



# HHS Public Access

Author manuscript

*Psychiatr Serv.* Author manuscript; available in PMC 2020 July 01.

Published in final edited form as:

*Psychiatr Serv.* 2019 July 01; 70(7): 538–544. doi:10.1176/appi.ps.201800519.

## User Engagement in Mental Health Apps: a Review of Measurement, Reporting and Validity

Michelle M Ng, BA<sup>1</sup>, Joseph Firth, Ph.D.<sup>2</sup>, Mia Minen, MD<sup>3</sup>, John Torous, MD<sup>1</sup>

<sup>1</sup>Division of Digital Psychiatry, Department of Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

<sup>2</sup>NICM, School of Science and Health, Western Sydney University, Australia

<sup>3</sup>NYU Langone Headache Center, Department of Neurology, NYU School of Medicine, New York, NY

### Abstract

**Objective:** Despite the potential benefits of mobile mental health apps, real world results indicate engageability issues due to low uptake and sustained use by patients. This study investigated how user engagement indicators of mental health apps are measured and reported.

**Methods:** A systematic review was performed in July 2018 of PsycINFO, Ovid MEDLINE, the Cochrane Central Register of Controlled Trials and AMED, Embase and HMIC databases for published papers presenting results of mental health apps for depression, bipolar disorder, schizophrenia and anxiety reporting user engagement indicators, namely “usability,” “user satisfaction,” “acceptability,” and “feasibility.” The subjective and objective criteria used to assess user engagement indicators, among other data, were extracted from each paper.

**Results:** The search returned 925 results from which 40 papers were found to be eligible. Every study reported positive results for the usability, satisfaction, acceptability and/or feasibility of the app in that study. 36 (90%) studies employed 371 indistinct subjective criteria that were assessed using surveys, interviews, or both. Of these, 23 studies employed custom subjective scales, as opposed to a standardized assessment tool. 24 studies (60%) employed objective criteria, amounting to 71 indistinct measures. No two studies used the same combination of subjective and/or objective criteria to determine the usability, satisfaction, acceptability and/or feasibility of their app.

**Conclusion:** The high heterogeneity and use of custom criteria to assess mental health apps in terms of usability, user satisfaction, acceptability or feasibility present a challenge for understanding real world low uptake of these apps. This discrepancy between every study claiming high user engagement indicators suggests a need for the field to focus on engagement through the creation of reporting standards and more careful consideration of claims.

---

jtorous@bidmc.harvard.edu.

**Disclosures and acknowledgements:** None of the authors declare any conflicts of interest.

**Previous presentations:** N/A

## INTRODUCTION

Mobile technologies are increasingly owned and utilized by people around the world. With this rise in pervasiveness comes the potential to increase access to and augment delivery of mental health care. This can occur in multiple ways, including patient-provider communication, self-management, diagnosis and even treatment (1). Early evidence concerning the efficacy of mobile mental health apps has created a wave of enthusiasm and support (2). The potential scalability of these app-based interventions has been proposed as a means of addressing the global burden of mental illnesses and offering services to those who are in need but previously have not been able to access care (3). Even in developed countries, where access to mental health services remains inadequate, app-based interventions have been proposed as innovative research, screening, preventive and care delivery platforms (4). The 10,000 mental health apps currently available for immediate download from the Apple iTunes or Google Android Play marketplaces speak to their easy availability, as well as people's high interest (5).

But neither potential, interest, nor availability alone has translated into the often forecasted digital revolution for mental health. While there are many possible explanations for this, one factor is the poor uptake of mental health apps (6). User engagement studies have the potential to shed valuable insight here. Many papers that evaluate mental health apps include a study of usability, user satisfaction, acceptability and/or feasibility. These "user engagement indicators" (UEI) are meant to represent the ability of an app to engage and sustain user interactions. However, the lack of guidelines, consensus, or specificity around "user engagement" in mental health research introduces the concerning potential for UEI to be selected inappropriately, presented with bias and/or interpreted incorrectly. Thus it is difficult to interpret, let alone compare or pool data on, engagement metrics related to these smartphone apps. For example, in one study (7), participants described an app as "buggy," "clunky," and "didn't really work" during qualitative interviews. Nevertheless, when the same participants were asked specifically whether the app was "user friendly" and "easy to use," five out of seven reported that the app was user friendly and easy to use. The study used the latter metrics as the basis of their conclusion that their app had positive UEI, masking potentially serious usability and safety concerns.

To both assess the current state of reporting and inform future efforts, we performed a systematic review about how the UEI of apps designed for people with depression, bipolar, schizophrenia and anxiety are evaluated. We hypothesized that there would be: one, confluences in the definitions and criteria for common types of UEI, namely "usability," "satisfaction," "acceptability," and "feasibility"; two, inconsistent subjective and objective criteria used to evaluate UEI; and three, inconsistent thresholds of UEI ratings across studies.

