 34.95 years, SD 5.39). The majority were employed (285/515, 55.3%) or worked in the household (180/515, 34.9%). In terms of family status, 27.4% participants (141/515) were unmarried and 68.5% (353/515) were married. At the time of the survey, 94.4% participants (486/515) were in a partnership. Four participants (0.8%) stated that they had confirmed COVID-19 infection. More detailed demographics are described in Table 1.

Table 1. Sociodemographic characteristics of the study participants (N=515), n (%). The average age of the participants was 34.95 years (SD 5.39).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>465 (90.3)</td>
</tr>
<tr>
<td>Male</td>
<td>50 (9.7)</td>
</tr>
<tr>
<td>Nationality</td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>177 (34.4)</td>
</tr>
<tr>
<td>Germany</td>
<td>312 (60.6)</td>
</tr>
<tr>
<td>Switzerland</td>
<td>6 (1.2)</td>
</tr>
<tr>
<td>Other</td>
<td>20 (3.8)</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
</tr>
<tr>
<td>Unmarried</td>
<td>141 (27.4)</td>
</tr>
<tr>
<td>Married</td>
<td>353 (68.5)</td>
</tr>
<tr>
<td>Divorced</td>
<td>21 (4.1)</td>
</tr>
<tr>
<td>In a relationship</td>
<td>486 (94.4)</td>
</tr>
<tr>
<td>Number of children</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>206 (40.0)</td>
</tr>
<tr>
<td>2</td>
<td>239 (46.4)</td>
</tr>
<tr>
<td>3</td>
<td>55 (10.7)</td>
</tr>
<tr>
<td>4</td>
<td>10 (1.9)</td>
</tr>
<tr>
<td>5 or more</td>
<td>5 (1.0)</td>
</tr>
<tr>
<td>Currently pregnant</td>
<td>22 (4.3)</td>
</tr>
<tr>
<td>Educational level</td>
<td></td>
</tr>
<tr>
<td>No degree</td>
<td>1 (0.2)</td>
</tr>
<tr>
<td>Lower secondary</td>
<td>30 (5.8)</td>
</tr>
<tr>
<td>Higher secondary</td>
<td>68 (13.2)</td>
</tr>
<tr>
<td>High school</td>
<td>109 (21.2)</td>
</tr>
<tr>
<td>University</td>
<td>307 (59.6)</td>
</tr>
<tr>
<td>Employment</td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>285 (55.3)</td>
</tr>
<tr>
<td>Working in household</td>
<td>180 (34.9)</td>
</tr>
<tr>
<td>Other</td>
<td>50 (9.8)</td>
</tr>
</tbody>
</table>

*aReflects the number and percentage of participants answering "yes" to this question.
Survey Procedure

Participants were recruited on the web during a 6-week period (June 29 to August 9, 2020), mainly via social media in parenting or child-related Facebook groups and message boards. The evaluation of the HTTP referers showed that half of the final sample (50.48%) found the survey through Facebook. In addition, more than 4000 kindergartens and parent-child centers in Germany, Austria, and Switzerland were contacted electronically and asked to distribute the link of the web-based survey. Participants were required to be ≥18 years of age and the parent of at least one child aged between 0 and 6 years; participants were excluded if they or their children had any chronic or acute diseases. This study was part of a larger study focused on understanding parental search behaviors for health information, and the inclusionary criteria were required for the purposes of the overall study. The survey took an average of 22 minutes to complete. Participants were offered the chance to win ten vouchers in the amount of €10 (US $11.91) at the end of the survey. They were provided with a separate link to enter the raffle that could not be connected with the data from the study, which was anonymous. The study was approved by the Institutional Review Board of the University of Klagenfurt. Informed consent was obtained before data were collected.

Data Cleaning

Prior to data analysis, the dataset of 1526 entries was cleaned in two waves. First, 978 participants (64.1%) failed to fill out the entire questionnaire. The majority of these participants ended the survey during or immediately after the demographics section (544/978). Only five participants dropped out during the COVID-19 items, which were presented on the last five pages. Second, 53 participants were excluded because they either stated that they had no children (n=5), or their youngest child was older than six years (n=48). This resulted in a final sample of 515 parents.

Translation

All COVID-19 scales were only available in English and were thus translated into German by one author (CK) and an American Studies student using the translation-back-translation procedure. To further ensure quality, two psychology doctoral students subsequently checked the correctness of the translations independently. Based on this check, some minor changes were made to individual items to improve readability and precision. Some of the questionnaires used different spellings for COVID-19 (eg, coronavirus-19, COVID-19, virus). We decided to use the term “Covid-19” consistently, as this spelling is common in German-speaking countries. All translated questionnaires can be found in Multimedia Appendix 1.

Data Analysis Strategy

All descriptive and correlational analyses were performed using SPSS version 25 (IBM Corporation). A two-tailed $P$ value ≤ 0.05 was considered statistically significant. We followed the Cohen [58] interpretation guidelines for Pearson correlations, with $r=0.10$ considered to be a small correlation, $r=0.30$ a medium correlation, and $r=0.50$ a large correlation. For the EFA model fit, the root mean square error of approximation (RMSEA) was calculated with JASP version 0.11.1 [59]. The RMSEA fit index was interpreted according to Browne and Cudeck [60].

In addition, we applied an item response theory (IRT) approach to provide measures for item discriminability and difficulty. The graded response model by Samejima [61] was used. This model is an extension of the two-parameter logistic model that is applicable for ordered polytomous variable data (eg, Likert scales). A sample size of N=500 is recommended for accurate parameter estimation [62]. Marginal maximum likelihood estimation [63] was used for estimation of the parameters. We calculated item discrimination (alpha) and item difficulty (beta) for each scale separately based on the initial proposed unidimensional factor structures of COVID-19–related anxiety scales. As a result, the PAS was only considered with the subscale “disease anxiety” for the IRT analysis. All other scales were included in the analysis in their entirety, as they were proposed to measure one factor. According to the guidelines of Baker and Kim [64], we interpreted alpha values ≤0.64 as low item discrimination, values between 0.65 and 1.34 as moderate, and values ≥1.35 as high. IRTPRO software was used to estimate the parameters of the IRT models [65].

Measures

The One-Item Covid-Fear Scale

The One-Item Covid-Fear scale (Covid-F) was developed for this study. The item assessed fear of COVID-19 (“How do you rate your fear of the coronavirus (Covid-19)?”) based on a 10-point Likert-scale (1-10), with a higher score indicating greater fear.

The FCV-19S

The FCV-19S, developed by Ahorsu et al [48], is a 7-item inventory using a 5-point Likert scale with scores between 7 and 35. The higher the score, the higher the fear of COVID-19. The scale showed good internal consistency (α=0.82). Moderate correlations with depression ($r=0.42$) and anxiety ($r=0.51$) were reported. Validation studies were performed with samples from Russia and Belarus [66], Italy [67], Bangladesh [68], Turkey [69], Saudi Arabia [70], Israel [71], India [72], Greece [73], the United States [74], Spain [75], Japan [76], Cuba [77], and Mexico [78]. Overall, the FCV-19S has shown robust psychometric properties across validation studies; the findings predominantly support a unidimensional factor structure.

The CAS

The CAS, developed by Lee [49], is a short 5-item screening instrument that assesses common physiological anxiety symptoms related to COVID-19 over the previous two weeks: dizziness, sleep disturbance, tonic immobility, appetite loss, and abdominal distress. Confirmatory factor analysis (CFA) indicated a single factor structure of the coronavirus anxiety construct. The scale showed excellent internal consistency (α=0.93) in the initial validation study. Scores can range between 0 and 20. Associations were found with COVID-19 diagnosis, functional impairment, and maladaptive coping strategies, but not with history of anxiety. The suggested cutoff score (≥9) identified burdened adults with 90% sensitivity and 85% specificity for dysfunctional levels of COVID-19–related anxiety.
anxiety. Validation studies were performed in Turkey [79] and Bangladesh [80].

**The PAS**
The PAS, developed by McElroy et al [81], is a 7-item scale for assessing anxiety experienced during a pandemic. In a validation study, 4793 parents with children aged between 4 and 16 years were included. Total scores can range between 0 and 28. EFA revealed a two-factor solution with four items regarding contracting and transmitting the virus (disease anxiety) and three items concerning worries about consequences of the pandemic (consequence anxiety). This factor structure was verified with CFA in the other half of the sample. Internal consistency across all items was acceptable ($\alpha=70$). Moderate correlation was found with a subset of items of the Depression, Anxiety, and Stress Scale [82]. The PAS was also tested in a sample of medical students and residents in the United Kingdom [83].

**The CSS**
The CSS, developed by Taylor et al [51], is a 36-item inventory that consists of five subscales: danger and contamination fears, socioeconomic consequences, xenophobia, traumatic stress symptoms, and compulsive checking related to COVID-19. Initially, the scale was validated in a Canadian and US sample. The internal consistencies varied from $\alpha=.83$ to $\alpha=.95$ for the different subscales, and the subscales were moderately to highly correlated.

In the initial 6-factor solution, the scales of danger and contamination were divided into 2 subscales; however, due to high cross-loadings, they were combined a posteriori. For our study, we only used the 6 items of the danger subscale. This subscale includes relational items that seem especially relevant for parents (eg, “I am worried that I can’t keep my family safe from the virus”). Further studies have been conducted in an additional US and Canadian sample [84] and in the Philippines [85].

**Validity Measures**

**The Modified Short Health Anxiety Inventory**
The Modified Short Health Anxiety Inventory (mSHAI), developed by Bailé et al [86], is a 14-item test instrument for the measurement of trait health anxiety as a single construct. A meta-analysis has shown that the original Short Health Anxiety Inventory by Salkovskis et al [87] is a valid, reliable, and useful instrument for assessing health anxiety in clinical and non-clinical samples [88]. In contrast to the original inventory by Salkovskis et al [87], the mSHAI has a simpler response format on a 5-point Likert scale. Total scores range between 0 and 56. The mSHAI showed excellent internal consistency in our sample ($\alpha=.94$). We expected COVID-19 anxiety scales to only weakly or moderately correlate with this measure of trait health anxiety because the COVID-19 pandemic is uniquely impacting parents, who otherwise would have low levels of health anxiety.

**The Couple Satisfaction Index**
The Couple Satisfaction Index (CSI), developed by Funk and Rogge [89], is a widely used measurement in research and practice for relationship satisfaction. The basic version contains 32 items (CSI-32); however, the short version with 16 items (CSI-16) demonstrates strong psychometric properties and precision in detecting couple satisfaction compared to other measures. Total scores can range between 0 and 81. Scores below the recommended cutoff score of 51.5 indicate substantial relationship distress. In our sample, the Cronbach coefficient of the CSI-16 was excellent, with $\alpha=.97$.

**Parental Stress Scale**
The Parental Stress Scale (PSS), developed by Berry and Jones [90], is an 18-item scale for assessing child-related burden in mothers and fathers. Scores range from 18 to 90. A higher score indicates a higher level of parental stress. Factor analysis identified four dimensions: parental rewards, parental stressors, loss of control, and parental satisfaction. Despite some discord in the literature about the initial factor structure [91], the PSS is a psychometrically robust and widely used measurement in both clinical and nonclinical samples. The internal consistency in the present sample was good, with Cronbach $\alpha=.86$.

**The General Functioning Scale of the Family Assessment Device**
The General Functioning Scale (GFS) [92] is a 12-item subscale of the McMaster Family Assessment Device [93] to assess family functioning. Parents evaluate statements about family life on a 4-point Likert scale. The total score is then divided by 12 to give the overall functional level. A score of 1.0 indicates healthy family functioning, while a score of 4.0 represents extremely poor family functioning. Byles et al [92] recommended 2.17 as a cutoff score to detect dysfunctional families. The measure is correlated with a variety of other measures of problems, including alcohol abuse, marital distress, partner violence, and parental separation. In our sample, the GFS showed good internal consistency ($\alpha=.87$).

**Results**

**Sample Descriptive Statistics and Gender Differences**
Overall, 27.4% respondents (133/486) scored below the distress cutoff of the CSI-16, indicating couple dissatisfaction. In the measure of family functioning, 17.5% respondents (90/515) were identified as reporting problematic family functioning. There was a significant difference in trait health anxiety (mSHAI) scores for mothers and fathers ($t_{513}=2.30, P=.02$); mothers had higher scores. However, there were no differences between mothers and fathers regarding COVID-19–related fear (Covid-F) ($t_{513}=0.49, P=.62$). No significant gender differences were found for couple satisfaction, parenting stress, or family functioning (all $P<.05$).

Table 2 shows the range, mean, SD, score range, skewness, and kurtosis for all scales. With the exception of the PAS, all COVID-19–related scales were right skewed. None of the scales were normally distributed as assessed by Shapiro-Wilk test, $P<.001$. In particular, the CAS showed the least variance. More than three quarters of the participants had zero variance (no endorsed symptoms) on this scale (402/515, 78.1%).
Next, independent sample t tests were conducted to compare the total scores of the COVID-19 scales between mothers and fathers. There were no significant differences in scores on the CAS (t\(_{513}=1.03, P=.30\)), PAS (t\(_{513}=-0.28, P=.77\)), or CSS-D (t\(_{513}=-0.08, P=.93\)). There was a significant difference in scoring for the FCV-19S, with higher scores among mothers than fathers (t\(_{513}= 2.98, P=.003\)).

### Table 2. Descriptive statistics of the scales (N=515 for all scales, except CSI-16, for which n=486).

<table>
<thead>
<tr>
<th>Scale</th>
<th>Range of the scale</th>
<th>Mean (SD)</th>
<th>Score range</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Shapiro-Wilk test P value of the Shapiro-Wilk test</th>
</tr>
</thead>
<tbody>
<tr>
<td>mSHAI(^a)</td>
<td>0-56</td>
<td>13.99 (10.66)</td>
<td>0-56</td>
<td>1.12</td>
<td>1.47</td>
<td>0.91 (&lt;.001)</td>
</tr>
<tr>
<td>Covid-F(^b)</td>
<td>1-10</td>
<td>4.10 (2.25)</td>
<td>1-10</td>
<td>0.50</td>
<td>-0.67</td>
<td>0.93 (&lt;.001)</td>
</tr>
<tr>
<td>FCV-19S(^c)</td>
<td>7-35</td>
<td>13.39 (4.96)</td>
<td>7-35</td>
<td>0.91</td>
<td>0.80</td>
<td>0.93 (&lt;.001)</td>
</tr>
<tr>
<td>CAS(^d)</td>
<td>0-20</td>
<td>0.67 (1.80)</td>
<td>0-15</td>
<td>3.96</td>
<td>18.89</td>
<td>0.43 (&lt;.001)</td>
</tr>
<tr>
<td>PAS(^e)</td>
<td>0-28</td>
<td>10.63 (5.29)</td>
<td>0-25</td>
<td>0.13</td>
<td>-0.48</td>
<td>0.98 (&lt;.001)</td>
</tr>
<tr>
<td>CSS-D(^f)</td>
<td>0-24</td>
<td>6.07 (5.47)</td>
<td>0-24</td>
<td>0.86</td>
<td>0.01</td>
<td>0.90 (&lt;.001)</td>
</tr>
<tr>
<td>CSI-16(^g)</td>
<td>0-81</td>
<td>58.81 (17.12)</td>
<td>3-81</td>
<td>-1.03</td>
<td>0.50</td>
<td>0.91 (&lt;.001)</td>
</tr>
<tr>
<td>PSS(^h)</td>
<td>18-90</td>
<td>39.21 (8.99)</td>
<td>18-73</td>
<td>0.33</td>
<td>0.06</td>
<td>0.99 (.003)</td>
</tr>
<tr>
<td>GFS(^i)</td>
<td>1-4</td>
<td>1.71 (0.52)</td>
<td>1-3.75</td>
<td>0.10</td>
<td>1.02</td>
<td>0.92 (&lt;.001)</td>
</tr>
</tbody>
</table>

\(^a\)mSHAI: modified Short Health Anxiety Inventory.
\(^b\)Covid-F: One-Item Covid-Fear scale.
\(^c\)FCV-19S: Fear of COVID-19 Scale.
\(^d\)CAS: Coronavirus Anxiety Scale.
\(^e\)PAS: Pandemic Anxiety Scale.
\(^f\)CSS-D: COVID Stress Scales–Danger subscale.
\(^g\)CSI-16: 16-item Couple Satisfaction Index.
\(^h\)PSS: Parenting Stress Scale.
\(^i\)GFS: General Functioning Scale.

### Reliability

Internal consistencies for each of the four scales are presented in Table 3. All four scales showed at least acceptable consistency (unstandardized Cronbach \(\alpha>.70\)) [94]. Inter-item average correlations were between 0.30 and 0.63.

### Table 3. Reliability of the COVID-19–related anxiety and distress scales (N=515).

<table>
<thead>
<tr>
<th>Scale</th>
<th>Cronbach (\alpha)</th>
<th>McDonald (\omega)</th>
<th>Gutmann (\lambda6)</th>
<th>Inter-item correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCV-19S(^a)</td>
<td>.87</td>
<td>.88</td>
<td>.89</td>
<td>.52</td>
</tr>
<tr>
<td>CAS(^b)</td>
<td>.83</td>
<td>.84</td>
<td>.82</td>
<td>.51</td>
</tr>
<tr>
<td>PAS(^c)</td>
<td>.73</td>
<td>.75</td>
<td>.78</td>
<td>.30</td>
</tr>
<tr>
<td>CSS-D(^d)</td>
<td>.91</td>
<td>.91</td>
<td>.91</td>
<td>.63</td>
</tr>
</tbody>
</table>

\(^a\)FCV-19S: Fear of COVID-19 Scale.
\(^b\)CAS: Coronavirus Anxiety Scale.
\(^c\)PAS: Pandemic Anxiety Scale.
\(^d\)CSS-D: COVID Stress Scales–Danger subscale.

