International Journal of Technology Assessment in Health Care

cambridge.org/thc

A Narrative Review of Methods for Applying User Experience in the Design and Assessment of Mental Health Smartphone Interventions

Christopher Lemon[1,2], Kit Huckvale[3], Kenneth Carswell[4] and John Torous[5]

[1]St Vincent's Hospital, 390 Victoria Street, Sydney, Australia; [2]Faculty of Medicine, University of New South Wales, Sydney, NSW, Australia; [3]Black Dog Institute, UNSW Sydney, Sydney, Australia; [4]Department of Mental Health and Substance Use, World Health Organization, Geneva, Switzerland and [5]Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, USA

Abstract

Objectives. User experience (UX) plays a key role in uptake and usage of mental health smartphone interventions, yet remains underinvestigated. This review aimed to characterize and compare UX evaluation approaches that have been applied in this specific context, and to identify implications for research and practice.

Methods. A narrative review was conducted of UX-themed studies published in PubMed, PsycINFO, and Scopus up to February 2019. Eligible studies reported on data reflecting users' interactions with a smartphone intervention for any mental health condition. Studies were categorized into "situated" versus "construct-based" methods according to whether or not an established UX construct was used to acquire and analyze data.

Results. Situated approaches used bespoke UX metrics, including quantitative measures of usage and performance, as well as grounded interview data. Construct-based approaches such as assessments of usability and acceptability were based on conceptual frameworks, with methodologically stronger versions featuring construct definitions, validated measurement tools, and an ability to compare data across studies. Constructs and measures were sometimes combined to form bespoke construct-based approaches.

Conclusions. Both situated and construct-based UX data may provide benefits during design and implementation of a mental health smartphone intervention by helping to clarify the needs of users and the impact of new features. Notable however was the omission of UX methods, such as split testing. Future research should consider these unaddressed methods, aim to improve the rigor of UX assessment, integrate their use alongside clinical outcomes, and explore UX assessment of more complex, adaptive interventions.

Assessment

Cite this article: Lemon C, Huckvale K, Carswell K, Torous J (2020). A Narrative Review of Methods for Applying User Experience in the Design and Assessment of Mental Health Smartphone Interventions. *International Journal of Technology Assessment in Health Care* **36**, 64–70. https://doi.org/10.1017/S0266462319003507

Received: 10 June 2019
Revised: 31 October 2019
Accepted: 10 December 2019
First published online: 24 January 2020

Key words:
User experience; Mental health; eHealth; Mobile health; Apps

Author for correspondence:
Christopher Lemon,
E-mail: christopheralemon@gmail.com

Smartphone interventions offer unprecedented opportunities to support access to various forms of health care almost anywhere in the world, especially where the rate of device ownership is high but the availability (1) or uptake (2) of face-to-face services is low. However, many users of smartphone health interventions delete them within weeks of download (3). Poor user experience (UX) is a contributing factor, and a consequent threat to ubiquitous smartphone-supported care. Indeed, negative UX leads to poor adoption and variable use of technology-based interventions (4–6). Analysis of UX can provide useful data for enhancing the design and implementation of smartphone interventions, as well as improving target outcomes through increased adherence, fidelity, and reach. Addressing poor UX can increase user satisfaction (7), while tailoring design to users' needs can improve treatment adherence and patient empowerment (8).

Mental disorders are a leading cause of disability, yet uptake of care services remains low. Interventions aiming to address barriers to improving mental health-related morbidity, such as stigma, poor mental health literacy (2), and fragmented health services (9), have a UX dimension. For example, the selection of language and imagery, as well as direct use of attitude-changing interventions (such as "myth-busting" (10)) are not only relevant to users' experience of stigma, but must also be designed through user-focused methods to be effective (11).

Smartphone interventions have the potential to assist in overcoming these barriers by providing confidential care, tailored toward users' literacy, needs, and availability (12). Design thinking (DT) (13) can help realize these advantages, yet remains underutilized in mental health research (14). DT involves building solutions based on empathy toward users' nuanced needs, multidisciplinary ideation, and experimentation using prototyping and iterative development (14). UX testing is based around the principles of DT, and aims to collect and utilize data reflecting user values and behaviors, and direct design to satisfy users' purposes (15). UX research requires balancing the demands of the scientific method with the subjectivity of user feedback.

CAMBRIDGE UNIVERSITY PRESS

While this can be challenging, it allows the development of bespoke products which more effectively meet the needs of users (16).