## METHODS

### Search string and selection criteria

We conducted a systematic search of PsycINFO, Ovid MEDLINE, the Cochrane Central Register of Controlled Trials and AMED, Embase and HMIC Databases, on 14th July 2018,

using terms synonymous with mobile apps for mental health. The full search algorithm is presented in Table 1.

Inclusion criteria was as follows: report original qualitative or quantitative data; primarily involve a mobile application; design for people with depression, bipolar, schizophrenia and/or anxiety (including post-traumatic stress disorder and obsessive-compulsive disorder); inclusion of a conclusion about user engagement indicators (including usability, satisfaction, acceptability and/or feasibility) for the app; and a study length of at least seven days. Exclusion criteria were: review, conference, protocol or dissertation papers; non-English language papers; and a lack of focus on the technologies or diseases of interest. All papers were screened by two authors (MN and JT) and any disagreements were resolved through discussion resulting in consensus.

### Data Extraction & Synthesis

A systematic tool was developed and the following data was extract by JT and MN:

*Study details:* study design (e.g. single-arm or RCT), sample size, inclusion criteria and clinical characteristics of participants.

*Intervention details:* Details of app, length of intervention, device type used

*Objective user engagement indicators:* usage frequency, response to prompts, trial retention

*Subjective user engagement indicators:* satisfaction questionnaires, interviews about usability, etc.

*Putative factors which could influence usability,* such as whether patients were involved in the app design process, incentives for participation, etc.

Institutional review board (IRB) approval was not required for this literature review.

## RESULTS

### Included Studies

The PRISMA chart in the online supplement shows the full study selection process. The initial database search returned 925 results. This was reduced to 882 after duplicates were removed. A further 778 articles were excluded after reviewing the titles and abstracts for eligibility. Full text versions were retrieved for 104 articles, of which 64 were ineligible for reasons shown in the PRISMA chart.

Thus, a total of 40 studies reporting UEI of mental health apps for people with mental illness were included (7–47). Of these, nine apps were designed for people with depression (11, 12, 19, 21, 27, 29, 40, 43, 46), four with bipolar (16, 20, 22, 47), seven with schizophrenia (23, 25, 28, 33, 38, 39, 45) and seven with anxiety (15, 26, 30, 31, 32, 35, 44). Thirteen apps were designed for two or more populations with a different mental illness (8, 9, 10, 13, 14, 17, 18, 24, 34, 36, 37, 41, 42)

The mean number of participants enrolled was 32 per study (range = 2 to 163). Of studies that report the length of study or mean length of study (since some studies last as long as participants want to use the app), the mean length of study was 58 days.

### **User Engagement Indicators: Usability, Satisfaction, Acceptability and Feasibility**

Every study performed an evaluation of the “usability,” “satisfaction,” “acceptability,” and/or “feasibility” of an app. While we refer to these criteria as UEI, the studies reviewed did not use UEI as a term or a framework.

Across studies, there were confluences in the definitions and criteria of “usability,” “satisfaction,” “acceptability,” and/or “feasibility.” Some papers referred to these types of UEI interchangeably. For example, multiple papers used the phrase “usability/acceptability” (23–24) (and another paper used the phrase “acceptability/usability” (25)). One refers to a “Satisfaction/Usability Interview” (44). Another first uses the phrase “Tolerability and usability” and later switches to “Acceptability and tolerability” (47). One paper first wrote, “Acceptability was measured by examining self-reports and user engagement with the program” but later stated, “Acceptability was measured by examining users’ self-reported attitudes and satisfaction” (43). Yet another paper used the Technology Acceptance Model to in part evaluate the usability of an app (44).

Some papers treated some UEI as determinants for others. One paper stated, “The BeyondNow app was also shown to be feasible given the high level of usability” (36). Another paper wrote, “To evaluate acceptability of using a smartphone application as part of EP outpatient care, participants completed self-report surveys at the end of the study evaluating satisfaction” (14). And under its subheading “Aim I–Feasibility: Mobile App Satisfaction,” one other paper reported, “Participants provided high usability ratings for the mobile app based on the SUS...” (8).

The majority of papers evaluate multiple UEI at once. Eight papers drew conclusions about one type of UEI (e.g. just usability) (10, 12, 18, 21, 27, 30, 37, 41), 11 papers about two types of UEI (e.g. feasibility and acceptability) (9, 17, 19, 29, 31, 36, 38, 40, 44, 46, 47), 11 papers about three types (11, 14, 15, 16, 20, 23, 24, 28, 34, 39, 45) and ten papers about four types (8, 13, 22, 25, 26, 32, 33, 35, 42, 43). Furthermore, most papers use the same criteria to evaluate multiple UEI. For instance, one paper described, “Satisfaction, usability and acceptability were calculated based on the percentage of answers of the Likert-scale” (22). The fact that the majority of studies used similar methods to evaluate more than one type of UEI speaks to the lack of precision and distinction between them.