### Correlations With COVID-19 Anxiety Scales

Prior to analyzing their validity, the correlations of the scales with the demographic characteristics of the participants and the COVID-19 scales were examined. The parents’ age, years in a relationship, age of the youngest child, and number of children were not significantly correlated with FCV-19S, CAS, PAS, or CSS-D (all \(P>.05\)).

To investigate the convergent validity, we examined bivariate correlations between the four COVID-19 anxiety scales (Table 4). Moderate to high correlations of the four COVID-19 anxiety scales were found, ranging between \(r=0.36\) and \(r=0.65\). Except for the CAS, all scales had moderate correlations with the One-Item Covid-Fear scale, indicating convergent validity. Small to medium positive correlations were found between...
health anxiety as a trait (mSHAI) and the different COVID-19 scales, ranging from \( r=0.21 \) to \( r=0.38 \).

Table 4. Pearson correlations for COVID-19 anxiety scales and other measures of anxiety and family variables. \( N=515 \) for all scales, except CSI-16, for which \( n=486 \).

<table>
<thead>
<tr>
<th>Variable</th>
<th>mSHAI(^a)</th>
<th>Covid-F(^b)</th>
<th>FCV-19S(^c)</th>
<th>CAS(^d)</th>
<th>PAS(^e)</th>
<th>CSS-D(^f)</th>
<th>CSI-16(^g)</th>
<th>PSS(^h)</th>
<th>GFS(^i)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r )</td>
<td>1</td>
<td>0.19</td>
<td>0.38</td>
<td>0.28</td>
<td>0.21</td>
<td>0.26</td>
<td>-0.13</td>
<td>0.19</td>
<td>0.15</td>
</tr>
<tr>
<td>( P ) ( \text{value} )</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Covid-F

| \( r \)  | 0.19 | 1    | 0.71    | 0.32      | 0.56      | 0.69        | -0.03       | 0.09    | 0.05    |
| \( P \) \( \text{value} \) | <.001 | —    | <.001   | <.001     | <.001     | <.001       | <.001       | .50    | .03     |

FCV-19S

| \( r \)  | 0.38 | 0.71 | 1       | 0.51      | 0.61      | 0.65        | -0.07       | 0.17    | 0.13    |
| \( P \) \( \text{value} \) | <.001 | <.001 | —       | <.001     | <.001     | <.001       | <.001       | .08    | <.001  |

CAS

| \( r \)  | 0.28 | 0.32 | 0.51    | 1         | 0.36      | 0.40        | -0.12       | 0.15    | 0.16    |
| \( P \) \( \text{value} \) | <.001 | <.001 | <.001   | —         | <.001     | <.001       | <.001       | .07    | <.001  |

PAS

| \( r \)  | 0.21 | 0.56 | 0.61    | 0.36      | 1         | 0.61        | -0.07       | 0.25    | 0.11    |
| \( P \) \( \text{value} \) | <.001 | <.001 | <.001   | <.001     | —         | <.001       | <.001       | .10    | <.001  |

CSS-D

| \( r \)  | 0.26 | 0.69 | 0.65    | 0.40      | 0.61      | 1           | -0.05       | 0.19    | 0.14    |
| \( P \) \( \text{value} \) | <.001 | <.001 | <.001   | <.001     | <.001     | —           | <.001       | .27    | <.001  |

CSI-16

| \( r \)  | -0.13 | -0.03 | -0.07    | -0.12     | -0.07     | -0.05       | 1           | -0.31   | -0.80   |
| \( P \) \( \text{value} \) | .002  | .50    | .08      | .007      | .10        | .27         | —           | <.001   | <.001   |

PSS

| \( r \)  | 0.19 | 0.09 | 0.17    | 0.15      | 0.25      | 0.19        | -0.31       | 1       | 0.35    |
| \( P \) \( \text{value} \) | <.001 | .03    | <.001   | <.001     | <.001     | <.001       | <.001       | —       | <.001   |

GFS

| \( r \)  | 0.15 | 0.05 | 0.13    | 0.16      | 0.11      | 0.14        | -0.80       | 0.35    | 1       |
| \( P \) \( \text{value} \) | .001  | .23    | .002    | <.001     | .007      | .001        | <.001       | <.001   | —       |

\(^a\)mSHAI: modified Short Health Anxiety Inventory.
\(^b\)Covid-F: One-Item Covid-Fear scale.
\(^c\)FCV-19S: Fear of COVID-19 Scale.
\(^d\)CAS: Coronavirus Anxiety Scale.
\(^e\)PAS: Pandemic Anxiety Scale.
\(^f\)CSS-D: COVID Stress Scales–Danger subscale.
\(^g\)CSI-16: 16-item Couple Satisfaction Index.
\(^h\)PSS: Parenting Stress Scale.
\(^i\)GFS: General Functioning Scale.
\(^j\)—: not applicable.

Validity Analyses With COVID-19 Anxiety Scales and Family Measures

As hypothesized, small positive correlations between the four COVID-19 scales were found with parenting stress (\( r=0.15-0.25 \)) and general family functioning (\( r=0.11-0.16 \)). No significant associations were found between the COVID-19 scales and couple satisfaction, except in the case of the CAS, which showed a small negative correlation with couple satisfaction. Among family measures, all scales correlated at least moderately.
EFA of All COVID-19 Anxiety Scales

We performed an additional EFA on all 25 items of the COVID-19 scales to examine the overall similarity of the constructs (Table 5). The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.93. The Bartlett test of sphericity was significant ($P<0.001$). Varimax (orthogonal) rotation was applied. The Kaiser criteria [95] and scree plot retained a 5-component solution with initial eigenvalues of 8.49, 2.62, 1.83, 1.48, and 1.01, which accounted for 67.10% of the total variance. The model showed acceptable fit (RMSEA=0.068, 90% CI 0.061-0.073, Tucker-Lewis index=0.909).

The first factor included eight total items from the CSS (CSS-1 and CSS-4), FCV-19S (FCV-1, FCV-2, and FCV-5), and PAS (PAS-1, PAS-2, and PAS-4) with loadings higher than 0.40, and it accounted for 39.24% of the total variance. This factor represented COVID-19–related fear of infection. The second factor accounted for 10.5% of the total variance and was formed by the six items of the CSS-D. However, two items had cross-loadings on the first factor. One item was nearly identical to another (CSS-1 and FCV1), and the other item relates to protecting one’s family from the virus. The third factor explained 7.32% of the variance and contained six items without clear content focus, including fear of dying, nervousness about news on social media, physical symptoms of anxiety, and insomnia. One item of the PAS, on fear of leaving the house, was also loaded on this factor. Furthermore, the five items of the CAS were all loaded uniquely on the fourth factor, representing physical symptoms of anxiety and explaining an additional 5.94% of the variance. Surprisingly, these items did not load sufficiently with those of the FCV-19S on physical anxiety symptoms. Finally, three items of the PAS (PAS5, PAS5, and PAS7) formed the fifth factor regarding the socioeconomic consequences of COVID-19, explaining 4.07% of the total variance.
### Table 5. Exploratory factor analysis of the CAS, CSS, FCV-19S, and PAS scales. The varimax rotation method was applied.

<table>
<thead>
<tr>
<th>Scale item</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
<th>Factor 5</th>
<th>Uniqueness</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAS(^a)_CAS1</td>
<td>_(^b)</td>
<td>—</td>
<td>—</td>
<td>.666</td>
<td>—</td>
<td>.414</td>
</tr>
<tr>
<td>CAS_CAS2</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.665</td>
<td>—</td>
<td>.449</td>
</tr>
<tr>
<td>CAS_CAS3</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.787</td>
<td>—</td>
<td>.312</td>
</tr>
<tr>
<td>CAS_CAS4</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.798</td>
<td>—</td>
<td>.324</td>
</tr>
<tr>
<td>CAS_CAS5</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.766</td>
<td>—</td>
<td>.385</td>
</tr>
<tr>
<td>CSS(^c)_CSS1</td>
<td>.610</td>
<td>.436</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.360</td>
</tr>
<tr>
<td>CSS_CSS2</td>
<td>—</td>
<td>.692</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.312</td>
</tr>
<tr>
<td>CSS_CSS3</td>
<td>—</td>
<td>.843</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.189</td>
</tr>
<tr>
<td>CSS_CSS4</td>
<td>.468</td>
<td>.694</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.244</td>
</tr>
<tr>
<td>CSS_CSS5</td>
<td>—</td>
<td>.826</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.206</td>
</tr>
<tr>
<td>CSS_CSS6</td>
<td>—</td>
<td>.711</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.339</td>
</tr>
<tr>
<td>FCV-19S(^d)_FCV1</td>
<td>.687</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.262</td>
</tr>
<tr>
<td>FCV-19S_FCV2</td>
<td>.727</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.334</td>
</tr>
<tr>
<td>FCV-19S_FCV3</td>
<td>—</td>
<td>—</td>
<td>.765</td>
<td>—</td>
<td>—</td>
<td>.347</td>
</tr>
<tr>
<td>FCV-19S_FCV4</td>
<td>—</td>
<td>—</td>
<td>.591</td>
<td>—</td>
<td>—</td>
<td>.392</td>
</tr>
<tr>
<td>FCV-19S_FCV5</td>
<td>.431</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.425</td>
</tr>
<tr>
<td>FCV-19S_FCV6</td>
<td>—</td>
<td>—</td>
<td>.813</td>
<td>—</td>
<td>—</td>
<td>.220</td>
</tr>
<tr>
<td>FCV-19S_FCV7</td>
<td>—</td>
<td>—</td>
<td>.818</td>
<td>—</td>
<td>—</td>
<td>.212</td>
</tr>
<tr>
<td>PAS(^e)_PAS1</td>
<td>.754</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.225</td>
</tr>
<tr>
<td>PAS_PAS2</td>
<td>.715</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.301</td>
</tr>
<tr>
<td>PAS_PAS3</td>
<td>—</td>
<td>—</td>
<td>.479</td>
<td>—</td>
<td>—</td>
<td>.525</td>
</tr>
<tr>
<td>PAS_PAS4</td>
<td>.648</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.522</td>
</tr>
<tr>
<td>PAS_PAS5</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.746</td>
<td>—</td>
<td>.354</td>
</tr>
<tr>
<td>PAS_PAS6</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.864</td>
<td>—</td>
<td>.222</td>
</tr>
<tr>
<td>PAS_PAS7</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.796</td>
<td>.350</td>
</tr>
</tbody>
</table>

\(^a\)CAS: Coronavirus Anxiety Scale.
\(^b\)Factor loadings below .40 are omitted from the table to improve readability.
\(^c\)CSS: COVID Stress Scales.
\(^d\)FCV-19S: Fear of COVID-19 Scale.
\(^e\)PAS: Pandemic Anxiety Scale.

**IRT Analysis**

Based on the results of the EFA, the overall set of items did not appear to have a common unidimensional latent structure. In addition, the response format options were not identical for all questionnaires. Therefore, we conducted an analysis for each scale separately. Parameter estimation for item discrimination (ie, slopes) and item difficulty (ie, thresholds) can be found in **Table 6**. Characteristic curves for the individual items are available in **Multimedia Appendix 2**.

<table>
<thead>
<tr>
<th>Item</th>
<th>Discrimination</th>
<th>Difficulty</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>α (SE)</td>
<td>β₁</td>
</tr>
<tr>
<td>CAS¹</td>
<td>3.01 (0.39)</td>
<td>1.49</td>
</tr>
<tr>
<td>CAS-2</td>
<td>2.66 (0.41)</td>
<td>1.35</td>
</tr>
<tr>
<td>CAS-3</td>
<td>3.98 (1.97)</td>
<td>1.26</td>
</tr>
<tr>
<td>CAS-4</td>
<td>4.03 (1.05)</td>
<td>1.74</td>
</tr>
<tr>
<td>CAS-5</td>
<td>3.35 (0.56)</td>
<td>1.82</td>
</tr>
<tr>
<td>CSS-Dᶜ</td>
<td>1.97 (0.16)</td>
<td>−0.34</td>
</tr>
<tr>
<td>CSS-D2</td>
<td>2.79 (0.24)</td>
<td>0.01</td>
</tr>
<tr>
<td>CSS-D3</td>
<td>3.86 (0.35)</td>
<td>0.08</td>
</tr>
<tr>
<td>CSS-D4</td>
<td>3.66 (0.31)</td>
<td>−0.62</td>
</tr>
<tr>
<td>CSS-D5</td>
<td>3.90 (0.34)</td>
<td>−0.34</td>
</tr>
<tr>
<td>CSS-D6</td>
<td>2.57 (0.22)</td>
<td>0.06</td>
</tr>
<tr>
<td>FCV-19Sᵈ</td>
<td>2.24 (0.19)</td>
<td>−0.83</td>
</tr>
<tr>
<td>FCV-19S-2</td>
<td>1.91 (0.16)</td>
<td>−1.42</td>
</tr>
<tr>
<td>FCV-19S-3</td>
<td>2.68 (0.27)</td>
<td>0.65</td>
</tr>
<tr>
<td>FCV-19S-4</td>
<td>2.48 (0.23)</td>
<td>0.29</td>
</tr>
<tr>
<td>FCV-19S-5</td>
<td>2.26 (0.19)</td>
<td>−0.31</td>
</tr>
<tr>
<td>FCV-19S-6</td>
<td>4.40 (0.57)</td>
<td>0.66</td>
</tr>
<tr>
<td>FCV-19S-7</td>
<td>4.47 (0.59)</td>
<td>0.70</td>
</tr>
<tr>
<td>PASᵉ</td>
<td>4.19 (0.76)</td>
<td>−0.72</td>
</tr>
<tr>
<td>PAS-2</td>
<td>4.10 (0.77)</td>
<td>−1.20</td>
</tr>
<tr>
<td>PAS-3</td>
<td>1.49 (0.16)</td>
<td>0.69</td>
</tr>
<tr>
<td>PAS-4</td>
<td>1.30 (0.12)</td>
<td>−1.06</td>
</tr>
</tbody>
</table>

¹CAS: Coronavirus Anxiety Scale. Difficulty parameters for responses on a 5-point Likert scale: β₁ (from “not at all” to “rare, less than a day or two”), β₂ (from “rare, less than a day or two” to “several days”), β₃ (from “several days” to “more than 7 days”), and β₄ (from “more than 7 days” to “nearly every day over the last 2 weeks”).

ᵇN/A: not applicable (β₄ could not be calculated for CAS-1, CAS-4, or PAS-3 due to the unused item response range).

ᶜCSS-D: COVID Stress Scales–Danger subscale. Difficulty parameters for responses on a 5-point Likert scale: β₁ (from “Not at all” to “slightly”), β₂ (from “slightly” to “moderately”), β₃ (from “moderately” to “very”), and β₄ (from “very” to “extremely”).

dFCV-19S: Fear of COVID-19 Scale. Difficulty parameters for responses on a 5-point Likert scale: β₁ (from “strongly disagree” to “disagree”), β₂ (from “disagree” to “neither agree nor disagree”), β₃ (from “neither agree nor disagree” to “agree”), and β₄ (from “agree” to “strongly agree”).

ᵉPAS: Pandemic Anxiety Scale. Difficulty parameters for responses on a 5-point Likert scale: β₁ (from “strongly disagree” to “disagree”), β₂ (from “disagree” to “neither agree nor disagree”), β₃ (from “neither agree nor disagree” to “agree”), and β₄ (from “agree” to “strongly agree”).

Item discrimination was high for all items except for the PAS-4 item on worry about transferring the infection to someone else, which had a moderate level of discrimination. High alpha values in all scales indicate that the items were able to discriminate parents with a high latent trait from those with a low latent trait. With respect to item difficulty, only the CAS provided exclusively positive threshold parameters, suggesting that these items perform best when measuring people with higher levels of the latent trait.

The test information function of each scale is presented in Figure 1. All scales have the tendency to provide more information between 0 and +2 SDs than between 0 and −2 SDs. The CAS provides insufficient information for parents with scores lower than the mean. The PAS and CSS-D achieved good accuracy.
between the mean and ±1 SD. High values of the latent trait with +3 SDs were measured accurately with CAS and FCV-19S, but less precise with PAS or CSS-D. Detailed item information function values at different theta levels can be found in Table 7.

Figure 1. Total information functions for the PAS, FCV-19S, CAS and CSS-D scales. CAS: Coronavirus Anxiety Scale; CSS-D: COVID Stress Scales–Danger subscale; FCV-19S: Fear of COVID-19 Scale; PAS: Pandemic Anxiety Scale.

Table 7. Item information function values for each scale at \( \theta \) values between –2.4 and 2.4.

<table>
<thead>
<tr>
<th>Scale</th>
<th>( \theta ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>–2.4</td>
</tr>
<tr>
<td>PAS(^a)</td>
<td>1.37</td>
</tr>
<tr>
<td>FCV-19S(^b)</td>
<td>1.62</td>
</tr>
<tr>
<td>CAS(^c)</td>
<td>1.00</td>
</tr>
<tr>
<td>CSS-D(^d)</td>
<td>1.11</td>
</tr>
</tbody>
</table>

\(^a\)PAS: Pandemic Anxiety Scale.  
\(^b\)FCV-19S: Fear of COVID-19 Scale.  
\(^c\)CAS: Coronavirus Anxiety Scale.  
\(^d\)CSS-D: COVID Stress Scales–Danger subscale.