UX testing remains underinvestigated in the digital mental health literature. Some studies have recognized the importance of optimizing the UX of digital interventions, but few provide guidance on how best to do this. For example, in a review paper, Bakker et al. (17) recommended using reminder functions, simple interface designs, and providing links to crisis support services to improve user engagement. However, methods available for collecting and using UX data in design to realize these recommendations were not explored. Torous et al. (18) reviewed eleven studies on smartphone apps for schizophrenia, and found some form of UX testing in most, but this commonly involved obtaining only general, unstructured feedback from users throughout the design process. Additionally, Feather et al. (19) conducted a systematic review, and examined 21 studies on web-based interventions for mental health conditions, such as depression. They found UX data was most commonly collected to improve the understandings of barriers to use. Some studies used qualitative methods, such as interviewing users during field trials of an intervention. Others applied questionnaires to examine UX through constructs, including user-satisfaction and acceptability. However, most measures were not assessed for rigor, and little was provided on the applications of data. Nicholas et al. (20) investigated how users formulate UX assessments, and examined trends in public user reviews of smartphone apps for bipolar disorder. Users focused on functionality, perceived effects on relationships with clinicians, and how easily apps could be integrated into existing care plans. However, potential applications for these data were not examined in detail.

Limited understanding of user characteristics, user-centric design practices, and field utility assessments remain barriers to realizing the full potential of smartphones to improve mental health outcomes (21). UX-based design methods, the reporting of UX data, and the development of standards for UX testing have been recognized as key issues (22). However, a synthesis of potential approaches to help guide researchers, clinicians, and patients is lacking. The objective of this review was to respond by providing a critical overview of UX methods that have been used by researchers and practitioners seeking to improve the design and implementation of mental health smartphone interventions and to use these findings to identify opportunities for future UX evaluations directed toward this goal.

Methods

Within this paper, "user experience" (UX) refers to the dynamic combination of subjective and contextual factors that shape users' engagement with products or systems (23;24). Approaches to examining UX vary, ranging from quantitative system performance data, such as time taken to complete tasks, to user opinions, and construct-based measures, including user-satisfaction questionnaires (13). A principled approach to the measurement of UX recognizes it as a multidimensional phenomenon that can be characterized using multiple, overlapping constructs from a wide range of theoretical backgrounds. In this context, a construct is a coherent, well-operationalized account of some aspect of subjective experience. Constructs relevant to UX include satisfaction, acceptability, feasibility, utility, likeability, learnability, credibility, and usability (21;25;26).

Reflecting the rapid pace of research on mental health smartphone interventions, the heterogeneity of study types, and a still-emerging focus on UX testing in relation to smartphone apps, we conducted a narrative review based on the research question: "In the research literature, what methods have been applied in practice to integrate user experience data into the design and implementation of mental health smartphone interventions?" We adopted a sample-based search strategy. We did not critically review every paper published in this area, but instead scanned available literature to identify key examples of major types of UX approaches within mental health. We focused on studies investigating the role of UX testing in mental health research, clinical work, and other settings involving mental health smartphone interventions.

PubMed, PsycINFO, and Scopus were searched from their inception to February 2019 with the following search terms: cell phone, smartphone, telemedicine, psychiatry, mental disorders, user experience, user perceptions, user testing, analytics, evaluation, and usability. We selected representative papers in English for each category from the last 6 years if they examined a mental health smartphone intervention, included an assessment of data reflecting the users' perception of the intervention, and drew conclusions specifically about UX. As UX testing is still emerging in the literature on mental health smartphone interventions, the quality of research was considered and discussed, but not used to exclude studies.

Currently accessible approaches are diverse, both in terms of the processes involved and the underlying ontological frameworks. Consequently, there is no universally agreed organizing principle. We therefore divided findings from our search into (a) studies using a situated approach to define bespoke UX measurement and evaluation and (b) studies using methods based on normative UX constructs, such as usability. We discuss how these approaches can be applied in design and implementation processes, as well as guide clinical use of mental health smartphone interventions. Table 1 summarizes our search findings.

Results

Situated Approaches

Situated approaches were characterized by the use of investigator-specified UX methods or metrics closely tied to the specific contexts where evaluation occurred. Quantitative methods within this group included bespoke measures of product adherence, for example, task completion rates that were used for subgroup analyses of study outcomes. Methods also included qualitative techniques, such as collating unstructured user opinions about functional and emotional aspects of design and implementation of specific products. Grounded approaches to qualitative analysis that did not draw on formal UX constructs in their subsequent analysis also fell into this group.