### **Types of Criteria: Subjective and Objective**

The criteria used to draw conclusions about UEI varied widely across studies, as shown in Figure 3. 16 (40%) out of the 40 included papers concluded that their app had positive UEI based entirely on subjective criteria. Four studies (10%) concluded that their app had positive UEI based entirely on objective criteria. 20 studies (50%) concluded that their app had positive UEI based on a combination of subjective and objective criteria.

**Subjective Criteria**—In the 36 studies (90%) that considered subjective criteria, these subjective criteria manifested as 371 indistinct questions (see online supplement) and were assessed using multiple methods, such as a survey, an interview, or a survey and interview. As shown in Table 2, 13 studies derived inspiration from one or more pre-existing assessment tools. The remaining 23 studies did not rely on pre-existing assessment tools to evaluate subjective criteria, suggesting that they developed their own custom questions. This assortment of both subjective criteria and methodologies for evaluating UEI demonstrates that there is no one gold standard.

**Objective Criteria**—The 24 studies (60%) that evaluated UEI entirely or partially based on objective criteria relied on 71 indistinct measures of usage data (see online supplement). Of these 24 studies, five set a target usage goal in advance (8, 28, 34, 38, 39) and 19 considered their usage data retrospectively (9, 11, 13, 14, 16, 17, 20, 21, 22, 24, 25, 26, 33, 35, 36, 40, 42, 43, 45) to determine positive UEI.

Across all studies, a wide array of objective criteria was taken into account. This includes “average number of peer and coach interactions” (11), “length of time in clinic at enrollment” (14), “(reliable) logging of location” (19), “(number of) active users” (22) and “percentage of participants who were able to use both system-initiated (i.e., in response to prompts) and participant-initiated (i.e., on-demand) videos independently and in their own environments for a minimum of 3 days after receiving the smartphone” (33).

### Thresholds of UEI

All 40 studies (100%) concluded that their app had positive UEI. However, they came to the same conclusion in different ways: they evaluate the different types of UEI with different methodologies, from the criteria (such as subjective ratings and objective data) to the means of assessment (such as a survey, interview, or usage data). In other words, inconsistencies in the UEI evaluation process cast doubt on studies’ ability to claim that their app was “usable,” “satisfactory,” “acceptable,” and/or “feasible.”

**Subjective Criteria**—Due to the range of both subjective criteria and their evaluation methods, it is impossible to compare the ratings of UEI across papers. What is possible to see is that papers utilized different thresholds for surmising that their app had positive UEI. For example, of papers that evaluate the subjective criteria “ease of use,” the percentage of users reportedly satisfied with “ease of use” range from 60% (18) to 100% (13). Similarly, the satisfaction scores for “ease of use” range from 79.7% (46) to 92.6% (16). Despite the range of perceptions about the “ease of use” of an app, every study concluded that their app had positive UEI.

**Objective Criteria**—There were differences across studies in objective criteria such as target usage goals and frequency of usage. For example, of studies that set a target usage goal pertaining to task completion, two studies sought completion of over 33% of prompted tasks (28, 39) while another study sought completion of over 70% of prompted tasks (38). Despite this variability, they all concluded that the apps had positive UEI on the basis of this usage data. Similarly, studies that considered frequency of usage as an objective criterion

reported frequencies ranging from once per day (8) to once every other day (45) to an average of 5.64 times accessed by each participant over the course of two months (36). Similar to above results, each of these trials concluded that their apps had positive UEI.

**Discrepancies between thresholds of criteria**—Even when an app seems to meet the threshold for positive UEI based on subjective criteria, it might not meet the threshold based on objective criteria. One paper raises the issue of possible discrepancies arising from evaluating UEI based on solely subjective versus objective criteria: “Analysis of objective use data for another study utilizing PTSD Coach indicates that although app users report positive feedback on usability and positive impact on symptom distress, only 80% of first-time users reach the home screen and only 37% progress to one of the primary content areas” (15).

This is an issue not only within studies, but also across studies. For instance, five studies that used retention rate as an objective criterion for UEI reported retention rates of 80% (35), 83% (11), 91.5% (21), 100% (38) and 100% (45). Yet studies that did not rely on retention rate as a criterion had retention rates as low as 35% (13) and 65.7% (27). All of these papers concluded that their apps have positive UEI.