Additional Analysis of Selected Items of the PAS Subscale

We selected items for further investigation based on the analysis of distributions, variance in the sample, exploratory factor analysis, and IRT analysis. This resulted in three items on infection worries regarding oneself (PAS-1) as well as family and friends (PAS-2), and the possibility of spreading the virus to someone else (PAS-4). These were all obtained from the PAS subscale on disease anxiety. We did not consider the other items in the first factor (Table 5) due to substantial cross-loadings above .4 on other factors or gender differences in scoring \( P<.05 \).
The items of the other factors were not considered for the following reasons. The second factor (see Table 5) consists mainly of items regarding the role of the health care system and information provided on containment. Further, the third factor lacked coherence due to high variability of content with different cognitive and behavioral dimensions. Another issue was the different mean scores for mothers and fathers ($P<.01$) on some of these items (FCV-5, FCV-6, and FCV-7), which indicate gender-specific differences. Moreover, although all items of the CAS loaded on an anxiety-related fourth factor, they showed insufficient variance in the sample, with 78.1% of participants (402/515) not endorsing a single item. An additional IRT analysis revealed a lack of sufficient test information for parents with scores equal to or lower than the mean. Finally, the items of the fifth factor (PAS5, PAS6, and PAS7) referred only to socioeconomic consequences of the COVID-19 pandemic without being related to anxiety.

Accordingly, we examined the psychometric properties of the 3-item PAS scale assessing disease anxiety. Scores can range between 0 and 12 (mean 5.25, SD 3.06). The scale showed acceptable internal consistency ($\alpha=.79$). There was no significant scoring difference for mothers and fathers ($t_{511}=-0.79$, $P=.42$). Parents with elevated health anxiety had higher scores ($t_{511}=-2.70$, $P=.007$). High correlations were found with the One-Item Covid-Fear scale ($r=0.69$), the FCV-19S ($r=0.79$), the PAS ($r=0.66$), and the CSS-D ($r=0.70$). Small to moderate correlation was found with the CAS ($r=0.28$), trait health anxiety ($r=0.18$) and parenting stress ($r=0.15$; all $P<.001$). Nonsignificant correlations were found with age ($r=0.04$, $P=.27$), length of partnership ($r=0.00$, $P=.88$), age of the youngest child ($r_{hs}=0.01$, $P=.79$), couple satisfaction ($r=0.00$, $P=.98$) and family functioning ($r=0.05$, $P=.20$).

**Discussion**

**Principal Findings**

Our aim was to evaluate various existing scales for COVID-19–related anxiety and fear (ie, basal anxiety regarding the COVID-19 pandemic and the infection itself, out of pre-existing scales). In our sample, all four scales (the FCV-19S, PAS, CAS, and CSS-D) had adequate psychometric properties. However, exploratory factor analysis revealed that different facets of anxiety and worries were measured across the scales. Based on our classical test theory and IRT analysis, the PAS subscale on disease anxiety for assessing COVID-19–related anxiety seems to be appropriate as a brief scale. However, factor analysis suggests using only the items PAS-1 (ie, self-infection), PAS-2 (ie, infection of family and friends), and PAS-4 (ie, spreading of infection) for unidimensional assessment. We were able to show that these three items are psychometrically sound for covering general infection anxiety related to COVID-19 in parents. Nonetheless, all the investigated inventories had strengths, and the selection of which scale to use may be dependent on the sample in which it will be used (eg, clinical vs nonclinical, parent vs nonparent, or families with toddlers vs families with older children).

Although the CAS has a one-dimensional structure without cross-loadings on other factors, floor effects were found for three-quarters of the participants (ie, zero variance). This inventory assesses distressing bodily symptoms and may not capture general COVID-19–related stress among community samples; however, it may be suitable for clinical samples. In the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders*, the constructs of somatic symptom disorder (F45.1) and illness anxiety disorder (F45.21) replaced hypochondria [96,97]. We suspect that the CAS best detects whether parents report a somatic expression of COVID-19–related anxiety, but not necessarily whether their fear is predominantly cognitive. Interestingly, the FCV-19S contains items on somatic symptoms of anxiety, such as clammy hands and tachycardia, which did not load on the same factor as the CAS items.

The initially proposed two-factor structure of the PAS was partially replicated with disease anxiety and consequence anxiety as two latent factors. The item assessing “worries about leaving the house” no longer loaded on either factor and can be explained by timing of the data collection in the original study. McElroy et al [81] collected data early during the pandemic outbreak in April 2020, when lockdowns were in effect. This suggests that the influence of COVID-19–related anxiety items may change when perception of risk situations changes over time in society. It may be important for longitudinal studies on understanding COVID-19–related anxiety to include and test items that are relevant regardless of changes in lockdowns and public health measures.

Further, we observed lower means for all items and scales than in other studies on these measures [48,49,81]. It should be noted that the overall level of fear is probably strongly dependent on the time of the survey, the country of assessment, local closeness to infection clusters, and media reporting. At the time of our survey period, in July 2020, the number of infections in German-speaking countries was relatively stable, with greater infection clusters in a subset of settings [98]. In contrast, the validation studies [48,49,51,81] all took place between March and April 2020, at the onset of the pandemic outbreak, when there was a high level of uncertainty regarding the course of the pandemic.

In addition, small to moderate bivariate correlations between health anxiety as a trait (measured with the mSHAI) and the instruments raised questions about COVID-19–related fear and its association with health anxiety. The One-Item Covid-Fear scale had a Pearson $r$ value of 0.19 with mSHAI. The correlations of mSHAI with the CAS, PAS, and CSS-D scales were in the range of $r=0.20$-$0.28$. This suggests that pandemic-related health anxiety is distinct from trait health anxiety and should be assessed separately.

There may be several explanations for the small associations between COVID-19–related anxiety and health anxiety. Previous studies found different antecedents for COVID-19–related fears: fear about economic consequences, fear of new measures, fear of health care collapse, fear of illness, fear of death, or fear of spreading the virus to risk groups [99-102]. We assume that these fears can appear independently from each other. Not all of these fears are health-related; therefore, they are not necessarily linked to an individual’s own health anxiety [103].
In addition, the construct of trait health anxiety is based on relatively stable negative health-related cognitions and preoccupation with one's own health [96,104]. COVID-19–related anxiety may affect cognitions differently due to the public attention paid to the virus in the media, which places the focus on a public health level rather than on an individual health level. Similarly, COVID-19–related anxiety may be perceived more as a threat to an individual's family than to their own health.

In our factor analysis, all items of the CSS-D loaded on a common unique factor that had cross loadings with the general COVID-19–related anxiety factor (see Table 5). We suspect that this perceived fundamental threat occurs regardless of health anxieties and is represented in this factor. For example, an early study on the H1N1 influenza pandemic from Jones and Salathé [29] found strong clustering of anxiety related to H1N1 influenza with anxiety over trauma. The operationalization of COVID-19–related anxiety as related to threat and traumatic event perception rather than health anxiety has implications for prevention and treatment (see [12,105]). The use of a traumatic stress framework was already noted during the previous H1N1 pandemic for families [31]; however, COVID-19–specific trauma research is needed [12].

A secondary goal of the study was to investigate the association between COVID-19–related anxiety and family variables (ie, couple satisfaction, family functioning, and parenting stress). We did not find significant associations between couple satisfaction and COVID-19 measures. Although there is some evidence of a link between couple distress and anxiety [37], findings related to general anxiety symptoms or disorders may not generalize to COVID-19–related anxiety, which is related to a population-level public health crisis. In addition, we suspect that the relationship between COVID-19–related anxiety and couple satisfaction may be moderated by other variables that were not assessed, such as social support or work stress [106].

As hypothesized, parenting stress and family functioning showed small correlations with COVID-19–related anxiety among parents. Intriguingly, neither the One-Item Covid-Fear scale (Covid-F) nor the PAS subscale on disease anxiety correlated with family functioning, although all other tested COVID-19 scales did. Certain families experienced chronic stress and anxiety from the pandemic, which one would expect to impact family well-being and functioning over time and should therefore be associated [107,108]. It is possible that the high education levels in our sample may have weakened the relationship between family functioning and COVID-19–related anxiety that might be seen among samples with a wider spectrum of socioeconomic status [109-111]. In contrast, the PSS correlated consistently with all COVID-19 scales. Parenting stress may be a better indicator of COVID-19–related impacts than a general family functioning measure, which may have a more distal relationship [107]. Especially, some items of the parental stress scale were highly relevant during the time of recruitment, with limited possibility of childcare offers (eg, “Having child(ren) leaves little time and flexibility in my life”), and may have captured the COVID-19–related burden. It is possible that prolonged exposure to increased parental stress would have an effect on worsening family functioning over time. Prospective designs are needed to best understand the impacts of COVID-19–related anxiety on the parental relationship, parenting stress, and family functioning.

Finally, more than one in four parents showed significant distress in their partner relationship, and almost every fifth family had poor family functioning. Parental stress was equally substantial across mothers and fathers. High numbers of burdened parents during the COVID-19 pandemic have been reported in other studies, along with serious warnings regarding increased violence potential in families [14,22,112]. We encourage policy makers to focus on families as an important societal functional unit. Initial support for burdened parents is urgently needed at all levels to mitigate the negative impact of COVID-19 on mental health in parents and children by providing public health education [113], offering positive parenting training and psychological support via telehealth [114,115], providing funding to mitigate economic hardship [116], strengthening couple relationships, and promoting general family functioning for building resilience [117].

**Limitations**

The study is cross-sectional; thus, we cannot make any statements about causalities. All measures were translated into German and tested in a German-speaking sample. It is conceivable that there are language or country-specific differences. We excluded people with self-reported acute medical conditions, which means that the results are only generalizable to a sample of parents without medical conditions. It is possible that the relationship between health anxiety and COVID-19–related anxiety would be different in a sample of parents with acute or chronic medical conditions. Another possible limitation is that all parents were recruited on the web. Therefore, our results could be biased through self-selection [118] and overrepresentation of parents using social media.

Further, more mothers participated than fathers; therefore, further validation work with fathers is needed.

**Conclusion**

This study highlights how some of the existing scales on COVID-19–related anxiety measure different facets of pandemic-related anxiety among parents of young children. The differences across highlighted measures can serve as a guide for future selection of brief measures that assesses COVID-19–related anxiety among parents, which may be useful for future research. This study also highlights the associations between family variables and COVID-19–related anxiety, particularly in the case of parental stress. Future research should examine how anxiety can impact family relationships over time to better understand the potential impact of the pandemic on both mental health and family health. The results should also be replicated in other countries and cultures to best understand additional contextual factors.
Acknowledgments

We thank Amina Halilović, Janina Müller, and Katharina Prandstetter for their support in translating the scales into German.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Translations of the questionnaires. [PDF File (Adobe PDF File), 33 KB - mental_v7i12e24507_app1.pdf]

Multimedia Appendix 2
Item response theory analysis (item characteristic curve, total information curve, and test characteristic curve). [PDF File (Adobe PDF File), 793 KB - mental_v7i12e24507_app2.pdf]

References


59. IRTPRO. Vector Psychometric Group. URL: https://vpgcentral.com/software/irtpro/ [accessed 2020-10-08]


69. IRTPRO. Vector Psychometric Group. URL: https://vpgcentral.com/software/irtpro/ [accessed 2020-10-08]


Abbreviations

CAS: Coronavirus Anxiety Scale
CFA: confirmatory factor analysis
Covid-F: One-Item Covid-Fear scale
CSI: Couple Satisfaction Index
CSS: COVID Stress Scales
CSS-D: COVID Stress Scales—Danger subscale
EFA: exploratory factor analysis
FCV-19S: Fear of COVID-19 Scale
GFS: General Functioning Scale
IRT: item response theory
mSHAI: Modified Short Health Anxiety Inventory
PAS: Pandemic Anxiety Scale
PCA: Principal Component Analysis
PSS: Parenting Stress Scale
RMSEA: root mean square error of approximation

©Christian Kubb, Heather M Foran. Originally published in JMIR Mental Health (http://mental.jmir.org), 03.12.2020. 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 Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/; as well as this copyright and license information must be included.
Viewpoint

Videoconferencing-Based Telemental Health: Important Questions for the COVID-19 Era From Clinical and Patient-Centered Perspectives

Emil Chiauzzi¹*, PhD; Ashley Clayton²*, MA; Jina Huh-Yoo³*, PhD

¹Tridiuum Inc, Philadelphia, PA, United States
²Yale School of Medicine, Yale University, New Haven, CT, United States
³Department of Information Science, College of Computing and Informatics, Drexel University, Philadelphia, PA, United States

*all authors contributed equally

Corresponding Author:
Jina Huh-Yoo, PhD
Department of Information Science
College of Computing and Informatics
Drexel University
3675 Market St
Philadelphia, PA, 19104
United States
Phone: 1 7346453664
Email: jh3767@drexel.edu

Abstract

The COVID-19 pandemic has intensified the search for digital approaches in mental health treatment, particularly due to patients and clinicians practicing social distancing. This has resulted in the dramatic growth of videoconferencing-based telemental health (V-TMH) services. It is critical for behavioral health providers and those in the mental health field to understand the implications of V-TMH expansion on the stakeholders who use such services, such as patients and clinicians, to provide the service that addresses both patient and clinical needs. Several key questions arise as a result, such as the following: (1) in what ways does V-TMH affect the practice of psychotherapy (ie, clinical needs), (2) to what extent are ethical and patient-centered concerns warranted in terms of V-TMH services (ie, patient needs), and (3) how do factors related to user experience affect treatment dynamics for both the patient and therapist (ie, patient and clinical needs)? We discuss how behavioral health providers can consider the future delivery of mental health care services based on these questions, which pose strong implications for technological innovation, the adaptation of treatments to new technologies, and training professionals in the delivery of V-TMH services and other digital health interventions.

(JMIR Ment Health 2020;7(12):e24021) doi:10.2196/24021

KEYWORDS

telehealth; telemental health; COVID-19; videoconferencing; ethics; privacy; mental health; psychotherapy; patient-centered; lived experience

Introduction

Videoconferencing-based telemental health (V-TMH) has been the subject of clinical and research discussions for many years. Since the advent of the COVID-19 pandemic however, it has entered public health discussions in a significant way [1,2]. Essentially, V-TMH services operate on platforms used for business meetings (eg, Doxy.me, Vidyo, VSee, and Thera-link), allowing users to connect to each other via their desktop computers or mobile devices. A number of these platforms are integrated into electronic health records (EHRs), which allow for the use of existing workflows, record-keeping systems, and portals [3]. To facilitate treatment, these platforms often include scheduling and check-in systems and so-called “waiting rooms” so that clients can begin their sessions when the clinician is available, as well as high-definition audio and video, adherence to Health Insurance Portability and Accountability Act (HIPAA) data privacy requirements, branding, and analytics. Some platforms do not require the user to download software or plugins, but simply use a secure browser-based link. For HIPAA-compliant platforms, audio and video communications
are encrypted, no recordings are stored, and no personal health information is collected unless an EHR interface is used.

Despite clinician concerns about adverse effects on the therapeutic alliance, studies have indicated high ratings of satisfaction and therapeutic alliance from both patients and clinicians [4,5]. Several systematic reviews have supported the efficacy of V-TMH and telehealth in general, indicating comparable effectiveness with face-to-face treatment for a variety of conditions (eg, anxiety, depression, posttraumatic stress disorder, and substance abuse) and age groups (eg, adult, child, and geriatric) [6,7]. However, although there is empirical support for delivering cognitive behavioral therapy through telehealth services, there is still limited information about which therapy modalities are most amenable to telemental health care delivery [7]. Additional data are needed to assess clinicians’ and patients’ satisfaction with more relationally focused interventions, such as psychodynamic and interpersonal therapy [8].

Until recently, the use of telemental health in psychiatry was progressing at a relatively slow pace. Practicing psychologists had only engaged in the sporadic use of telecommunication technologies, such as videoconferencing (25%), instant messaging (3%), or smartphone apps (7%), and had spent less than 10% of their time delivering web-based psychotherapy [9]. Furthermore, only 5% of psychiatrists treating Medicare patients have conducted at least 1 telemedicine visit [10]. Many care providers lack technical knowledge, are not incentivized financially to provide telemental health services, or are concerned about safety and security issues with such services [9,11,12].

Before the COVID-19 pandemic era, care providers and institutions generally implemented V-TMH by choice. However, physical distancing regulations require most nonacute psychiatric care to be delivered remotely, resulting in the dramatic growth of telemental health services. For example, the Blue Cross Blue Shield of Massachusetts Foundation reported that the number of daily telehealth claims increased from 200 in February 2020 to 38,000 in May 2020, and about half of all telehealth claims were for behavioral health services [13].

These trends have increased the urgency to examine the effects of telemental health expansion on those who directly use telemental health services—patients and clinicians. The focus in current literature has been on demonstrating the equivalence of outcomes between in-person and telemental health care. Although these outcomes may be similar, it has become increasingly clear that the experience is different for all concerned. Reports published since the beginning of the COVID-19 pandemic have suggested that V-TMH alters the treatment dynamics in psychotherapy, introduces new ethical concerns, and shapes new interactions between clinicians, patients, and technology. Accordingly, we explored 3 perspectives that are particularly involved with V-TMH—clinical, ethical, and human-computer interaction perspectives. This viewpoint article addresses the following 3 questions: (1) in what ways does telemental health affect the practice of psychotherapy, (2) to what extent are ethical and patient-centered concerns warranted in terms of video-based therapy services, and (3) how do human-computer interaction factors affect treatment dynamics for both the patient and therapist? This paper should not be read as a literature review, but as an opinion piece that addresses the important treatment implications of the expanded use of V-TMH resulting from the COVID-19 pandemic. Our objective is to acknowledge and address the experiences of both clinicians and patients in order to optimize this very promising technology.