Simple bespoke UX-focused feedback offers a pragmatic and potentially efficient way to summarize perceptions concerning the UX of an existing app design. Mackintosh et al. (27) examined the UX of a mobile app for the treatment of emotional dysregulation among veterans. The app had functions for practicing emotional regulation skills, monitoring symptoms as well as recording physiological data, and was used in-between face-to-face anger management therapy sessions. A bespoke technology feedback questionnaire sought to assess "ease of use," "frustration," and "helpfulness" using a six-point scale with data summarized as score averages. Treatment engagement was also measured according to the dropout rate (attendance of less than twelve therapy

https://doi.org/10.1017/S0266462319003507 Published online by Cambridge University Press

Table 1. Summary of approaches to user experience testing of mental health smartphone interventions and potential applications in design and implementation

	Examples	Pros	Cons	Potential applications
Situated UX assessment	• Rates of use over time • Download and deletion rates • Semistructured interviews about technology issues with potential users	• Can use easily accessed background usage data, e.g., time spent using particular functions or on key tasks • Allows for user-centric design by allowing users to determine the most important issues in user experience testing • Easy to explain data to others who don't have expertise in user experience testing	• Difficult to meaningfully compare data from multiple interventions that serve the same purpose • Difficult to meaningfully examine changes in user experience over time or as interventions are updated	• Early stages of design of an intervention to determine the needs of users, e.g., requirement analysis phase • Screen for new issues arising in the user experience at various points in design and implementation
Approaches based on constructs	• User satisfaction • Usability • Acceptability	• Data are specific and can be collected using robust measures and interpreted with reference to an established definition • Can meaningfully compare user experience of multiple interventions for the same purpose • Some constructs allow measurement of change in user experience over time or as interventions are updated	• Some constructs lack definitions and robust measures, giving poor quality data • Requires some degree of user experience expertise to understand data, or other tools for interpretation can be needed to translate data to others • Studies often use several constructs interchangeably	• Measuring how user experience changes as interventions are modified • Tracking changes in user experience over time • Potential for testing the relationship between user experience and other measures of effectiveness, such as clinical outcomes

sessions), session attendance, and treatment completion. According to questionnaire data, users' perceptions of the app's ease and helpfulness increased after 3 months of consistent use, while levels of frustration decreased. There was no difference in treatment engagement amongst those using the app alongside face-to-face therapy. On the basis of these overall positive scores, the app was considered helpful, easy to use, and efficient. Summary statistics such as these have particular value as repeated measures and may provide a basis for iterative UX improvement. However, using single-dimension measures offers little diagnostic potential. Had users indicated poor subjective UX along the dimensions being measured, this approach offers limited insight as to *why*.

Potentially addressing this limitation, grounded qualitative methods allow intervention designers to understand the detail of UX issues considered important to users. Orlowski et al. (28) used thematic analysis of interviews about smartphone-based care with youth mental health workers. Participants were asked about current use of technology at work, barriers to increased use, and how technology could impact one's professional role. Responses revealed youth mental health workers felt technology-based care was not "standard" practice and could challenge their ability to make interpersonal connections with patients. Concerns were raised about the amount of specific training needed to ensure staff could access maximum benefits from such technology. These results highlight how UX insights reveal areas for improving the design beyond the technical intervention, such as ensuring that technology-led change is integrated into clinical workflows. Practical design strategies that respond to these issues might include training and workflow process engineering, in addition to changing the app design itself.

A situated alternative to user feedback is to use observable data, such as automatically collected user analytics. Attwood et al. (29) evaluated the UX of an app designed to help monitor and reduce alcohol consumption. Download and deletion rates were used as indicators of usage patterns. About 42% of users deleted the app within 1 week. By the end of a 12-week program, however, only 5% of the initial number of users remained. The use of particular app functions was also examined. For instance, female users, those aged 35–44, users exhibiting "high-risk" drinking behaviors (more than twice the upper limit of recommended intake on four or more occasions within 1 week), and those with specific goals of reducing alcohol intake tended to use goal-setting functions more frequently than others. An advantage of quantitative UX data is that it can be explicitly linked to outcomes through modeling or subgroup analysis. Multiple linear regression analyses on the patterns of app usage associated with drinking behaviors identified gender, age, season of download, and baseline drinking as explaining more than a quarter of the variance in alcohol consumption at 1 month of app use.

Approaches Based on UX Constructs

In contrast to situated measures, constructs can provide a normative conceptual framework for investigating and interpreting the UX. Multiple constructs can be applied, each examining a different UX dimension. Universal construct definitions and validated methods add precision to UX assessment and allow comparison of apps built for the same purpose. Usability is a popular construct for measuring the UX because of its commonly cited universal definition and specific assessment methodologies. Usability is defined by the International Organisation for Standardisation (ISO) as "the extent to which a system, product or service can be used by specified users to achieve a specified goal with effectiveness, efficiency and satisfaction in a specified

https://doi.org/10.1017/S0266462319003507 Published online by Cambridge University Press

context of use" (30). Systems with high usability are associated with increased occupational value, while low usability is linked to user frustration and workflow disruption (31;32). For other constructs, definitions are not always agreed upon, and evaluation methods not always robust.