## DISCUSSION

Despite the real-world challenges of mental health app usability, engagement and usage, all 40 studies included in this study reported that their app had positive user engagement indicators. This occurred despite all papers: claiming that their app was “usable,” “satisfactory,” “acceptable,” and/or “feasible,” using subjective, objective, or subjective and objective criteria and unfailingly interpreting positive UEI ratings despite a wide range of participant reports and usage data. This suggests that either studies do not know the threshold constituting positive UEI, or that the threshold is quite low. The inconsistency of these methodologies makes it difficult to understand what user engagement is and how to best design for it. Furthermore, it calls the practices used to evaluate mental health apps into question.

These results indicate the lack of consensus about what constitutes “usability,” “satisfaction,” “acceptability,” and “feasibility” for mental health apps. This makes it difficult to compare results across studies, hinders understanding of what makes app engaging for different users and limits their real world uptake. The online supplements demonstrate the ambiguity that currently characterizes the distinctions between different types of UEI, which reduces the usefulness of these descriptors. There is thus a clear and urgent need to formulate standards for reporting and sharing user engagement indicators so that accurate assessments and informed decisions regarding app research, funding and clinical use can be made.

It is concerning that 16 of the 40 (40%) studies concluded that their app had positive UEI without considering objective data (Figure 1). Qualitative data is unquestionably valuable for creating a fuller, more nuanced picture of participants, since their characteristics—such as language, disorder and age—largely inform their ability to use an app and their unique experience of an app. But there is also a need for objective measurements that can be

reproduced to validate initial results and create a baseline for generalizing results of any single study. Consequently a combination of both subjective and objective criteria may be most useful for offering insight into user engagement.

All studies concluded that their apps had positive UEI on the bases of vastly different subjective and objective criteria, as shown in the online supplements. While of the thresholds for positive UEI must depend on the specific purpose of an app (e.g. one paper argues that a single usage of a suicide prevention app at a critical moment could be adequate (36)), predetermined thresholds for interpreting UEI are urgently required for any meaningful conclusions to be drawn. With every single app claiming positive UEI, understanding the current challenges surrounding usability, engagement and usage is difficult and progress in the field is hindered.

This study has several limitations. We only reviewed papers from academic sources that focus on depression, bipolar disorder, schizophrenia and anxiety, after our search retrieved 925 papers. This restricts our discussion to how the academic community views engagement, as opposed to other industries, and limits the types of mental health apps we took into account. In addition, we assumed that it would be possible and useful for at least some dimensions of engageability to be measured and reported consistently across mental health apps. Of course, apps that are developed for different purposes require their own specific criteria for determining whether they are engaging or not. However, if every single paper claims their app has positive user engagement indicators and no papers use the same evaluation methods—as found in this review—it makes it difficult to understand and improve the real world low uptake of these apps. While publication bias may explain some of these results, the need for some reporting standards is still clear. With over 10,000 mental health apps in the commercial marketplaces and few of them having ever been studied or published on (5), there is an immense number of black boxes when it comes to user engagement in mobile mental health apps. Looking at different conditions beyond those targeted in this review may have also yielded different results.

## CONCLUSION

The experience of mental illness is personal and the technology literacy of individuals is variable, meaning that there will never be a single scale or measurement that perfectly captures all engagement indicators for all people. But the future of the field of mobile mental health apps depends on user engagement; and the lack of clear definitions and standards around UEI is harmful not only to the field, whose progress is impeded, but also to patients, who may not know which app to trust. This review confirms the necessity of generating more clarity around UEI, which can both promote app usage and enable researchers to learn from each other's work and design better mental health apps.

This challenge is compounded by the need to design specifically for the needs of people with mental illness. On the topic of website design, one study reports, “Commonly prescribed design models and guidelines produce websites that are poorly suited and confusing to persons with serious mental illnesses” (59). Given that smartphone apps are often more complex and interactive than websites, it is reasonable to assume that truly

usable apps for mental illnesses may look different than apps designed for the general population. The inconsistencies illustrated in this study raise the possibility that no engagement indicators were designed to take into account the potentially unique cognitive, neurological, or motor needs arising from mental illnesses. For example, schizophrenia can lead to changes in cognition, depression can impact reward learning and anxiety can affect working memory. Furthermore, it is important to consider how the intersectional identities of people with mental illness also shape their engagement with mental health apps. Applying the lessons from co-designing technology with those with mental illnesses, like schizophrenia (60), to the methods for evaluating UEI could serve as a useful starting point. Other fields have found solutions, and the popularity of the engineering field-derived System Usability Scale (employed by several studies in this review) indicates the potential of simple but well-validated metrics. A representative body of patients, clinicians, designers and technology-makers proposing collaborative measures would be a welcome first step.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## REFERENCES