### V-TMH and the Practice of Psychotherapy

#### Advantages of V-TMH

Previous studies on V-TMH have noted the convenience of this technology. As V-TMH has improved, people in rural and underserved areas have been able to access care more readily [14]. V-TMH helps overcome mobility issues due to age or disabilities, scheduling problems, and limitations in transportation [15]. Furthermore, people who are concerned about the stigma of being seen at a mental health clinic or those who want greater privacy than what face-to-face services can offer may prefer V-TMH [15]. V-TMH may also be particularly advantageous to women due to their child care, spousal care, and older adult caregiving responsibilities [16]. The effectiveness of telemental health also crosses generational lines, as it has been shown to improve behavioral health access for children and adolescents [17] and has been well-accepted in geriatric inpatient and nursing home consultations [18]. Additionally, recent accounts have indicated that V-TMH enables clinicians to maintain clinical volumes despite social distancing mandates, allows patients to attend sessions when they cannot leave home due to health reasons or scheduling conflicts, and reduces no-show rates [1]. Based on these studies, it is clear that V-TMH has strong potential in retaining those who already receive mental health services and attracting new patients who might be unable to access face-to-face services.

There is also evidence that indicates V-TMH has a number of benefits beyond convenience and easy access. V-TMH is appropriate for a broad range of mental health patients, but moderators for determining which patients and patient settings may benefit the most from V-TMH services have yet to be determined [19]. At this point, it appears that that the benefits of V-TMH outweigh its deficiencies, but the former can become the latter unless clinical adaptations are considered. These adaptations apply to clinical conditions that may be more amenable to V-TMH, strategies for managing serious mental illnesses or behavioral issues, and considerations in conducting formal assessments and psychological testing.

#### Diagnostic Considerations

Telemental clinical outcomes are comparable to conventional treatment outcomes across diverse patient diagnostic groups [20]. V-TMH may be particularly advantageous for patients with conditions that detract from patient involvement in office-based treatment. These conditions include severe anxiety and avoidant behavior, severe trauma-related disorders, and agoraphobia [15]. Telemental care may be preferable for patients with severe anxiety disorders [2,19]. Telemental care may also be preferable for patients whose behaviors pose a risk to others, such as those with violent/homicidal tendencies or a history of...
Addressing Serious Mental Illness

Although people with serious mental illnesses have a lower rate of smartphone ownership than those without serious mental illnesses, the ability to use apps and social media is comparable between the 2 groups [21]. Despite a variety of clinical challenges, there is a consensus that V-TMH may be feasible, usable, and acceptable for people with serious mental illnesses [22-24]. Promising outcomes for self-management, medication adherence, psychoeducation, and symptom monitoring have been obtained from a broad set of telemental and mental health apps [21]. The use of V-TMH has also been shown to reduce self-reported psychiatric symptoms, emergency room visits, and hospital admissions [25]. However, despite these positive outcomes, a review on the use of telespsychiatry for a broad range of psychiatric disorders (eg, schizophrenia, schizoaffective disorder, and bipolar disorder) found little evidence that this modality affected hospital readmission rates [26]. The American Psychological Association (APA) Task Force on Serious Mental Illness/Severe Emotional Disturbance recommends several strategies, as follows: (1) practice sessions to familiarize patients with telemental health technology, (2) the use of brief and frequent sessions to manage periods of distress and reinforce the therapeutic alliance, and (3) the promotion of team collaboration to help patients maintain access to prescriptions and lab testing [27].

Assessment and Psychological Testing

The administration of interview-based assessments and psychological testing through V-TMH presents unique challenges. Psychological testing procedures that require the physical manipulation of materials, standardized interactions, or observations in a physical environment may require adaptations [28]. This is particularly true for cognitive or neuropsychological assessments, which may be sensitive to the lack of a physical presence and the technological comfort of the tester and patient [29]. Many of these assessments may be time sensitive or involve high stakes. Therefore, there is a need to identify potential factors that influence test reliability and validity and find ways to account for them so that testing procedures can be optimized [29].

Emerging findings under controlled conditions have suggested that equivalence between face-to-face and videoconference-based assessments can be attained [30]. This has been demonstrated with brief cognitive test batteries [30], and obtaining such equivalence may work better with neuropsychological testing involving visually-dependent tasks rather than motor-dependent tasks [31]. When providing V-TMH services, the APA suggests carefully considering how V-TMH affects the presentation of materials (eg, shadowing and blurriness), substituting subtests that do not require the physical manipulation of objects, and widening confidence intervals when interpreting results [28]. In recent years there has been a proliferation of mobile technology–based cognitive assessments that have demonstrated psychometric properties comparable to those of laboratory-based assessments [32]. These apps can support repeated measurements that increase detection sensitivity in a patient’s natural environment and can supplement more formal traditional methods.

Ethical Considerations in V-TMH

The forced use of V-TMH during, and potentially after, the COVID-19 global pandemic raises ethical concerns about patient safety and privacy.

Patient Safety Risk

One of the primary ethical considerations in V-TMH is adequately ensuring the safety of the patient. Suicide prediction is challenging even under ideal assessment conditions, and it may be more difficult through V-TMH services. Psychiatrists and the mental health community have invested in the development of tools and predictive models to help determine who is at an elevated risk of suicide, but little success has been achieved [33-35]. Relying on clinical intuition to assess suicidality is less than ideal in face-to-face encounters, but V-TMH may introduce subtle changes to interpersonal relationships, trust, and decision making. For example, distortions or delays in the video can increase anxiety and feelings of disconnection in patients experiencing crises [36].

Due to the difficulty in determining treatment allocation based on suicide assessments [34], high-stakes decisions, such as seeking the involuntary detainment of a patient, require careful consideration. Involuntary commitment may be a tragic necessity during a psychiatric emergency, and it can cause potential harm and suffering for patients [37]. One of the most difficult decisions—if not the most stressful decision—for any physician or licensed mental health provider can make is to infringe on an individual’s fundamental civil rights; however, there is precedence. In 1993, a patient was involuntarily committed to a psychiatric hospital by use of videoconferencing. The patient petitioned that his rights to due process were violated, alleging that “the quality of information available through videoconferencing was limited and increased the risk of erroneous result” [38,39]. Upon appeal, the court determined that videoconferencing was equal to a face-to-face evaluation, asserting that the patient’s facial expressions and demeanor were easily observable. Therefore, videoconferencing did not increase the risk of erroneous results [39].

We do not yet know how V-TMH impacts clinicians’ ability to predict who is acutely at risk of suicide. However, actions can be taken to assist in managing such risks. Guidelines produced by the APA and American Telemedicine Association provide best practices for managing suicide risk [2]. Many of the guidelines mirror those for in-person care, such as obtaining patient contact information; monitoring the patient’s risk factors, use of substances, and access to lethal means; and having a safety plan available [40,41]. The major difference between V-TMH and in-person care is the use of telehealth technologies, as an interruption in a connection between the clinician and patient at an inopportune time can interfere with a needed intervention. The APA guidelines stress the need for the identification and consistency of the patient’s location during sessions and the availability of multiple forms of contact.
information (eg, phone, email, and SMS texts) in case contact is interrupted [2]. Despite these guidelines, clinicians remain skeptical about whether mental health professionals can screen at-risk patients through telemental health methods [9]. At present, it is incumbent on providers to carefully consider the risks involved in client safety and liberty when providing remote services to treat individuals in acute crises.

There has been a resurgence of enthusiasm for eased access to both in-home and cross-state telehealth care [42]. We must carefully consider the significant implications that such a model would have on individuals at elevated risk of suicide, especially when there is a heightened potential for detention and involuntary hospitalization. Enlisting a second care provider who is geographically closer to the client and building relationships with local services can help clinicians provide telemental health services across state lines and increase patient safety [43].

Intimate Partner Violence

There is limited data on the safety and efficacy of telemental health treatment for individuals who actively experience intimate partner violence (IPV) [44]. A systematic review of studies on using V-TMH for people who have experienced IPV only found a small number of studies, and most of these studies focused on women [44]. Other such studies on telehealth were designed for women in domestic violence shelters or mental health clinics [45,46]. There are no data available on providing V-TMH interventions at home for this population, and the risks and unintended negative consequences of V-TMH interventions are currently unknown.

It has been reported that people who have experienced previous natural disasters (eg, Hurricanes Harvey and Katrina), shelter-in-place orders, restrictions on travel, increased basic needs insecurity, and loss of jobs are likely to experience IPV [47,48]. Individuals who experience violence at home need access to support and services. For individuals who received mental health treatment before the COVID-19 pandemic, the transition to telemental health services is exceedingly complicated. Perpetrators of IPV often closely monitor and restrict access to electronic communication, making it difficult for individuals who experience IPV to access help, especially while at home. Such individuals who can attend V-TMH appointments may be unable to obtain any privacy at home, making it extremely difficult to disclose IPV experiences and placing them in imminent danger. Similarly, the telehealth provider must be scrupulous with their words, lest they say something that would upset the perpetrator and elevate the risk of danger. Data regarding the safety of V-TMH at home are not yet available. However, crisis support and other such lines have been supporting individuals who experience IPV for decades.

It is easier to move around and find privacy while on the phone than doing so while in a videoconference, and the care provider on the other end cannot be seen during phone calls. Therefore, covers such as “I’m talking to my sister” can be used. As such, phone calls may be safer than V-TMH. At present, there is no empirical evidence that favors video-based care over audio-only care [1]. However, voice-only communications may enhance empathic accuracy and the assessment of emotions [49]. Many insurers have temporarily allowed mental health providers to bill for telephone-delivered services during the COVID-19 pandemic, and this should be viewed as a viable option. However, as temporary policy relaxations begin to roll back while the public health crisis continues, it is essential to carefully consider how different communication modalities impact access to care and support for vulnerable populations.

Protecting Privacy

Confidentiality is a core obligation and ethical standard for health care providers, especially for mental health clinicians who by the nature of their work treat vulnerable and historically marginalized — and perhaps presently marginalized — populations. Breaches of confidentiality within mental health care can decrease the efficacy of therapy [50]. The US government temporarily relaxed HIPAA regulations, which allowed for the quick scaling of telehealth care to increase access to medical care while practicing shelter-in-place orders and physical distancing guidelines during the COVID-19 pandemic. The loosening of regulations correctly prioritized access to care over the tools that enabled it [51].

Several V-TMH services, especially those provided by hospitals and large health systems, now use EHRs. Before the COVID-19 pandemic, there was some integration of psychiatric care into EHRs, yet confidentiality remains an issue. Documenting diagnoses and treatment plans within the EHR helps ensure that care providers are on the same page. However, the majority of EHRs do not have a section with restricted access for care providers to write confidential, detailed notes about an encounter, thereby violating federal statute 42 of the Code of Federal Regulations Part 2 [52]. This statute was intended to protect patient privacy and avoid adverse outcomes, but its unintended consequences have made integrated care difficult. While the HIPAA allows for the sharing of patient information for care coordination, the sharing of psychiatric notes falls into ethical and legal grey areas. Care providers should openly discuss the unique limitations of privacy when using an EHR, even if these limitations are described within written informed consent forms [53]. With the lightning-speed implementation of telemental health services, care providers must take the time to research and share the benefits and risks of V-TMH with their patients, especially those that differ in face-to-face treatments. This is critical when using non-HIPPA-compliant platforms during the pandemic. Since informed consent requirements vary state by state and not all states require additional consent, care providers need to stay attuned to changes in regulations.

User Experience With V-TMH

Clinician Perspectives

In 1 survey, more than 75% of psychologists were slightly or not at all confident that they could use telecommunication modalities (ie, video, phone, email, and chat and messaging apps) without an initial in-person assessment [9]. With the widespread uptake of telemental health during the COVID-19 pandemic, we are gaining a better understanding of everyday clinicians’ experiences with providing treatment via V-TMH instead of just their intentions or attitudes toward hypothetical...
V-TMH situations. Recent interviews with psychiatrists who deliver telemental care during the COVID-19 pandemic have found that clinical challenges, such as the inability to conduct physical examinations, difficulties in evaluating extrapyramidal symptoms from antipsychotics, and home distractions, may negatively impact the quality of provider-patient interactions [54]. These challenges may be partially due to the Ryan Haight Act regulations that previously required psychiatrists to conduct an initial in-person medical evaluation prior to prescribing controlled substances. However, as of March 2020, the Drug Enforcement Administration has suspended these regulations, thereby requiring psychiatrists to adjust to remote assessments when prescribing controlled medications. On the positive side, psychiatrists and other clinicians have reported increases in their understanding of family and home dynamics, access to underserved patients, and their ability to facilitate treatment when patients are more relaxed in their homes [1].

Beyond the clinical advantages and disadvantages of V-TMH, clinicians have also been sharing their individual experiences related to the phenomenological aspects of V-TMH. Recent newspaper and magazine articles have highlighted the newfound intimacy that telemental health has introduced to some therapist-client relationships. However, this newfound intimacy may not be welcomed by all. Like their patients, clinicians have had to optimize workspaces to balance professionalism and privacy, particularly when their own family members may be at home due to work and school closures [1]. In Self Magazine, Dr Jessica Gold recently shared her experiences as a psychiatrist now practicing teletherapy and she highlighted the “very significant things” she misses about in-person treatment, as follows: “It turns out I went into a field of talking to humans and listening to them because, quite simply, I like people. Online interactions aren’t the same” [55]. Furthermore, a recent article in The New Yorker features multiple psychologists discussing their new consciousness of patients seeing them and highlighting the difficulty of holding the therapeutic frame without the regular rituals of in-person care [56]. Additionally, therapists have reported the disconcerting and sometimes distracting experience of seeing their own faces during videoconferencing. A mental health clinician in Washington, DC experienced the loss of a safe space and spoke about how the transition from the refuge of the therapy office to receiving therapy within one’s home is difficult for both therapists and clients alike [57].

**Patient Perspectives**

Evaluations of patients’ attitudes toward telehealth have focused primarily on satisfaction data and therapeutic alliance ratings [8]. Although important, these are only 2 factors for assessing patients’ satisfaction [8]. As V-TMH takes on an increasingly influential role in mental health treatment during the COVID-19 pandemic, it is useful to consider patient accounts that reflect the less easily measured aspects of V-TMH interactions. V-TMH brings nuanced contexts that might interfere with the quality and privacy of mental health therapy sessions.

In many cases, clinicians may not be aware of the practical decisions and dilemmas that their patients face when attending these sessions. For instance, finding a private and comfortable space at home may force individuals to consider using private, intimate spaces. Due to a lack of privacy within some homes, patients have taken to using their bathrooms and closets for telehealth visits [58]. Some have also chosen to seek privacy in their cars or out on walks [59].

The use of V-TMH may also cause patients to feel unanticipated self-consciousness and discomfort during V-TMH sessions. This may potentially affect self-disclosure or cause heightened emotional reactions. Self-disclosure is defined as “the verbal revealing of personal information, thoughts, or feelings about oneself” [60], and it is one of the most salient behaviors in computer-mediated communication (CMC) [61]. Self-disclosure plays a critical role in relationship development and maintenance in CMC and face-to-face relationships [62]. Depending on how comfortable patients are with self-disclosure during V-TMH appointments, the results may be controversial, similar to how there is disagreement in literature about the consistency, size, and direction of differences of self-disclosure in CMC [61].

Self-view is a unique V-TMH feature that allows for the viewing of one’s own face during conversations. Since this is dramatically different from in-person conversations and treatment, the self-view feature may interfere with participants’ self-disclosure during V-TMH appointments. In most cases, the visual anonymity provided by CMC has fostered significantly higher self-disclosure levels than in face-to-face conversations [63]. However, the self-view feature of V-TMH tools directly counteracts the critical role that CMC plays in self-disclosure. Furthermore, V-TMH generates an environment in which gestural interactions and nonverbal cues (eg, eye contact) may be distorted and inadequately communicated due to technological factors, such as video bandwidth, camera viewpoint, and image resolution quality [64-67]. Technological features that might be considered minor, such as screen size, can affect perceptions of trust toward objects shown in the video calls [68]. Additionally, a sense of proximity can be gained through various methods, such as matching the patient’s eye gaze or showing more of the clinician’s body in video calls as opposed to just a headshot. This sense of proximity can help provide an effective V-TMH experience [69]. As such, features in video calls, such as the background shown behind the speaker and being able to see one’s own face, can be moderating factors for establishing affective and cognitive trust, a sense of proximity, good telepresence, and social and emotional connections.

People will eventually get accustomed to the affordances available in V-TMH technologies and learn to better regulate their interactions [70]. However, during the COVID-19 era, patients lack options for seeking mental health services and may feel forced to use V-TMH services over other modalities because of public health regulations. This raises several unresolved questions and many implications to consider when thinking about the post-COVID-19 era.
Implications of V-TMH in the Peri- and Post-COVID-19 Era

Current Status of V-TMH

Since the COVID-19 pandemic has forced patients and clinicians to shift quickly from face-to-face mental health therapy to V-TMH therapy, the focus of V-TMH has turned to implementation and access. This has allowed little consideration for the nuances of the V-TMH treatment experience, leaving many unresolved questions that software developers, health care administrators, organizations, patients, and clinicians need to consider. As we navigate through the current pandemic and prepare for a post-pandemic recovery, these questions are highly relevant to the future viability of V-TMH.

How Can V-TMH Services Assist Those Whose Mental Health Is Affected by COVID-19?

Past studies on severe acute respiratory syndrome survivors have found increases in the incidence of posttraumatic stress disorder, depression, chronic pain, obsessive-compulsive disorder, and chronic fatigue syndrome [71]. Negative quarantine effects increase with the perceived difficulty of compliance, length of quarantine, compliance with quarantine requirements, fear of infection, lack of supplies, and financial pressures [72]. Unemployment is also of particular concern. Past data have shown that a 5% increase in unemployment is associated with 4000 additional suicides and 775 additional deaths for each additional percentage point [73]. The Well Being Trust estimates that between 27,644 and 154,037 additional deaths have been due to suicide and substance abuse, and these additional deaths are related to the effectiveness of recovery efforts [74]. These data show the high demand for mental health services. Future models of mental health care delivery, including V-TMH, will need to address the influx of new cases by providing more flexible treatment options. Since mental health clinicians have been treating patients during the pandemic, many clinicians have gained experience in addressing both COVID-19–related and mental health–related challenges through V-TMH. Cost-effective, readily available, and socially distanced treatment will be necessary for those dealing with mental health crises, the aftereffects of quarantine, unemployment, and social adjustment issues. Since many clinicians have backgrounds in behavior change techniques, they may also have a role in supporting the behavioral strategies used to prevent the transmission of SARS-CoV-2 (eg, mask use, hand washing, etc) [75].