1. Bhugra D, Tasman A, Pathare S, et al.: The WPA-lancet psychiatry commission on the future of psychiatry. *The Lancet Psychiatry*. 2017 10 31;4(10):775–818. [PubMed: 28946952]
2. Firth J, Torous J, Nicholas J, et al.: The efficacy of smartphone based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry*. 2017 10;16(3):287–98. [PubMed: 28941113]
3. Naslund JA, Aschbrenner KA, Araya R, et al.: Digital technology for treating and preventing mental disorders in low-income and middle-income countries: a narrative review of the literature. *The Lancet Psychiatry*. 2017 6 1;4(6):486–500. [PubMed: 28433615]
4. Chan S, Godwin H, Gonzalez A, et al.: Review of use and integration of mobile apps into psychiatric treatments. *Current psychiatry reports*. 2017 12 1;19(12):96. [PubMed: 29082425]
5. Torous J, Roberts LW: Needed innovation in digital health and smartphone applications for mental health: transparency and trust. *JAMA psychiatry*. 2017 5 1;74(5):437–8. [PubMed: 28384700]
6. Torous J, Nicholas J, Larsen ME, et al.: Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. *Evidence-based mental health*. 2018 8 1;21(3):116–9. [PubMed: 29871870]
7. Mackie C, Dunn N, MacLean S, et al.: A qualitative study of a blended therapy using problem solving therapy with a customised smartphone app in men who present to hospital with intentional self-harm. *Evidence-based mental health*. 2017 10 13;ebmental-2017.
8. Levin M, Haeger J, Pierce B, et al.: Evaluating an Adjunctive Mobile App to Enhance Psychological Flexibility in Acceptance and Commitment Therapy. *Behav Modif*. 2017;41(6):846–867. [PubMed: 28689449]
9. Fortuna K, DiMilia P, Lohman M, et al.: Feasibility, Acceptability, and Preliminary Effectiveness of a Peer-Delivered and Technology Supported Self-Management Intervention for Older Adults with Serious Mental Illness. *Psychiatric Quarterly*. 2017;89(2):293–305.
10. Korsbek L, Tønder E: Momentum: A smartphone application to support shared decision making for people using mental health services. *Psychiatr Rehabil J*. 2016;39(2):167–172. [PubMed: 27030907]
11. Schlosser D, Campellone T, Truong B, et al.: The feasibility, acceptability, and outcomes of PRIME-D: A novel mobile intervention treatment for depression. *Depress Anxiety*. 2017;34(6):546–554. [PubMed: 28419621]



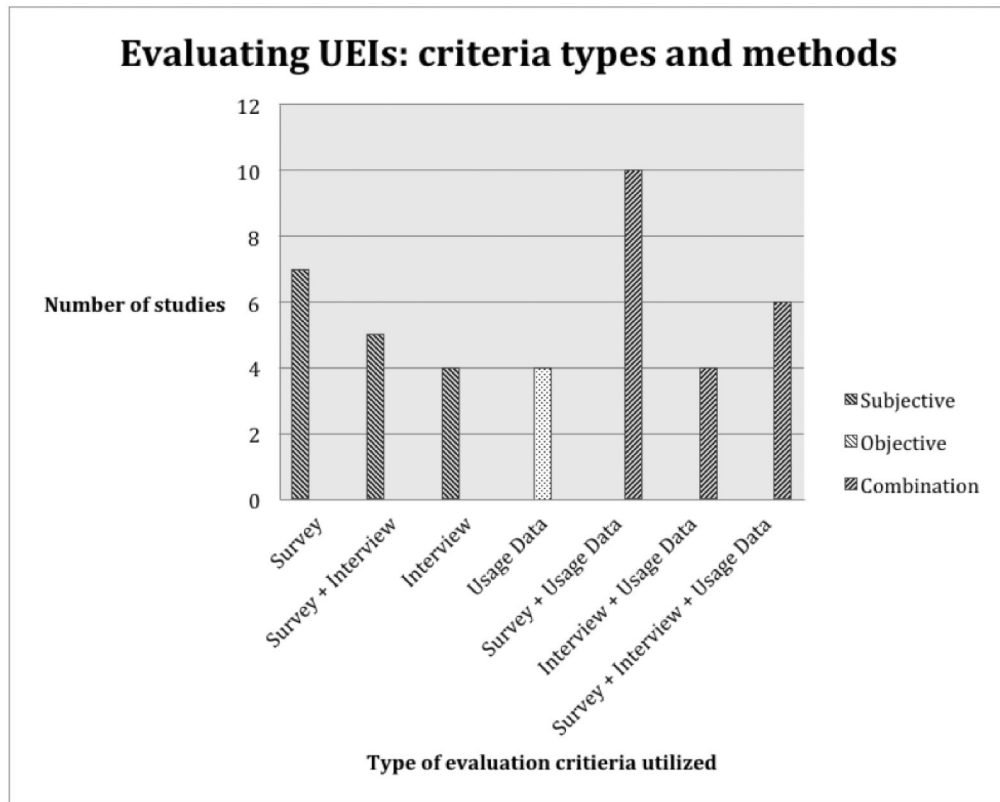
12. Dahne J, Kustanowitz J, Lejuez C: Development and Preliminary Feasibility Study of a Brief Behavioral Activation Mobile Application (Behavioral Appivation) to Be Used in Conjunction With Ongoing Therapy. *Cogn Behav Pract*. 2018;25(1):44–56. [PubMed: 29725198]
13. Bauer A, Iles-Shih M, Ghomi R, et al.: Acceptability of mHealth augmentation of Collaborative Care: A mixed methods pilot study. *Gen Hosp Psychiatry*. 2018;51:22–29. [PubMed: 29272712]
14. Niendam T, Tully L, Iosif A, et al.: Enhancing early psychosis treatment using smartphone technology: A longitudinal feasibility and validity study. *J Psychiatr Res*. 2018;96:239–246. [PubMed: 29126059]
15. Possemato K, Kuhn E, Johnson E, et al.: Development and refinement of a clinician intervention to facilitate primary care patient use of the PTSD Coach app. *Transl Behav Med*. 2016;7(1):116–126.
16. Wenze S, Arney M, Weinstock L, et al.: An Open Trial of a Smartphone-assisted, Adjunctive Intervention to Improve Treatment Adherence in Bipolar Disorder. *J Psychiatr Pract*. 2016;22(6):492–504. [PubMed: 27824786]
17. Mohr D, Tomasino K, Lattie E, et al.: IntelliCare: An Eclectic, Skills-Based App Suite for the Treatment of Depression and Anxiety. *J Med Internet Res*. 2017;19(1):e10. [PubMed: 28057609]
18. Rohatagi S, Profit D, Hatch A, et al.: Optimization of a Digital Medicine System in Psychiatry. *J Clin Psychiatry*. 2016:e1101–e1107. [PubMed: 27487251]
19. Dang M, Mielke C, Diehl A, et al.: Accompanying Depression with FINE-A Smartphone-Based Approach. *Stud Health Technol Inform*. 2016;228:195–9. [PubMed: 27577370]
20. Schwartz S, Schultz S, Reider A, et al.: Daily mood monitoring of symptoms using smartphones in bipolar disorder: A pilot study assessing the feasibility of ecological momentary assessment. *J Affect Disord*. 2016;191:88–93. [PubMed: 26655117]
21. Hung S, Li M, Chen Y, et al.: Smartphone-based ecological momentary assessment for Chinese patients with depression: An exploratory study in Taiwan. *Asian J Psychiatr*. 2016;23:131–136. [PubMed: 27969071]
22. Hidalgo-Mazzei D, Mateu A, Reinares M, et al.: Psychoeducation in bipolar disorder with a SIMPLe smartphone application: Feasibility, acceptability and satisfaction. *J Affect Disord*. 2016;200:58–66. [PubMed: 27128358]
23. Ben-Zeev D, Wang R, Abdullah S, et al.: Mobile Behavioral Sensing for Outpatients and Inpatients With Schizophrenia. *Psychiatric Services*. 2016;67(5):558–561. [PubMed: 26695497]
24. Macias C, Panch T, Hicks Y, et al.: Using Smartphone Apps to Promote Psychiatric and Physical Well-Being. *Psychiatric Quarterly*. 2015;86(4):505–519. [PubMed: 25636496]
25. Ben-Zeev D, Brenner C, Begale M, et al.: Feasibility, Acceptability, and Preliminary Efficacy of a Smartphone Intervention for Schizophrenia. *Schizophr Bull*. 2014;40(6):1244–1253. [PubMed: 24609454]
26. Pramana G, Parmanto B, Kendall P, et al.: The SmartCAT: An m-Health Platform for Ecological Momentary Intervention in Child Anxiety Treatment. *Telemedicine and e-Health*. 2014;20(5):419–427. [PubMed: 24579913]
27. Watts S, Mackenzie A, Thomas C, et al.: CBT for depression: a pilot RCT comparing mobile phone vs. computer. *BMC Psychiatry*. 2013;13(1).
28. Palmier-Claus J, Ainsworth J, Machin M, et al.: The feasibility and validity of ambulatory self-report of psychotic symptoms using a smartphone software application. *BMC Psychiatry*. 2012;12(1).
29. Burns M, Begale M, Duffecy J, et al.: Harnessing Context Sensing to Develop a Mobile Intervention for Depression. *J Med Internet Res*. 2011;13(3):e55. [PubMed: 21840837]
30. Cernvall M, Sveen J, Bergh Johannesson K, et al.: A pilot study of user satisfaction and perceived helpfulness of the Swedish version of the mobile app PTSD Coach. *Eur J Psychotraumatol*. 2018;9(sup1):1472990. [PubMed: 29805783]
31. Reger G, Skopp N, Edwards-Stewart A, et al.: Comparison of Prolonged Exposure (PE) Coach to Treatment as Usual: A Case Series With Two Active Duty Soldiers. *Military Psychology*. 2015;27(5):287–296.
32. Boisseau C, Schwartzman C, Lawton J, et al.: App-guided exposure and response prevention for obsessive compulsive disorder: an open pilot trial. *Cogn Behav Ther*. 2017;46(6):447–458. [PubMed: 28565937]