How Can We Better Address the Digital Divide by Helping Patients to Effectively Use Technology?

Since most Americans, including those who previously lacked internet access, are increasingly gaining access to the internet through smartphones, V-TMH has the potential to become more readily available. Telehealth has typically been viewed as a service that is useful for hard-to-reach locales (eg, rural areas) or locations with underserved populations. This has shown no signs of abating. For example, since the pandemic began, the telehealth adoption rate for primary care visits was 28% higher in urban areas than in rural areas [76]. However, telehealth adoption rates in the Medicaid population are lower than those in populations who do not have Medicaid [76]. Due to the COVID-19 pandemic, there is a broader need for accessing telehealth care among mental health patients, as evidenced by the 61.8% telehealth adoption rate among psychiatrists [77].

Solutions for the digital divide issue must be more nuanced than providing mere access to the internet, as improved access to broadband internet is critical [78]. The quality of one’s connection (ie, high-speed vs low-speed bandwidth) brings more richness and variation to the internet experience [79]. With V-TMH, it is about making software usable and relevant to a broad audience, thereby enabling active participation in web-based discussions and the interpretation of social cues from an interpersonal context that is more limited than in face-to-face sessions.

Social determinants of health, such as cultural expectations about technology, digital literacy, economic factors, and comorbidities that affect the use of technology, may perpetuate inequalities [80]. Programs, such as the Digital Opportunities for Outcomes in Recovery Services (DOORS), have been developed to address low digital literacy in patients with serious mental illnesses [81]. This interactive training program helps patients learn safe smartphone usage, use technology to build wellness habits, and learn new skills through web-based resources. Although DOORS was specifically designed to support the use of digital health apps, its guiding principles and framework could easily be adapted to support the use of V-TMH services. Older adults who live in assisted and independent living communities and have basic training on using computers and navigating through the internet are able to reduce loneliness and increase social contact [82]. These findings provide critical evidence for how technology training can resolve negative outcomes that stem from the isolation that vulnerable populations may face with the COVID-19 quarantine. Beyond addressing the digital literacy gap, it is also important to note that the 2018 American Community Survey found that 15% of households lack broadband internet access, and a 2020 Brookings Report asserted that “broadband is the country’s most inequitable infrastructure” [83]. Expanded access can be achieved through a government subsidy program, similar to the Federal Communications Commission Lifeline program, but additional solutions are needed.

What Have We Learned About the User Experience and Clinical Implications for Those Who Have Transitioned From Face-To-Face Treatment to V-TMH Treatment?

Treatment effectiveness depends on how well clinicians and patients can adapt to new ways of delivering and receiving mental health services. New technology adoption is not simply driven by efficiency or technological advancement, it is also governed by multiple contextual factors, including the meanings that people put on such factors. The implications of particular gestures portrayed by a limited view of a person’s face and chest differ from those portrayed by hands and legs [65]. Beyond the lack of V-TMH quality due to limited internet bandwidth, the more important problem lies in how the contexts delivered through V-TMH are being perceived and how people interpret
what it means to use V-TMH for the purpose of mental health therapy. What would it mean to not meet the care provider in person anymore? What does it mean to blend the therapy hour into one’s daily schedule instead of having delineated time in a professional office? For clinicians, what does it mean to not be able to physically hand over a tissue box to the patient? How do the meanings that people put on a technological tool affect treatment outcomes? These are questions that we have not dealt with on the scale we have been anticipating. Studies on the effectiveness of mental health service adaptations during the COVID-19 pandemic remains mainly anecdotal, and quality measures for assessing effectiveness have yet to be proposed [84].

Given the Popularity and Efficacy of V-TMH, What Are the Privacy Implications for Loosening Post-COVID-19 Federal and State Regulations Governing V-TMH Delivery and Reimbursement?

During a global pandemic, the risk of in-person treatment far outweighs the benefit (with some exclusions), which has led to a variety of changes in telehealth regulations [85]. The Coronavirus Preparedness and Response Supplemental Appropriations Act 2020 has allowed Medicare patients to receive telehealth services in their homes, waived HIPAA violation penalties for providers who treat patients on platforms that do not meet HIPAA standards (eg, Skype and Facetime), and provided federal waivers to a number of states to allow providers to treat their patients out of state.

These changes, as well as others, have significantly reduced barriers to telemental health service provision, but they have also increased risks to personal privacy. It may indeed be the case that the risks involved are tolerated by both the patient and clinician alike, especially when telemental health enables access to care for people who do not otherwise have the resources and luxury to regularly attend in-person treatment [8]. However, the use of digital technology generates a large amount of data that can reveal a lot of personal information, especially when data from multiple data sources are combined. All digital technology data have the potential to be considered personal health information [86]. In addition, there are unresolved issues surrounding the reidentification, marketability, and invisibility (ie, people being unaware of how their data are being used and tracked) of digital technology data [87,88]. Experts in digital research and privacy regulation have a difficult time understanding the risks of using digital technology, and there are no concrete laws to effectively regulate privacy surrounding digital technology, which is constantly evolving [86]. We expect these challenges to be even greater for practicing clinicians and their patients. In order to facilitate transparency, informed consent resources have become available, such as the Telehealth Consent Teach-back Documentation sheet developed by the Agency for Healthcare Research and Quality [89].

How Can We Prepare Clinicians for the More Effective Use of V-TMH and Development of More V-TMH-Based Treatment Models?

Since the skills used in face-to-face sessions do not necessarily translate to V-TMH sessions, recent studies on mental health professionals have indicated the need for training to deal with clinical, legal, and technical issues [9,90]. Clinicians may be particularly challenged by emergencies in a V-TMH environment [9]. Driven by the relaxing of regulations, the need for an immediate pandemic response has led to the rapid and successful deployment of V-TMH in large academic centers [1] and clinical training programs [91]. These case studies have identified important recommendations for optimizing the physical arrangement of sessions (eg, eliminating visual and auditory distractions), setting up technology (eg, reducing open programs on one’s screen), and communicating effectively with patients (eg, clarifying expectations, benefits, and backup plans in case access to V-TMH is lost).

Expanded V-TMH deployment may also require the consideration of new specialties or service offerings, such as those provided by professional clinical technologists (ie, individuals familiar with a wide variety of digital health resources) [92]. Their main functions would be to match and train patients on health technologies that address their clinical needs, as well as train, educate, and support care providers. The development of web-based clinics is another potential vehicle for delivering digital care, as treatment would be constructed around the use of technology with patients while increasing the access to and offerings of clinical services [93].

Conclusion

V-TMH, as well as telemental health in general, has a great amount of potential for providing opportunities to increase access to mental health care. However, with the COVID-19 pandemic, the use of V-TMH has been forcefully expanded to those who might benefit more from in-person mental health care. Understanding V-TMH users’ perspectives can help shape future services to be rendered and delivered in a safe and effective manner. This paper provides expert opinions from those with expertise in clinical psychology, service design, and research. Given the uncertain future of the COVID-19 pandemic, addressing the important questions presented in this paper will be critical for maximizing the value of V-TMH for patients and providers. V-TMH may have been disseminated through forced use, but it is unlikely to disappear once the current public health crisis has ended. The issues we have described will have strong implications for technology innovation, the adaptation of treatments to new technologies, and training professionals in delivering V-TMH services and other digital health interventions. Regulations and reimbursement policies need to encourage the broader use of V-TMH, which has the potential to expand access to treatment. The infrastructure for using V-TMH services needs to be better developed; people need reliable internet access and technology with the capabilities for V-TMH. This is an issue of social justice and should be a primary concern. High-speed internet, innovative delivery tools, and training programs for professionals who use such tools can equitably address the mental health needs of the currently served, underserved, and unserved.
Authors' Contributions
EC, AC, and JH were all involved in manuscript conceptualization, literature review, and writing.

Conflicts of Interest
EC is an employee of Tridiuum, Inc. AC and JH have no conflicts to declare.

References


43. Chiauzzi et alJMIR MENTAL HEALTH


Abbreviations

APA: American Psychological Association
CMC: computer-mediated communication
DOORS: Digital Opportunities for Outcomes in Recovery Services
EHR: electronic health record
HIPPA: Health Insurance Portability and Accountability Act
IPV: intimate partner violence
V-TMH: videoconferencing-based telemental health

©Emil Chiauzzi, Ashley Clayton, Jina Huh-Yoo. Originally published in JMIR Mental Health (http://mental.jmir.org), 08.12.2020. 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 Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.
Flip the Clinic: A Digital Health Approach to Youth Mental Health Service Delivery During the COVID-19 Pandemic and Beyond

Tracey A Davenport¹, BA(Hons), eMBA; Vanessa Wan Sze Cheng¹, PhD; Frank Iorfino¹, PhD; Blake Hamilton², MS, DClinPsy; Eva Castaldi², BSc, PostGradDip; Amy Burton², BSc(Hons), PhD; Elizabeth M Scott¹, MBBS, FRANZCP; Ian B Hickie¹, AM, MD, FRANZCP, FASSA

1Brain and Mind Centre, The University of Sydney, Sydney, Australia
2headspace Camperdown, Sydney, Australia

Corresponding Author:
Tracey A Davenport, BA(Hons), eMBA
Brain and Mind Centre
The University of Sydney
94 Mallett Street
Sydney, 2050
Australia
Phone: 61 404 83 9897
Email: tracey.davenport@sydney.edu.au

Abstract

The demand for mental health services is projected to rapidly increase as a direct and indirect result of the COVID-19 pandemic. Given that young people are disproportionately disadvantaged by mental illness and will face further challenges related to the COVID-19 pandemic, it is crucial to deliver appropriate mental health care to young people as early as possible. Integrating digital health solutions into mental health service delivery pathways has the potential to greatly increase efficiencies, enabling the provision of “right care, first time.” We propose an innovative digital health solution for demand management intended for use by primary youth mental health services, comprised of (1) a youth mental health model of care (ie, the Brain and Mind Centre Youth Model) and (2) a health information technology specifically designed to deliver this model of care (eg, the InnoWell Platform). We also propose an operational protocol of how this solution could be applied to primary youth mental health service delivery processes. By “flipping” the conventional service delivery models of majority in-clinic and minority web-delivered care to a model where web-delivered care is the default, this digital health solution offers a scalable way of delivering quality youth mental health care both in response to public health crises (such as the COVID-19 pandemic) and on an ongoing basis in the future.

(JMIR Ment Health 2020;7(12):e24578) doi:10.2196/24578

KEYWORDS

health information technologies; clinical staging; youth; mental health; transdiagnostic; eHealth; routine outcome monitoring; adolescent; mental health services; health services; telemedicine; monitoring; outcome; young adult; COVID-19

Introduction

The COVID-19 pandemic has created a major public health crisis that has led to multidisciplinary national coordination in which physical distancing and social isolation are being mandated to flatten predicted morbidity and mortality curves [1]. This has led to the emergence of a second public health crisis due to the onset or exacerbation of poor mental health symptoms caused by the social and economic impacts of the COVID-19 pandemic, leading to increased mental health service demand [2].

In response to the pandemic, the Australian health system has embraced existing telehealth solutions in which consumers receive care through telephone calls, video calls, or the internet. In March 2020, Medicare moved to support telehealth sessions through temporary Medicare Benefits Schedule (MBS) item numbers that include receiving care via information and communication technologies [3]. Importantly, this has provided a mechanism for many Australian mental health services to continue to operate “business as usual” even though they have been required to close their front doors.

The COVID-19 pandemic has necessitated the establishment and improvement of the technological infrastructure required
to execute these temporary, stopgap solutions. Now that the groundwork has been laid, it is time to capitalize on these infrastructure developments and adapt our public mental health system to allow true digital health that fully integrates health information technologies (HITs) into service delivery models of care.

**The Need for Digital Health**

Pre–COVID-19, the World Economic Forum highlighted massive problems in mental health service provision (eg, stigma, consumer-reported poor care experiences, late intervention, mental health treatment isolated from other physical and social needs, poor resource allocation, and service fragmentation) and called for the “rapid deployment of smarter, digitally enhanced health services” as a means to address issues of demand. More recently, the World Health Organization has highlighted the need to urgently increase mental health service capacity in response to the COVID-19 pandemic. Physical distancing, social isolation, fear of contagion, and the sudden loss of family members are being further impacted by distress caused by loss of income and employment, with increased rates of stress, anxiety, depression, anger, and fear already being reported worldwide. Within Australia, the Federal Treasury forecasts an unemployment peak of 10%-11%, which translates to an approximate jobless rate of 24% for Australian young people. These trends are particularly concerning, as major mental disorders and substance use disorders affect at least one in four young people by 25 years of age and are associated with significant disability and premature death in this group worldwide. The increase in youth unemployment as a result of the COVID-19 pandemic will further compound the vulnerability of young people as a group.

Already, many Australian mental health and suicide prevention services have reported an increase in demand during the COVID-19 pandemic, highlighting the need to increase service capacity to mount an effective response. Further, a recent dynamic simulation modelling study estimated that for young people, the impact of this secondary public health crisis will result in a 21% increase in mental health–related emergency department (ED) presentations, a 22% increase in self-harm hospitalizations, and a 23% increase in deaths by suicide. Importantly, this model also found that increasing the provision of general practitioners, psychiatrists, and allied services by 11% per year and Community Mental Health Centre service capacity by 10% (resulting in a total youth mental health service capacity increase of 40%) alone would not have a major impact. However, if this capacity increase were combined with digital health (or technology-enabled care) and post–suicide attempt aftercare, this combination of strategies would result in an 8%-10% reduction in ED presentations, self-harm hospitalizations, and deaths by suicide.

A key priority should therefore be to immediately increase the speed and quality of the responses of the mental health system to increases in service demand. Previous research has established the utility of “blended” approaches to mental health care that combine face-to-face treatment with web-based tools such as internet-delivered mental health interventions and self-monitoring apps. Redesigning youth mental health service delivery pathways to integrate digital health solutions would build on the progress that has been made in the past (as well as our initial response to the COVID-19 pandemic) and establish a mental health workforce that is equipped to respond to surges in demand, both now and on an ongoing basis in the future.

**Digital Health as a Demand Management Strategy**

In academic literature, digital health has been defined as the “cultural transformation of how disruptive technologies that provide digital and objective data accessible to both caregivers and patients lead to an equal level physician-patient relationship with shared decision-making and the democratization of care”. In clinical practice, our research team has more recently defined digital health as the “provision of guided mental health care where consumers navigate a rapid and more effective system experience of service entry, skilled assessment, and multidisciplinary and coordinated care, as well as ongoing outcome-based monitoring.” Importantly, the latter definition situates consumers more strongly within the service operational processes through which mental health care is delivered, and it places additional focus on system-level factors such as demand management.

Over the past decade, the mental health sector has experienced a substantial surge in the development of numerous apps and technologies for mental health. On an individual level, web-based cognitive behavior therapy modules and web-based programs have been found to lower anxiety and depression symptoms in young people; also, many studies have found that young people are receptive to accessing HITs for their mental health in conjunction with face-to-face care. Providing these HITs as an adjunct resource to support young people in early stages of illness could promote agency instead of reliance on an overburdened mental health care system, and it would be particularly helpful for young people who experience barriers to regularly accessing in-clinic care, such as geographic remoteness.

The Australian mental health system currently funds a limited number of psychologist appointments per year for each individual consumer, and senior clinicians in conventional mental health services spend extensive time conducting thorough face-to-face assessments. The successful integration of HITs to support this process could represent significant time-saving and increase efficiencies in assessment, care coordination, and monitoring; thus, the time of senior clinicians could be reallocated to delivering skilled interventions. On the service level, HITs have already demonstrated potential for enabling an appropriate and timely response for young people reporting higher levels of suicidality, facilitating broader assessment of the totality of a young person’s needs, and enabling senior clinicians to move away from traditional evaluations toward detailed data-driven assessments.

It is now time to take the next step of integrating HITs into mental health service delivery models of care in a way that...
meets increasing consumer demand and expectations, maximizes use of senior clinician skill, enables truly person-centered and collaborative care, and preserves service quality [26,27]. Although it remains important to develop and assess the efficacy of web-based interventions, an equally important task is to determine how to integrate and use these new technologies effectively [28]. Digitizing mental health service delivery, or “flipping” conventional models of service delivery so that a larger proportion of care is delivered via HIT-enabled solutions, would therefore address many barriers to help-seeking and service access, and take advantage of many of the benefits listed above.

The effectiveness of this approach would vary across mental health services; the optimal “flip” of in-clinic versus web-based service delivery is likely to be different for each service. We theorize that optimal level of “flip” will depend on the following factors: intensity of care provided by a mental health service (ie, primary and specialist services); digital familiarity of service staff and consumers; continuity of the service’s workforce; the aggregate physician-patient relationship; level of communication for change management; and local technology infrastructure and internet connectivity. It is likely that each service will need to iteratively identify its optimum level of “flip” over time. Regardless, the basic principles of using HITs to enhance key service operational processes and improve the quality of mental health care would apply.

**Presenting an Innovative Digital Health Solution**

Our research team has recently developed an innovative digital health solution that comprises two components: (1) a measurement-based (data-driven) model of highly personalized youth mental health care, the Brain and Mind Centre (BMC) Youth Model; and (2) a dedicated HIT designed to deliver this model of care (exemplified by, but not limited to, the InnoWell Platform).