33. Ben-Zeev D, Brian R, Aschbrenner K, et al.: Video-based mobile health interventions for people with schizophrenia: Bringing the “pocket therapist” to life. *Psychiatr Rehabil J*. 2018;41(1):39–45. [PubMed: 27295133]
34. Ramsey A, Wetherell J, Depp C, et al.: Feasibility and Acceptability of Smartphone Assessment in Older Adults with Cognitive and Emotional Difficulties. *J Technol Hum Serv*. 2016;34(2):209–223. [PubMed: 27683018]
35. Hicks T, Thomas S, Wilson S, et al.: A Preliminary Investigation of a Relapse Prevention Mobile Application to Maintain Smoking Abstinence Among Individuals With Posttraumatic Stress Disorder. *J Dual Diagn*. 2016;13(1):15–20. [PubMed: 27918881]
36. Melvin G, Gresham D, Beaton S, et al.: Evaluating the Feasibility and Effectiveness of an Australian Safety Planning Smartphone Application: A Pilot Study Within a Tertiary Mental Health Service. *Suicide and Life-Threatening Behavior*. 2018.
37. Ben-Zeev D, Brian R, Jonathan G, et al.: Mobile Health (mHealth) Versus Clinic-Based Group Intervention for People With Serious Mental Illness: A Randomized Controlled Trial. *Psychiatric Services*. 2018;69(9):978–985. [PubMed: 29793397]
38. Kreyenbuhl J, Record E, Himelhoch S, et al.: Development and Feasibility Testing of a Smartphone Intervention to Improve Adherence to Antipsychotic Medications. *Clin Schizophr Relat Psychoses*. 2016;(aop).
39. Bucci S, Barrowclough C, Ainsworth J, et al.: Actissist: Proof-of-Concept Trial of a Theory-Driven Digital Intervention for Psychosis. *Schizophr Bull*. 2018;44(5):1070–1080. [PubMed: 29566206]
40. Corden M, Koucky E, Brenner C, et al.: MedLink: A mobile intervention to improve medication adherence and processes of care for treatment of depression in general medicine. *Digit Health*. 2016;2:205520761666306.
41. Bauer A, Hodsdon S, Bechtel J, et al.: Applying the Principles for Digital Development: Case Study of a Smartphone App to Support Collaborative Care for Rural Patients With Posttraumatic Stress Disorder or Bipolar Disorder. *J Med Internet Res*. 2018;20(6):e10048. [PubMed: 29875085]
42. Kumar D, Tully L, Iosif A et al.: A Mobile Health Platform for Clinical Monitoring in Early Psychosis: Implementation in Community-Based Outpatient Early Psychosis Care. *JMIR Ment Health*. 2018;5(1):e15. [PubMed: 29487044]
43. Baumel A, Tinkelman A, Mathur N, et al.: Digital Peer-Support Platform (7Cups) as an Adjunct Treatment for Women With Postpartum Depression: Feasibility, Acceptability, and Preliminary Efficacy Study. *JMIR Mhealth Uhealth*. 2018;6(2):e38. [PubMed: 29439944]
44. Price M, van Stolk-Cooke K, Ward H et al.: Tracking Post-trauma Psychopathology Using Mobile Applications: a Usability Study. *J Technol Behav Sci*. 2017;2(1):41–48. [PubMed: 29109968]
45. Schlosser D, Campellone T, Kim D, et al.: Feasibility of PRIME: A Cognitive Neuroscience-Informed Mobile App Intervention to Enhance Motivated Behavior and Improve Quality of Life in Recent Onset Schizophrenia. *JMIR Res Protoc*. 2016;5(2):e77. [PubMed: 27125771]
46. Birney A, Gunn R, Russell J: MoodHacker Mobile Web App With Email for Adults to Self-Manage Mild-to-Moderate Depression: Randomized Controlled Trial. *JMIR Mhealth Uhealth*. 2016;4(1):e8. [PubMed: 26813737]
47. Saunders K, Bilderbeck A, Panchal P: Experiences of remote mood and activity monitoring in bipolar disorder: A qualitative study. *European Psychiatry*. 2017;41:115–121. [PubMed: 28135594]
48. Brooke J: SUS-A quick and dirty usability scale. *Usability Evaluation in Industry*. 1996:189–194.
49. Larsen DL, Attkisson CC, Hargreaves WA, et al.: Assessment of client/patient satisfaction: development of a general scale. *Eval Program Plan*. 1979; 2:197–207.
50. Devilly GJ, Borkovec TD: Psychometric properties of the credibility/expectancy questionnaire. *J Behav Ther Exp Psychiatry*. 2000; 32:73–86.
51. Lund AM: Measuring usability with the USE questionnaire. *Usability and User Experience*. 2001;8:8.
52. Kuhn E, Greene C, Hoffman J, et al.: Preliminary evaluation of PTSD Coach, a smartphone app for post-traumatic stress symptoms. *Military Medicine* 2014, 179(1), 12–18. [PubMed: 24402979]
53. Lewis JR: Psychometric evaluation of the poststudy system usability questionnaire: the PSSUQ. In: *Proceedings of the Human Factor Society Annual Meeting*. 1992;36:1259.

54. Davis FD: User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. *Int. J. Man-Mach. Stud* 1993; 38(no. 3):475–487.
55. Venkatech V, Davis FD: A theoretical extension of the technology acceptance model: four longitudinal field studies. *Manage Sci.* 2000;46:186–204.
56. Shirk SR, Gudmundsen G, Kaplinski HC, et al.: Alliance and outcome in cognitive-behavioral therapy for adolescent depression. *J Clin Child Adolesc Psychol* 2008;37:631–639. [PubMed: 18645753]
57. Nguyen TD, Attkisson CC, & Stegner BL (1983): Assessment of patient satisfaction: Development and refinement of a service evaluation questionnaire. *Evaluation and Program Planning*, 6, 299–313. [PubMed: 10267258]
58. Lewis J: IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *Int J Hum Comput Interact.* 1995;7(1):57–78.
59. Rotondi A, Sinkule J, Haas G et al.: Designing websites for persons with cognitive deficits: Design and usability of a psychoeducational intervention for persons with severe mental illness. *Psychol Serv.* 2007;4(3):202–224. [PubMed: 26321884]
60. Lindberg S, Jormfeldt H, Bergquist M: Unlocking design potential: Design with people diagnosed with schizophrenia. *Informatics for Health and Social Care.* 2017:1–17.

**Highlights:**

- In response to the low uptake and sustained use of mobile mental health apps, this study conducted a systematic review of papers presenting results of mental health apps for depression, bipolar disorder, schizophrenia and anxiety.
- 90% of papers employed at least some subjective criteria (371 total) and 60% of papers employed at least some objective criteria (71 total).
- Every single paper concluded that their app had positive user engagement indicators, yet no papers used the same combination criteria or thresholds to evaluate their apps.



**Figure 1.** Studies evaluated UEI based on different subjective and/or objective criteria, as well as different methods.

**Table 1:**

Search terms organized to PICO's framework

Category	Search words used ('OR' terms)
<i>Population</i>	depression; depressive; mental illness; mental health; mood disorder; affective disorder; anxiety; phobia; bipolar; psychosis; schizophren*
<i>Intervention</i>	Smartphone*; smart phone*; mhealth; mobile phone*; iphone*; android; mhealth; mobile app*; phone app*.
<i>Comparator</i>	[any]
<i>Outcomes</i>	usability; user interface; ui; feasib*; pilot; engag*; acceptability;

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2:**

Of the minority of studies that rely on pre-existing assessment tools, there was no consensus on the best tool for evaluating UEI.

UEI assessment tools utilized	N	Study
System Usability Scale (48)	4	(8, 25, 41, 46)
Client Satisfaction Questionnaire (49)	3	(15, 16, 32)
Credibility and Expectancy Scale (50)	3	(16, 27, 32)
Usefulness, Satisfaction, and Ease Questionnaire (51)	2	(25, 40)
Adaptation of Another Study's Assessment Tool (25, 52)	2	(23, 30)
Post Study System Usability Questionnaire (53)	1	(25)
Technology Acceptance Model (54)	1	(44)
Technology Assessment Model Measurement Scale (55)	1	(25)
Therapeutic Alliance Scales for Children—Revised (56)	1	(26)
Client Evaluation of Services Questionnaire (57)	1	(26)
Computer System Usability Questionnaire (58)	1	(26)