The BMC Youth Model [29] recognizes that while it is crucial to intervene at early stages of mental illness (ie, early intervention), these early stages also tend to be characterized by nonspecific symptoms that overlap disease categories and do not meet diagnostic criteria. Under the BMC Youth Model, a young person’s needs are assessed on a multidimensional outcomes framework, and their underlying pathophysiological mechanisms and illness trajectories are considered. They are also allocated to a clinical stage that represents their relative state of illness severity and persistence, with Stage 1a representing nonspecific symptoms, Stage 1b representing attenuated syndromes, and Stage 2+ indicating more discrete and persistent disorders [30]. The BMC Youth Model integrates over 10 years of neurobiological, neuropsychological, and clinical research into a single framework of youth mental health care, and it was designed to be supported by two mechanisms: education and training for service staff (Scott et al, in submission) and a dedicated HIT whose key functionalities would be specially developed to support the core concepts of the BMC Youth Model. We developed the InnoWell Platform in our work in Australia for this purpose.

**Details of the InnoWell Platform**

The InnoWell Platform is an industrial-grade HIT manufactured by InnoWell Pty Ltd (a joint venture between The University of Sydney and PricewaterhouseCoopers [PwC] Australia). This platform is listed in the Australian Register of Therapeutic Goods (software as a medical device, class 1, ARTG ID 315030) as a “customisable digital toolkit to assist assessment, monitoring, and management of mental ill health and maintenance of wellbeing. It does this by collecting, storing, scoring, and reporting personal and health information back to consumers and their health professionals to promote collaborative care partnerships” [31]. Since its inception (wireframe and prototype stage), the InnoWell Platform has been continually co-designed with target end users, including consumers and their supportive others, health professionals, and other service staff [24].

Importantly, the BMC Youth Model is not wedded to one HIT; it can be supported by any HIT that has been designed to provide highly personalized and measurement-based care according to its key clinical and scientific principles. The development of this dedicated HIT would ideally involve a co-design process with target end users (consumers and their supportive others, clinicians, and other service staff). Co-design is the deliberate, considered, and nontokenistic involvement of target end users in the design and development of a technology. In addition to boosting the voices of marginalized groups, co-design increases the acceptability of the end product [32]. Because its primary purpose is to support the delivery of the BMC Youth Model, the dedicated HIT would not provide diagnoses and medical advice; instead, it would support the provision of diagnoses and medical advice based on the specifications of the mental health services in which it is implemented. In other words, each service can customize the care options recommended by the dedicated HIT and the conditions under which these care options are recommended.

**Operating a “Flipped Clinic”**

Although the predicted surge in demand following the COVID-19 pandemic will exacerbate existing difficulties faced by young people in accessing quality mental health care and by service providers in delivering said care, this can be mitigated with HIT-enabled demand management strategies. In conventional service delivery models, in-clinic care tends to be emphasized and provided regardless of symptom severity and persistence, with web-based service delivery modalities (eg, voice or video calling) offered as “backup.” A redesigned service delivery model that “flips” to a default of web-based service delivery, and the integration of HITs into service delivery pathways to take advantage of their enhanced assessment, care coordination, and monitoring capabilities, could result in quicker waiting times (or even no more waitlists) and more efficient resource allocation. Offering young people at earlier and milder stages of illness (ie, Stage 1a) the opportunity to access psychoeducation and web-based interventions with the support of a junior clinician would free up service resources (including senior clinician time), enabling...
them to be reallocated toward delivering higher-intensity interventions to young people at more severe stages of illness. Figure 1 shows a redesigned HIT-enabled mental health service model for primary youth (aged 12 to 25 years) that adopts our digital health solution. First, standardized service entry uses web-based assessment, triage, and sophisticated suicide escalation protocols [25] to help immediately determine best care pathways based on the young person’s risk and clinical stage. Care decisions are driven by the young person’s needs and focus on getting the right team involved as well as on identifying the type and length of care required, modes of delivery, and need for further assessment. Routine outcome monitoring also enables the care team to respond to the young person’s changing needs (eg, increasing the intensity of care or changing the care team); this outcome monitoring can occur at a micro (daily or weekly) or macro (monthly or yearly) level.

Table 1 provides further operational details on how each service process could be digitally “flipped” with specific reference to the dedicated HIT, and it outlines the minimum functionalities the HIT should possess. In addition to detailing the intake, assessment, triage, care coordination, tracking, and monitoring processes depicted in Figure 1, Table 1 adds an additional step for service quality improvement, whereby key performance data can be regularly reviewed by service management who can then enact changes to service delivery to improve the quality of mental health care.
Table 1. Operational protocol demonstrating how to “flip” a primary youth mental health service using our digital health solution.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Actions</th>
</tr>
</thead>
</table>
| Intake  | The young person or their supportive other contacts the service (eg, telephone, web, email, walk-in).  
          | The service conducts an intake screen and invites them to use the digital health solution (via the dedicated HIT).  
          | The young person accepts the invite and sets up an account on the dedicated HIT. |
| Assess  | The young person completes a web-based multidimensional assessment that covers key mental health domains as well as social and occupational function, physical health, and substance misuse (20-40 minutes), ideally within 72 hours.  
          | The young person invites their supportive others to also contribute data through a shorter “summary” assessment (5 minutes). |
| Triage  | Triage is conducted by a senior clinician, such as a psychiatrist (and registrar), clinical psychologist, or mental health nurse.  
          | Triage is determined by real-time “escalations” that trigger upon detection of clinical risk (eg, meeting a certain threshold for suicidal thoughts/behaviors or abnormal mental states such as mania and psychosis), clinical staging, and current level of need.  
          | Triage is completed the next business day after a young person has finished the web-based multidimensional assessment.  
          | Urgent cases (suicidal thoughts/behaviors, mania, psychosis) are prioritized to be seen immediately using video-visit functionality [23], of which a small proportion are referred straight to acute care services. |
| Care    | Ongoing care pathways are matched to the appropriate type, intensity, and duration of intervention [29]. Using a multidisciplinary care team approach, these critical decisions are made early in the care pathway by senior clinicians to ensure accurate and efficient allocation of young people to care. This represents a secondary “flip” in the clinic, whereby senior clinicians are involved earlier in the care pathway as opposed to later.  
          | Because Stage 1b cases are at greater risk of transitioning to more severe stages of illness compared to Stage 1a cases [33], these should be reviewed with video-visit functionality [23] and directed to use care options in partnership with their multidisciplinary care team for a minimum of 12 months.  
          | Stage 2+ cases receive more specialist care in face-to-face settings for a minimum of two years. |
| Track   | Active tracking of symptoms/functioning by encouraging young people (and invited supportive others) to complete a “check-in assessment” at least every 21 days.  
          | Innovative use of the dedicated HIT should be considered, wherein a service could use the video-visit functionality to track young people in real time (eg, up to three 10-minute sessions per week). |
| Monitor | In conventional primary youth mental health services, only 20%-30% of cases show reliable improvement; 10%-25% of cases will deteriorate significantly over approximately six months; and, the majority of cases are left with persistent distress and/or impairment (i.e. no change) [22].  
          | Therefore, the dedicated HIT should be used for real-time review of deteriorating or non-changing cases through routine outcome monitoring that encourages care plans to change in response to outcome data, such as changing the type, intensity, and duration of intervention. |
| Review  | As part of a service’s “quality improvement cycle,” management could review service-level data collected by the dedicated HIT during routine care to evaluate (overall and by clinician) the clinical safety; accessibility and equity; effectiveness and outcomes; acceptability and satisfaction; efficiency, expenditure and cost; appropriateness; continuity and coordination; and workforce competence and capability [24]. |

Further Considerations for a “Flipped Clinic”

Our proposed “flipped clinic” service delivery model was developed following many years of clinical and co-design research with participants in the Australian mental health system, including consumers and their supportive others, health professionals, and service providers, to address major problems faced by users of this system: a limited supply of mental health professionals who are burdened with administration and assessment [22]; a lack of communication between mental health providers [4,18]; and the “tyranny of distance” [18,21]. This model proposes an innovative redesign of mental health service delivery that harnesses the capabilities of digital technologies (including, but not limited to, eHealth or mobile health [mHealth] interventions and other HITs) to enhance the provision of care for all participants in the system. However, some key considerations remain for those wishing to adopt this model or to enact similar widespread digital health system reforms.

As this model was developed using the Australian health system (a predominantly fee-for-service system) as a reference, it is likely to be generally applicable to similar systems. However, further work would still be required to adapt the model to a different country (with differences in service delivery mechanisms as well as facilitators and barriers to mental health care delivery) or funding model (eg, blended or pay-for-performance approaches). Within our digital solution, the “quality improvement cycle” (Table 1) could be particularly relevant for payment systems that are not primarily fee-for-service, and it could be further developed.

Our digital solution was also developed primarily from research activities at and around a primary youth mental health service aimed at providing care to Australian young people aged 12-25 years. The developmental heterogeneity of this age group has been widely researched [34,35]; as such, any solution for this
age group would have to cater to a mixture of children (aged <15 years), adolescents (aged 15-18 years), and transition-aged youth (aged 18-25 years) with a diversity of life experiences, clinical presentations, and clinical trajectories. Because our research team was working with a mental health service that specialized in treating youth, we were able to integrate our digital solution into existing service protocols for youth of different ages. However, other services or systems looking to adopt our solution may be required to adapt our approach, for example by developing different versions of our protocol for different youth age groups, and involving supportive others more heavily for young people below a certain age (with this age varying across jurisdictions and cultures).

Finally, although our digital solution was developed to address inequalities in access to youth mental health care in Australia, such as the “tyranny of distance” faced by young people in rural and remote areas [18,21], it does not yet address other existing inequalities, such as English literacy and different levels of access to digital technologies or the internet. A mixture of solutions will be required to address these inequalities, such as translating the dedicated HIT to different languages or designing offline functionality. However, these inequalities are also inherent to the broader Australian mental health system [36], and widespread government action will be required to address them (such as through infrastructure upgrades to improve nationwide internet connectivity).

Conclusion

Our digital health solution of digitally “flipping” clinics is a promising new demand management strategy for primary youth mental health services that aims to provide quality mental health care by leading consumers through a rapid experience of service entry, comprehensive assessment, multidisciplinary care, and routine outcome-based monitoring. Adopting this solution and establishing the required technological infrastructure could increase efficiencies in accessing and delivering quality mental health care and could also enable a dynamic response to public health crises such as the COVID-19 pandemic.

Acknowledgments

The authors would like to acknowledge the Australian Government Department of Health for funding Project Synergy (2017-2020), which aims to transform mental health service through the use of new and emerging technologies. Project Synergy (2017-2020) is being delivered by InnoWell Pty Ltd in a joint venture between The University of Sydney and PwC Australia. The authors also thank headspace Camperdown, a primary youth mental health service operated by The University of Sydney’s Brain and Mind Centre, for their input into the flow chart and operation protocol.

Authors’ Contributions

TD, VWSC, FI, and IH wrote and revised the manuscript. All authors contributed to developing the protocol. All authors read, revised, and approved the final manuscript.

Conflicts of Interest

IH was an inaugural Commissioner on Australia’s National Mental Health Commission (2012-2018). He is the Co-Director, Health and Policy at the BMC at The University of Sydney. The BMC operates early-intervention youth services at Camperdown under contract to headspace. IH is the Chief Scientific Advisor to, and a 5% equity shareholder in, InnoWell Pty Ltd. InnoWell was formed by The University of Sydney (45% equity) and PwC (Australia; 45% equity) to deliver the $30 million Australian government–funded Project Synergy (2017-2020; a three-year program for the transformation of mental health services) and to lead transformation of mental health services internationally through the use of innovative technologies.

References


Abbreviations
BMC: Brain and Mind Centre
ED: emergency department
HIT: health information technology
MBS: Medicare Benefits Schedule
mHealth: mobile health
PwC: PricewaterhouseCoopers

©Tracey A Davenport, Vanessa Wan Sze Cheng, Frank Iorfino, Blake Hamilton, Eva Castaldi, Amy Burton, Elizabeth M Scott, Ian B Hickie. Originally published in JMIR Mental Health (http://mental.jmir.org), 15.12.2020. 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 Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.
Original Paper

Real-time Mental Health Impact of the COVID-19 Pandemic on College Students: Ecological Momentary Assessment Study

Evan M Kleiman, PhD; April L Yeager, BA; Jeremy L Grove, PhD; John K Kellerman, BA; Joanne S Kim, BA
Rutgers, The State University of New Jersey, Piscataway, NJ, United States

Corresponding Author:
Evan M Kleiman, PhD
Rutgers, The State University of New Jersey
53 Avenue E
Tillett Hall Room 627
Piscataway, NJ, 08854
United States
Phone: 1 2153596148
Email: evan.kleiman@rutgers.edu

Abstract

Background: College students’ mental health may be disproportionately affected by the COVID-19 pandemic because of the abrupt shift off campus and subsequent loss of a social network and potential long-term impact on job prospects.

Objective: We sought to assess the nature of COVID-19’s mental health impact among a sample of undergraduates who were experiencing the pandemic as it occurred in real time.

Methods: In total, 140 college students completed smartphone-based ecological momentary assessments of anxiety and optimism related to COVID-19 and other generic mental health variables 6 times daily.

Results: Participants completed >23,750 surveys. Overall, >75% of these surveys indicated at least some level of anxiety about COVID-19. On average, the proportion of responses each day at the highest levels of anxiety about COVID-19 was 7 times greater than the proportion of responses at the highest levels of non–COVID-19–specific anxiety. Structural change analyses indicated a significant downward trend in COVID-19 anxiety after the first week of June, but even at the lowest point, >15% of the participants in the sample still reported high levels of COVID-19 anxiety each day. Participants felt more anxious about COVID-19 on days when the number of new cases and deaths due to COVID-19 were higher. When participants felt anxious about COVID-19, they also felt sad, anxious (in general), and had a greater desire to drink and use drugs. Participants felt more optimistic about COVID-19 when they received more support from others and from their university.

Conclusions: This study demonstrated the widespread mental health impact that COVID-19 has had on college students.

(JMIR Ment Health 2020;7(12):e24815) doi:10.2196/24815

KEYWORDS
ecological momentary assessment; college students; COVID-19, anxiety; real-time; mental health; impact; student

Introduction

Background

The COVID-19 pandemic has had a large impact on mental health [1-4] particularly among those on the front line combating the pandemic [5,6] and those who are staying at home due to social distancing mandates and have been displaced from routine activities and social contact (eg, young children [7] and older adults [8]). One group that has received relatively less attention but whose mental health may be disproportionally and uniquely impacted by COVID-19 in the short term and long term is college students [9-11]. In the early weeks of the pandemic hitting the United States, college students faced abrupt closures of schools and subsequent displacement from their on-campus housing. The economic downturn in the later phases of the pandemic influenced students’ ability to afford returning to campus (eg, if students’ on-campus jobs were eliminated), to obtain internships, and to procure stable employment after graduation. These impacts place a large mental health burden on college students.

Some empirical research on college student mental health during COVID-19 has been published. However, with few exceptions [12], this work has been predominantly cross-sectional [13,14]. Although it is useful to determine the impact of the pandemic...
at one point in time, this research does not allow us to understand the dynamics of the mental health impact of COVID-19. The COVID-19 pandemic was (and still is, as of December 2020) evolving daily and the impact on mental health seen one day could be quite different from what was seen on another day. Accordingly, the goal of this study is to explore how the rapidly changing impact of COVID-19 on mental health evolves over time using high-resolution, real-time data collection.

Research Questions

What is COVID-19’s Impact on Mental Health?

We assessed two COVID-19–specific mental health variables: anxiety about COVID-19 and optimism about COVID-19. Given the rapidly changing nature of the COVID-19 pandemic, we wanted to capture, over time, the severity and reach of the mental health impact (eg, “What percent of students experience severe COVID-19 anxiety throughout the day?”). We also wanted to see how these COVID-19 variables compared to other non–COVID-19–specific mental health variables, such as anxiety or worry. This allowed us to determine whether the pandemic’s mental health impact differed (eg, in rates, severity) from other mental health experiences that college students may face. We had no specific a priori hypotheses about this question, since it is primarily descriptive.

What Factors Characterize Periods of Anxiety and Optimism About COVID-19?

We were interested in the role of day-level contextual factors that could characterize days when the mental health impact of COVID-19 would be particularly severe. Such information can inform when and how to prevent or reduce any deleterious effects of COVID-19 on mental health. We focused on constructs that reflected exposure to negative news about COVID-19, including the number of new COVID-19 cases and deaths due to COVID-19 and the amount of COVID-19–related media consumed each day. We expected that participants would report higher anxiety and lower optimism on days when the numbers of new COVID-19 cases and/or deaths were higher and when they consumed more news related to COVID-19. One cross-sectional study conducted during the pandemic found that greater social media usage (not specific to COVID-19) was associated with higher anxiety [15].

We were also interested in the role that interpersonal and institutional support regarding COVID-19 could have in blunting the adverse mental health effects of COVID-19. Natural disasters have a lesser impact on the mental health of those who perceive more social support [16,17]. We assumed the same may be true here: those who perceive more support may be less impacted by COVID-19. Social support may be particularly relevant to undergraduates during the pandemic given that their on-campus social networks were quickly and unexpectedly disrupted. We also expected that support from institutions, especially the university, may play a role in how students respond to the pandemic. During the pandemic, students were faced with incredible uncertainty about whether they would be able to receive the same education remotely as they would in person, as well as when and how they would be able to return to campus. Institutions like universities play a large role in providing support around this uncertainty (eg, by creating resources to help students cope) and thus we expected that greater perceived institutional support would be associated with a reduced impact on mental health. To our knowledge, this is the first study to assess daily perceptions of how supported students feel from institutions like the university they attend.

What Are the Proximal Consequences of Anxiety and Optimism About COVID-19?

Although anxiety and optimism about COVID-19 are relevant end points, it is likely that COVID-19’s mental health impact does not stop at anxiety and optimism specific to COVID-19. We thought it was likely that anxiety and optimism about COVID-19 would lead to other proximal mental health consequences, such as extended periods of anxiety and attempts to cope, possibly in maladaptive ways. Understanding these consequences further allows us to characterize COVID-19’s mental health impact. Specifically, we were interested in whether anxiety and optimism about COVID-19 would have any appreciable impact on other mental health variables (ie, ratings of sadness or anxiety), health behavior variables (ie, urge to drink and urge to use drugs), and interpersonal variables (ie, feeling connected to others). We hypothesized that COVID-19 anxiety would be positively associated with adverse mental health consequences (sadness, anxiety, urge to drink, urge to use drugs) and negatively associated with social connection. We expected the opposite pattern for optimism about COVID-19.

Methods

Recruitment

Included in this manuscript are 140 participants from an ongoing study who were recruited between April 24 and May 26, 2020. The sample of 140 came from a total sample of 143 participants who completed the consent form and baseline; of the 143 participants, 3 did not complete the smartphone monitoring portion and were thus excluded from the study.

Procedure

Recruitment and Baseline

Participants were recruited remotely (stay-at-home orders were in place before the beginning of the study) from the undergraduate psychology pool and several large psychology classes. Participants first completed a screener for primary study inclusion criteria (aged ≥18 years, compatible smartphone, willing and able to do the surveys). They then completed a baseline assessment that assessed demographics and other constructs not used in this manuscript. After baseline, they received instructions for the app (MetricWire) we used to send the smartphone surveys.

Smartphone Surveys

Using their smartphones, participants completed a brief (<5 minutes) ecological momentary assessment (EMA) assessing general affect at the present moment (eg, anxious, worried) on a scale from 0 (not at all present) to 10 (very much), 6 times per day at random times within prespecified windows. This
survey also included a COVID-19–specific anxiety question ("How worried or anxious are you about the coronavirus outbreak?") that was answered on a scale from 0 (not at all) to 5 (very much).

In addition to the surveys assessing the current moment, the last assessment of each night included questions that asked participants to reflect over the entire day. In this study, we used 3 COVID-19–specific items that asked participants to rate their answers to the following questions on a scale from 0 (not at all) to 5 (very much).

1. How frequently did you see or read news or media about COVID-19 today?
2. Which of the following best describes how supported you feel by friends, family, or other individuals you know in dealing with the coronavirus outbreak?
3. Which of the following best describes how supported you feel by groups, organizations, or institutions you belong to in dealing with the coronavirus outbreak (eg, school, workplace, religious institution, community organization)?

This assessment also measured optimism about COVID-19 ("What best describes how optimistic you feel about the coronavirus outbreak?") on a scale from 1 (very pessimistic) to 5 (very optimistic).

**Compensation**

For the baseline, participants were given the option to receive a US $15 Amazon gift card or extra class credit. For the EMA portion, participants were paid $0.25 (in the form an Amazon gift card) for each of the surveys except for the longer nightly portion, participants were paid $0.50 (in the form an Amazon gift card) for each of the surveys except for the longer nightly survey, for which they were paid $0.50. If participants completed ≥4 surveys each day, they received a $0.50 bonus. In total, participants had the opportunity to be paid as much as $141.

**COVID-19 Data Extraction**

We used data on the number of new COVID-19 cases and deaths from COVID-19 for each day in the United States and New Jersey (nearly all participants resided in New Jersey during the study). These data were obtained using the COVID-19 R package [18], which uses the COVID-19 Data Hub to obtain gold-standard COVID-19 data from the repository hosted by the Center for Systems Science and Engineering at Johns Hopkins University [19].

**Statistical Analysis**

**What is COVID-19’s Impact on Mental Health?**

As COVID-19 anxiety was measured at a higher resolution than COVID-19 optimism (momentary versus daily), we primarily focused on anxiety for the descriptive analyses. To explore trends over time, we calculated for each day the percentage of non-zero and most-severe responses and the intraclass correlation (ICC) showing the amount of variability from observation-to-observation (versus person-to-person). To put the COVID-19–specific variables into context, we calculated similar statistics for two similar EMA questions: ratings of general worry and anxiety.

**What Factors Characterize Periods of Anxiety and Optimism About COVID-19?**

We were interested in which day-level contextual variables affected COVID-19 anxiety and optimism, specifically the following: the number of new cases/deaths announced each day in New Jersey (nearly all participants were in the state), perceived support related to COVID-19 from others and from organizations, and time spent consuming news media related to COVID-19. We conducted separate models with daily average level of COVID-19 anxiety (ie, the mean of all ratings for each participant on each day) and daily level of COVID-19 optimism as outcome variables.

All analyses for this aim have two levels: days (level-1) nested within people (level-2), which were tested using the lme4 [22] R package. We centered on person means of all self-reported predictors (support from others and organizations, frequency of upsetting news about COVID-19).

**What Are the Proximal Consequences of Anxiety and Optimism About COVID-19?**

We conducted contemporaneous and temporal multilevel models for each set of associations between COVID-19 anxiety or optimism and one of the outcome variables (anxious, sad, desire to use drugs, use alcohol, feeling close/connected to others). We restricted temporal analyses regarding COVID-19 anxiety to response pairs that were spaced by <6 hours. Since COVID-19 optimism was assessed once daily, we aggregated the momentary variables to create a daily average for each construct. We restricted the temporal analyses to pairs of consecutive days. As in the prior aim, we used the lme4 package with person-centered predictors.

**Results**

**Overview**

Between April 24 and July 19, 2020, 140 participants contributed 23,793 data points. Participants answered at least one survey on 5728 days (ie, each participant, on average, gave at least one survey response on a total of 40.91 days). This led to a response rate of 69.23% of surveys completed across those days.

**Demographics**

Among the 140 participants, 77.6% (n=109) were female and the average age was 19.98 years (SD 1.61, range 18.44-33.23). Regarding race, 48.59% (n=68) of the sample identified as White, 36.62% (n=51) Asian, 7.04% Black/African American, 1.41% American Indian/Alaskan Native, and the remaining 6.34% identified as multiple or other races. At the baseline assessment, 69.2% of the sample lived either on campus in the dorms or immediately off-campus and 30.1% of the sample were commuter students (eg, living with family). Although none
of the participants reported COVID-19 diagnoses at baseline, 1.4% of the sample reported experiencing symptoms of COVID-19. A large proportion of the sample reported knowing someone who had a COVID-19 diagnosis or symptoms; overall, 12.1% of the sample knew someone who had been diagnosed with COVID-19 and 5.7% of the sample knew someone who was experiencing potential COVID-19 symptoms.

What is COVID-19’s Impact on Mental Health?

Figure 1 shows the percent of daily responses each day that were at the highest levels of the scale for anxiety about COVID-19 and non–COVID-19–specific anxiety. The proportion of the sample endorsing the highest levels of COVID-19 anxiety on any given day was, on average, 6.89 (SD 4.12) times the rate of those endorsing the highest levels of anxiety.

The structural change model identified three segments with breakpoints at the 38th (June 8, 2020) and 64th (July 4, 2020) day. The first segment, from May 3 to June 8, 2020, was statistically flat ($b<0.001$, $t=-1.16$, $P=0.26$, $M=23.5\%$ of responses were at the highest level of the scale). The second segment, from June 9 to July 3, 2020, showed a decrease in worry ($b=-0.002$, $t=-3.69$, $P=0.001$, $M=15.3\%$). The third segment, from July 4 to 9, 2020, was also statistically flat ($b=0.009$, $t=-0.53$, $P=0.62$). However, the average proportion of people who had high levels of anxiety about COVID-19 each day during this third segment was significantly higher than the prior period (24.78% versus 15.3%, $t=3.78$, $P=0.006$).

Multimedia Appendix 1 shows participant-level data across the study. These data followed the same trend as the day-level data. Across the study, significantly more responses were non-zero ($\chi^2=5023.4$, $P<.001$) for COVID-19 anxiety (78.5% of all responses were $>0$) than for “anxious” (47.0%). Similarly, a greater number of people reported the highest levels of COVID-19 anxiety ($\geq4$ out of 5) at least once in the study (78.5%) than they did the highest levels of “anxious” (51.7% of participants reported a score of $\geq8$ out of 10 at least once). Nearly 75% of the variability in ratings of COVID-19 anxiety occurred from response-to-response (ICC=0.74, 95% CI 0.70-0.79). This ICC is significantly higher (ie, because the confidence intervals do not overlap) than the within-person variability for ratings of “anxious” (ICC=0.51, 95% CI 0.45-0.57).

What Factors Coincide With Anxiety and Optimism About COVID-19?

The left column of Table 1 shows which factors coincided with daily anxiety about COVID-19. Anxiety about COVID-19 was positively associated with the number of new cases announced in the state each day and the amount of news consumed relating to COVID-19. Counterintuitively, perceived support from others regarding COVID-19 was positively associated with daily anxiety about COVID-19. The right column of Table 1 shows which factors coincided with daily COVID-19 optimism. Daily optimism about COVID-19 was positively associated with both support from others and organizations, but unassociated with all other variables.

![Figure 1. Daily proportion of responses at highest levels of anxiety about COVID-19 and non–COVID-19–specific anxiety. Vertical lines indicate breakpoints in COVID-19 anxiety from the structural change model. Horizontal lines indicate mean during each segment.](image-url)
Table 1. Result of multilevel models showing which factors coincide with anxiety and optimism about COVID-19.

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Daily anxiety about COVID-19</th>
<th></th>
<th>Daily optimism about COVID-19</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Value</td>
<td>P value</td>
<td>Value</td>
<td>P value</td>
</tr>
<tr>
<td>Intercept, β (95% CI)</td>
<td>0.079 (–0.076 to 0.234)</td>
<td>.32</td>
<td>0.015 (–0.099 to 0.128)</td>
<td>.80</td>
</tr>
<tr>
<td>New cases in New Jersey, β (95% CI)</td>
<td>0.021 (0.008 to 0.034)</td>
<td>.002</td>
<td>–0.004 (–0.024 to 0.017)</td>
<td>.73</td>
</tr>
<tr>
<td>New deaths in New Jersey, β (95% CI)</td>
<td>–0.005 (–0.017 to 0.006)</td>
<td>.37</td>
<td>–0.015 (–0.033 to 0.004)</td>
<td>.12</td>
</tr>
<tr>
<td>Support regarding COVID-19 from others, β (95% CI)</td>
<td>0.027 (0.014 to 0.040)</td>
<td>&lt;.001</td>
<td>0.348 (0.329 to 0.368)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Support regarding COVID-19 from organization, β (95% CI)</td>
<td>0.008 (–0.005 to 0.021)</td>
<td>.21</td>
<td>0.092 (0.073 to 0.112)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Frequency of upsetting news regarding COVID-19, β (95% CI)</td>
<td>0.085 (0.073 to 0.097)</td>
<td>&lt;.001</td>
<td>–0.010 (–0.029 to 0.009)</td>
<td>.30</td>
</tr>
</tbody>
</table>

Random effects

| σ²                        | 0.39               | N/A                  | 0.68               | N/A                  |
| τ00                      | 2.04 participant  | N/A                  | 0.74 participant  | N/A                  |
| Intraclass correlation   | 0.84               | N/A                  | 0.52               | N/A                  |
| Marginal R² / conditional R² | 0.010 / 0.841    | N/A                  | 0.152 / 0.593      | N/A                  |

What Are the Proximal Consequences of Anxiety and Optimism About COVID-19?

Consequences of Anxiety About COVID-19

The first row of Table 2 shows the positive contemporaneous associations between anxiety about COVID-19 and ratings of anxiety, sadness, urge to drink, and urge to use drugs. The second row of Table 2 shows the temporal associations between anxiety about COVID-19 (at time T) and ratings of anxiety, sadness, urge to drink, and urge to use drugs (at time T+1). There was no association between anxiety about COVID-19 and desire to use drugs.

Consequences of Optimism About COVID-19

The third row of Table 2 shows the negative contemporaneous associations between optimism about COVID-19 and ratings of sad and urge to drink and the positive contemporaneous association between optimism about COVID-19 and feeling close/connected. There were no associations between optimism about COVID-19 and feelings of anxiety and the urge to use drugs. The fourth row of Table 2 shows the temporal associations between optimism about COVID-19 (at time T) and our outcome variables of interest (at time T+1). We found no significant temporal associations.

Table 2. Contemporaneous and short-term temporal associations between anxiety/optimism about COVID-19 and mental health, behavioral health, and connectedness dependent variables.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Dependent variable: anxiety</th>
<th></th>
<th>Dependent variable: sadness</th>
<th></th>
<th>Dependent variable: desire to use alcohol</th>
<th></th>
<th>Dependent variable: desire to use drugs</th>
<th></th>
<th>Dependent variable: close/connected</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>COVID-19 anxiety (contemporaneous)</td>
<td>0.055 (0.044 to 0.066)</td>
<td>&lt;.001</td>
<td>0.054 (0.043 to 0.064)</td>
<td>&lt;.001</td>
<td>0.037 (0.026 to 0.049)</td>
<td>&lt;.001</td>
<td>0.019 (0.009 to 0.029)</td>
<td>&lt;.001</td>
<td>0.015 (0.005 to 0.025)</td>
<td>.003</td>
</tr>
<tr>
<td>COVID-19 anxiety (temporal)</td>
<td>0.033 (0.021 to 0.046)</td>
<td>&lt;.001</td>
<td>0.035 (0.023 to 0.047)</td>
<td>&lt;.001</td>
<td>0.023 (0.010 to 0.036)</td>
<td>.001</td>
<td>0.011 (–0.010 to 0.022)</td>
<td>.07</td>
<td>0.015 (0.003 to 0.027)</td>
<td>.02</td>
</tr>
<tr>
<td>COVID-19 optimism (contemporaneous)</td>
<td>–0.004 (–0.024 to 0.017)</td>
<td>.72</td>
<td>–0.027 (–0.046 to –0.007)</td>
<td>.007</td>
<td>–0.024 (–0.046 to –0.002)</td>
<td>.03</td>
<td>0.007 (–0.009 to 0.023)</td>
<td>.40</td>
<td>0.015 (0.003 to 0.032)</td>
<td>.01</td>
</tr>
<tr>
<td>COVID-19 optimism (temporal)</td>
<td>–0.001 (–0.023 to 0.021)</td>
<td>.93</td>
<td>–0.003 (–0.023 to 0.017)</td>
<td>.78</td>
<td>0.00 (–0.023 to 0.022)</td>
<td>.97</td>
<td>0.013 (–0.004 to 0.031)</td>
<td>.13</td>
<td>0.007 (–0.012 to 0.026)</td>
<td>.47</td>
</tr>
</tbody>
</table>

Each row in each section (anxiety/optimism and contemporaneous/temporal) represents separate analyses. As optimism was assessed on the daily level, we aggregated (averaged) the momentary dependent variables for each day.


Discussion

Overview

We sought to better understand the nature of the mental health impact of the COVID-19 pandemic on a sample of college students as they were experiencing the pandemic in real time. We had three key goals: (1) to describe the variability of mental health variables related to COVID-19, (2) to identify which day-level factors coincided with anxiety and optimism about COVID-19, and (3) to identify the downstream consequences of the mental health impact of COVID-19. We discuss below the specific findings related to each of our three key questions and then conclude with a more general discussion of the study.

What is COVID-19’s Impact on Mental Health?

The findings regarding the frequency of anxiety about COVID-19 demonstrated that, compared to other similar states like nonspecific anxiety and worry, anxiety about COVID-19 in particular happened more often (non-zero responses occurred more than 75% of the time compared to less than 45% of the time for the other affect states), for more people (almost 80% of people reported the highest levels of COVID-19 anxiety at some point versus almost 50% of people reporting the highest levels of nonspecific anxiety), and changed more frequently throughout the day. This is notable because it means that students who otherwise may not be experiencing much distress are now experiencing at least some level of distress related to COVID-19. Although we did find a decrease in overall anxiety as time went on, even at its lowest point in the middle of June, more than 15% of all responses each day indicated severe anxiety about COVID-19. This increased anxiety among students could lead to a greater demand for counseling services as students adapt to the new reality of campus life during COVID-19. Methodologically, these findings are important because studies that assess only “anxiety” but not anxiety specifically related to COVID-19 may risk missing a sign of impaired functioning among students during this time.

What Factors Coincide With Anxiety and Optimism About COVID-19?

Participants reported greater anxiety about COVID-19 on days when the number of new cases announced in the state were higher and when they consumed more upsetting news specific to COVID-19. Some factors (the number of cases, deaths) are not under students’ control, while others are (eg, the frequency of watching upsetting news about COVID). This could point to the need to consume news in smaller quantities to avoid it having an undue impact on mental health [23].

Our findings highlighted the role of support during the pandemic, especially from organizations (eg, the university), which was positively associated with optimism about COVID-19. It should be noted, however, that literature on traumatic events [24] cautions that when universities do respond to events like this, care should be taken not to assume that all students are experiencing the event in the same way or experience it adversely at all. Moreover, although feeling more optimistic about COVID-19 may have potential mental health benefits, such optimism, especially if unrealistic, may carry potential health risks. There is a long history of literature on the idea that unrealistic optimism leads people to underestimate their risk of a negative health outcome and thus to engage in risky health behaviors [25,26]. Thus, in this case, those who are more optimistic about COVID-19 may underestimate their risk of exposure, leading them to engage in behaviors that would actually increase their risk of exposure.

Interestingly, we found that students feel more (rather than less) anxious on days when they perceived more support about COVID-19 from others. This could indicate that days of high anxiety are times when others would need to be more responsive. In other words, it may be that anxiety about COVID-19 leads to more support from others (to address that anxiety) rather than the other way around.

What Are the Proximal Consequences of Anxiety and Optimism About COVID-19?

We found that feeling anxious about COVID-19 now is associated with distress (anxiety, sadness) and negative health behaviors (the desire to drink and use drugs) both in the moment and a few hours later. The increased urge to drink and use drugs may indicate an elevated likelihood of coping with this distress in maladaptive ways. We also found that optimism about COVID-19 was similarly associated with most same-day consequences but was not associated with any of the outcome variables when examined one day later. The lack of next-day effects for COVID-19 optimism could reflect some difference between optimism and anxiety, but it is probably more likely that this reflects a difference in the timescale of interest (ie, we measured anxiety related to COVID-19 six times daily and optimism once daily, about the entire day). This could mean that the direct mental health impact of COVID-19 is most relevant in the short term. This echoes the earlier findings in this study that these constructs are highly variable over a short period of time. It should be noted that although the direct impact is only present for a few hours, the cumulative effect of anxiety associated with COVID-19 could be deleterious in the long term in ways not assessed here (eg, through increased drinking over months, a transition into more general anxiety).

Limitations and Strengths

There were several limitations to this study that should be acknowledged. First, because this study began during the COVID-19 pandemic, we were not able to establish a “baseline” for the level of worry or anxiety that students typically experience. Second, although we had >23,750 responses, these responses only came from 140 students at one university. Although the sample was ethnically and racially diverse, the experiences of these students may not generalize to all students. There were several strengths of this study. It is the most fine-grained observation of COVID-19’s mental health impact to date and one of the few studies to explicitly collect data on COVID-19–relevant constructs (eg, anxiety specifically about COVID-19).

Conclusions

There are several tangible recommendations that come from this work. Students are anxious about this pandemic and the university may be a key source of support during this time.
College counseling centers may need to provide COVID-19–specific coping skills, which would be particularly useful for those students who are not typically anxious (and for whom coping skills for general anxiety/worry might not resonate). Moreover, it is possible that substance use behavior could increase in the context of a crisis and there is a need to assess for this as the pandemic continues. Finally, even after the pandemic is under control, there are long-lasting consequences that may continue to contribute to anxiety among college students (eg, finding a job after college, paying for college after the economic impact of COVID-19); further research is needed on these consequences. Such research will inform universities’ long-term recovery plans, which may need to help students address this anxiety.

Conflicts of Interest
None declared.

Multimedia Appendix 1
COVID-19 anxiety and non–COVID-19–specific anxiety by participant. Only participants with >50 responses are shown. Anxiety about COVID-19 was multiplied by 2 because it was on a 0-5 scale and not a 0-10 scale. Multilevel correlation between COVID-19 anxiety and non–COVID-19–specific anxiety was small (r=.11, P<.001).

References


Abbreviations

EMA: ecological momentary assessments
ICC: intraclass correlation

Edited by J Torous; submitted 06.10.20; peer-reviewed by J Dworkin, R Shalaby; comments to author 20.10.20; revised version received 25.10.20; accepted 05.11.20; published 15.12.20.

Please cite as:
Kleiman EM, Yeager AL, Grove JL, Kellerman JK, Kim JS
Real-time Mental Health Impact of the COVID-19 Pandemic on College Students: Ecological Momentary Assessment Study
JMIR Ment Health 2020;7(12):e24815
URL: http://mental.jmir.org/2020/12/e24815/
doi:10.2196/24815
PMID:33207308

©Evan M Kleiman, April L Yeager, Jeremy L Grove, John K Kellerman, Joanne S Kim. Originally published in JMIR Mental Health (http://mental.jmir.org), 15.12.2020. 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 Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.
Patient Attitudes Toward Telepsychiatry During the COVID-19 Pandemic: A Nationwide, Multisite Survey

Daniel Guinart, MD; Patricia Marcy, BSN; Marta Hauser, PhD; Michael Dwyer, MA; John M Kane, MD

1Department of Psychiatry, The Zucker Hillside Hospital, Northwell Health, New York, NY, United States
2Center for Psychiatric Neuroscience, Feinstein Institutes for Medical Research, Manhasset, NY, United States
3The Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Manhasset, NY, United States
4Vanguard Research Group, Glen Oaks, NY, United States
5Ambulatory Care Division, The Zucker Hillside Hospital, Northwell Health, New York, NY, United States

Corresponding Author:
Daniel Guinart, MD
Department of Psychiatry
The Zucker Hillside Hospital
Northwell Health
7559 263rd St
New York, NY
United States
Phone: 1 7184704139
Email: daniguinart@gmail.com

Abstract

Background: The COVID-19 pandemic and its associated movement restrictions forced a rapid and massive transition to telepsychiatry to successfully maintain care continuity.

Objective: The aim of this study is to examine a large number of patients’ experiences of, use of, and attitudes toward telepsychiatry.

Methods: An anonymous 11-question survey was delivered electronically to 14,000 patients receiving telepsychiatry care at 18 participating centers across 11 US states between the months of April and June 2020, including questions about their age and length of service use, as well as experience and satisfaction with telepsychiatry on a 5-point Likert scale. Descriptive statistics were used to analyze and report data.

Results: In total, 3070 patients with different age ranges participated. The overall experience using telepsychiatry was either excellent or good for 1189 (82.2%) participants using video and 2312 (81.5%) using telephone. In addition, 1922 (63.6%) patients either agreed or strongly agreed that remote treatment sessions (telephone or video) have been just as helpful as in-person treatment. Lack of commute (n=1406, 46.1%) and flexible scheduling/rescheduling (n=1389, 45.5%) were frequently reported advantages of telepsychiatry, whereas missing the clinic/hospital (n=936, 30.7%) and not feeling as connected to their doctor/nurse/therapist (n=752, 24.6%) were the most frequently reported challenges. After the current pandemic resolves, 1937 (64.2%) respondents either agreed or strongly agreed that they would consider using remote treatment sessions in the future.

Conclusions: Telepsychiatry is very well perceived among a large sample of patients. After the current pandemic resolves, some patients may benefit from continued telepsychiatry, but longitudinal studies are needed to assess impact on clinical outcomes and determine whether patients’ perceptions change over time.

(JMIR Ment Health 2020;7(12):e24761) doi:10.2196/24761

KEYWORDS

telehealth; telepsychiatry; telemedicine; attitude; patients; survey; COVID-19; mental health

Introduction

The outcomes and cost-effectiveness of telepsychiatry are overall comparable to in-person care across multiple treatment modalities, disorders, and patient groups [1-11]. However, widespread implementation of telepsychiatry has been challenging [12-14], partially due to mental health care professionals’ concerns about patients’ ability to use
conferencing devices, lack of sense of closeness/connection, technical problems, and reimbursement and privacy concerns [15,16]. However, barriers related to patient preference are also possible, and the patient perspective is crucial to further characterize implementation challenges. Previous studies showed positive patient satisfaction [17-19] but potential limitations including relatively small sample sizes and/or selection biases in the context of pilot programs or specific services may have limited their generalizability.

Due to the COVID-19 crisis, our health care system and others around the world rapidly transitioned all or almost all in-person visits to remote assessments [20], in an unprecedented context of mental health care professional stress and increased need for mental health services [21,22]. This revolution in telepsychiatry provided a unique opportunity to assess how patients that may not have initially opted for telepsychiatry feel about it. Hence, the aim of this study was to qualitatively assess opinions and attitudes about telepsychiatry of a large sample of patients.

**Methods**

In collaboration with the Vanguard Research Group (VRG), a research consortium specializing in behavioral health, an anonymous survey was distributed to patients using telepsychiatry in 18 hospitals and community centers located in rural, suburban, small urban, and large urban areas in 11 different states across the United States (Connecticut, Florida, Maine, Michigan, New Hampshire, New York, Oregon, Rhode Island, South Carolina, Texas, and Utah). Surveys were distributed through email and/or embedded into the video platform scheduling invitations between the months of April and June 2020, and could be completed electronically, with computers, tablets, or smartphones. Study procedures were deemed exempt by the local Institutional Review Board (IRB#20-0397). Further details can be found on the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) [23] listed in Multimedia Appendix 1.

The survey included 11 questions about telepsychiatry use and satisfaction using a 5-point Likert scale, as well as inquiries about both potential challenges and positive experiences (see survey in Multimedia Appendix 2). Descriptive statistics were used to report qualitative survey results. Chi-square tests were used to compare categorical variables. First, omnibus comparisons were conducted by age range and length of care at the same institution. If statistically significant differences were detected, we then tested the individual interactions of interest post hoc. All analyses were conducted using JMP (Version 13, SAS Institute Inc).

**Results**

The survey was distributed to approximately 14,000 patients, of which 3070 (22%) completed it. In total, 18 surveys were excluded due to the subject disclosing not having used telemedicine. Hence, 3052 surveys were included in the analysis. Patient characteristics are listed in Table 1.

### Table 1. Characteristics of the patients included in the study.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Patients, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age range, years (N=3040)</strong></td>
<td></td>
</tr>
<tr>
<td>&lt;25</td>
<td>304 (10.0)</td>
</tr>
<tr>
<td>25-34</td>
<td>494 (16.3)</td>
</tr>
<tr>
<td>35-44</td>
<td>576 (18.9)</td>
</tr>
<tr>
<td>45-54</td>
<td>680 (22.4)</td>
</tr>
<tr>
<td>55-64</td>
<td>721 (23.7)</td>
</tr>
<tr>
<td>65-74</td>
<td>232 (7.6)</td>
</tr>
<tr>
<td>&gt;74</td>
<td>33 (1.1)</td>
</tr>
<tr>
<td><strong>Duration of care, years (N=2994)</strong></td>
<td></td>
</tr>
<tr>
<td>&lt;1</td>
<td>793 (26.5)</td>
</tr>
<tr>
<td>1-5</td>
<td>1335 (44.6)</td>
</tr>
<tr>
<td>5-10</td>
<td>493 (16.4)</td>
</tr>
<tr>
<td>&gt;10</td>
<td>373 (12.5)</td>
</tr>
</tbody>
</table>

Briefly, 55% of the sample (n=1666) were aged >45 years and the majority of participants (n=2128, 71.1%) had been under care at the institution where the survey took place for ≤5 years (Table 1). Respondents were mostly using telephone (n=1924, 63.7%), followed by video (n=708, 23.4%), and a combination of telephone and video (n=390, 12.9%). When asked about their preferred method of receiving care, respondents preferred the telephone over video (n=1908, 64.1% versus n=1066, 35.9%). The overall experience was either good or excellent for 2312 (81.5%) of respondents when asked about telephone only and for 1189 (82.2%) when asked about video (Figure 1). Only 127 (4.5%) and 74 (5.1%) respondents rated their experience as “poor” or “very poor” for telephone and video, respectively (Figure 1).
We detected differences in the overall experience by age range in the case of telephone, $\chi^2_{24} (n=2831)=46.3, P=.004$; a lower proportion of patients aged 55-64 years declared their experience as excellent compared to other age groups (n=257, 38.2% versus n=960, 44.3%), $\chi^2_{4} (n=2840)=12.8, P=.01$. In addition, a higher proportion of patients aged 45-54 years rated their experience as poor compared to other age groups (n=27, 4.2% versus n=55, 2.5%), $\chi^2_{4} (n=2840)=10.5, P=.03$.

Further, 1922 (63.6%) patients either agreed or strongly agreed with the statement that remote treatment sessions (telephone or video) have been just as helpful as in-person treatment, whereas 1937 (64.2%) of respondents either agreed or strongly agreed with the statement that they would consider using remote treatment sessions in the future (Figure 2). Patients using video were more likely to strongly agree with that statement than those using telephone (n= 570, 38.9% versus n= 273, 29.9%), $\chi^2_{4} (n=2605)=29.6, P<.001$.

Patients endorsed the lack of commute (n=1406, 46.1%), flexible scheduling/rescheduling (n=1389, 45.5%), reduced likelihood of missing appointments (n=1064, 39.9%), and feeling more confidence/comfort than in person (n=601, 19.7%) as positive elements/advantages of telepsychiatry (Table 2), which did not vary by age ($\chi^2_{9} [n=4447]=15.4, P=.64$) or length of time under care ($\chi^2_{5} [n=4413]=10.7, P=.30$). Some of the challenges that patients endorsed were related to missing the clinic/hospital (n=936, 30.7%) and not feeling as connected to their doctor/nurse/therapist (n=752, 24.6%), among others (Table 2). Patients under care for less than one year endorsed missing the clinic and feeling connected to it less frequently than other groups (n=195, 21.6% versus n=741, 28%), $\chi^2_{6} (n=3550)=21.5, P=.002$. 

Figure 1. (A) Patients' experience of receiving mental health care via telephone, rated on a 5-point Likert scale ranging from 0=very poor to 5=excellent. (B) Patients' experience of receiving mental health care via video, rated on a 5-point Likert scale ranging from 0=very poor to 5=excellent.

Figure 2. (A) Patients' degree of agreement with the statement "I feel the remote treatment sessions have been just as helpful as in-person treatment," rated on a 5-point Likert scale ranging from "strongly disagree" to "strongly agree." (B) Patients' degree of agreement with the statement "I would consider using remote treatment sessions in the future," rated on a 5-point Likert scale ranging from "strongly disagree" to "strongly agree."
In the free-text comment section, patients generally found telepsychiatry to be safe and convenient, and expressed their gratitude to mental health care professionals for providing uninterrupted care during a very challenging time. Many suggested remote assessments should be maintained, mentioning that they feel more comfortable at home, can express themselves more freely, save transportation time and costs, and/or request less time off work. Others expressed feeling disengaged, feeling frustrated with technical difficulties and having a lack of resources to address them (eg, not owning a laptop or smartphone), difficulty finding a quiet setting (eg, children interrupting, shared housing), getting tests done or filling out forms.

Discussion

Principal Findings

In this study, we report highly favorable attitudes toward telepsychiatry in its diverse forms, across a large sample of patients across the United States. To our knowledge, this is the largest evaluation of patient attitudes toward telepsychiatry to date, by at least an order of magnitude, which is timely in the context of the current COVID-19 pandemic and the widespread stay-at-home and travel restriction orders, the duration of which is unclear.

Our results are aligned with other surveys very recently validated based on quality of care domains [24], showing high levels of satisfaction with telepsychiatry services. Other recent studies in older [25,26] and younger [27] adults showed similar results, all in smaller samples. Further, most of our respondents would like to continue using telepsychiatry. This finding is highly relevant given the diversity and size of our sample, drawn from a large network of community, real-world, and academic mental health centers, and should encourage allowing telepsychiatry to continue for some patient populations after the current pandemic is resolved. However, some respondents expressed a desire to resume usual in-person care as soon as possible and/or lean toward hybrid models. The option of telepsychiatry should remain tailored to individual patient needs and be the result of shared decision making.

Interestingly, subjects were more likely to strongly agree to consider using telepsychiatry in the future when using video. Concerns raised about lack of closeness and fear of a reduction in the doctor’s ability to detect subtle signs of body language, nonverbal cues, and/or physical signs of disease could be some of the reasons behind this preference [16]. Whereas the widespread use of the telephone may be the result of an abrupt transition related to COVID-19, access to technology may have been a potential barrier to the implementation of telepsychiatry that will need to be considered. Videoconferencing should be preferred over telephone whenever possible, particularly given the currently available technology, which allows for encrypted private communications [15]. Further, patients with sensory and/or cognitive limitations such as mutism, hearing difficulty, or visual or cognitive impairment would potentially require deployment of additional technologies and/or human resources.

Limitations

This study has some limitations. First, this study was conducted during the COVID-19 pandemic and associated movement restrictions, which may have made hospital/doctor visits less appealing, adding safety as a confounder, possibly overestimating real user satisfaction. Second, our survey was short, the completion rate was relatively low, and our sample was not random, so selection, nonresponse, and response biases are possible [28]. Third, the influence of additional sociodemographic factors as well as symptom severity and/or previous telepsychiatry experience could not be ascertained. Longitudinal studies will be needed to assess impact on clinical
outcomes and determine whether patients’ perceptions change over time.

Mental health professionals were already implementing digital technologies and advocating for more widespread use of telehealth [29], and the current scenario has accelerated its use. Thus, even after the COVID-19 pandemic ends, telepsychiatry is here to stay. However, patient concerns need to be heard and addressed, and positive experiences need to be acknowledged and echoed.

Conclusion

Patients had a generally positive attitude toward telepsychiatry and many would like to continue using it after the COVID-19 restrictions recede. Longitudinal studies are needed to assess whether patient perceptions change over time. However, some patients may benefit from continuous use of telepsychiatry. Results of this study should help shape policies regarding its use.

Acknowledgments

We thank the patients that took the time to participate in our study in such challenging times.

Conflicts of Interest

DG has been a consultant for and/or has received speaker honoraria from Otsuka America Pharmaceuticals and Janssen Pharmaceuticals. JMK has been a consultant and/or advisor for or has received honoraria from Alkermes, Allergan, LB Pharmaceuticals, H Lundbeck, Intracellular Therapies, Janssen Pharmaceuticals, Johnson and Johnson, Merck, Minerva, Neurocrine, Neuron, Otsuka, Pierre Fabre, Reviva, Roche, Sumitomo Dainippon, Sunovion, Takeda, Teva, and UpToDate and is a shareholder in LB Pharmaceuticals and Vanguard Research Group. PM has been a consultant to Otsuka and has received research funding from Alkermes, Boehringer-Ingelheim, Janssen, Lundbeck, NeuroRx, Otsuka, Takeda, and Roche. The other authors declare no conflicts.

Multimedia Appendix 1

Checklist for Reporting Results of Internet E-Surveys (CHERRIES).

Multimedia Appendix 2

Telepsychiatry Patient Satisfaction Survey.

References


23. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]


**Abbreviations**

CHERRIES: Checklist for Reporting Results of Internet E-Surveys