Similar to situated approaches, construct-based approaches to UX evaluation can be based on instrument-based measurement, qualitative inquiry, or both. In mixed approaches, the quantitative method provides normative data to benchmark and compare UX along the selected dimension, while the qualitative method is used to fill in contextual details to deepen understanding of UX problems and inform product refinement. For instance, Vilardaga et al. (33) had patients with a mental illness evaluate the usability of a popular smoking cessation app. Usability was measured through the System Usability Scale (SUS), a 10-item questionnaire that gives a total score out of 100 to represent overall usability. The SUS has been validated, found to be reliable ($\alpha = .91$), and there are agreed schedules for categorizing scores (34). A separate structured interview explored specific issues drawn from the underlying construct of usability. These related to navigation, motivation, utility of features, barriers to use, potential situations where the app may be pleasant or engaging as well as opportunities for giving feedback about design and implementation. The app scored 65.6/100 on the SUS. Interview data mapped to this moderate–poor rating included content being difficult to understand for some users, and differences in perceptions of functions, as some found smoking cessation reminders helpful, while others found them annoying. However, there was no subsequent refinement of the app, nor measurement of the UX.

One of the potential benefits of construct-based approaches is comparison of normative findings between competing intervention variants or across studies. Dubad et al. (35) conducted a systematic review assessing the psychometric properties, usability, and clinical effects of mood monitoring smartphone apps for young people aged 10–24 years. They used a previous, broader ISO definition of usability as "the capability of the software product to be understood, learned, used and attractive to the user, when used under specified conditions" (36). However, studies were compared mainly according to participation rates of app use, and available data reflecting users' perceptions of apps, such as whether they found them easy to use. Based on these data, Dubad et al. (35) concluded that users have overall positive perceptions of mental health apps and are willing to use them in real life. Although the usability construct provided a framework for study comparisons, data from usability metrics themselves were not compared. This may reflect the still limited use of standardized tools across the literature, yet remains a major potential advantage of the usability construct and associated measures, such as the SUS (37).

Construct-based measures can also be combined to form bespoke UX measures. Ben-Zeev et al. (38) drew on several constructs and validated standard instruments to analyze the UX of a smartphone app for schizophrenia which collected background behavioral data, such as location and physical movement to map behavioral tendencies. A 26-item questionnaire was used to measure acceptability and usability. Items were drawn from construct-based questionnaires, including the SUS (39), Post Study System Usability Scale (40), Technology Assessment Model Measurement Scales (41), and Usefulness, Satisfaction and Ease questionnaire (42). Data indicated users were comfortable using the app (95%), understood its functions (70%), did not find it difficult to use (70%), were interested in the app having functions to provide feedback to them about their behavior (65%) as well as functions suggesting coping mechanisms in times of distress (65%). However, there was no measure of change in acceptability or usability over time, only that the app was considered highly acceptable and usable.

Not all UX constructs have established definitions or conceptual and operational frameworks, which can limit the interpretation and utility of UX data. For instance, Povey et al. (43) conducted a qualitative study evaluating the acceptability of a cognitive behavioral therapy app and suicide prevention app for Aboriginal and Torres Strait Australians in controlled settings. Povey et al. (43) reported acceptability was influenced by personal factors (e.g., aspects of users' illness), environmental factors (e.g., stigma toward mental health conditions), and app functionality (e.g., "ease of use"). However, acceptability was not defined, and the qualitative methodology used was not specific to the construct. An established definition is needed to benchmark the UX of apps, and could have helped determine if the app's UX met a standard where clinicians could recommend use. Construct-specific, validated methods of analysis within a defined conceptual framework could also have enabled practical design and implementation recommendations. For instance, if Povey et al. (43) demonstrated that environmental factors are stronger predictors of app acceptability than personal factors, more focus could be given to implementation rather than design in communities where mental illness is strongly stigmatized.

Discussion

Mental health smartphone interventions are increasingly used in clinical practice. One survey estimated that there were almost 250 apps for depression alone (44). Despite the potential value of UX data across the process of intervention design, evaluation, and commissioning, there is no gold standard for assessing the UX of smartphone-based mental health interventions. In this narrative review, our objective was to summarize the major types of approaches that have been used in practice. Our review demonstrates how a range of methods, which we partitioned according to underlying theoretical basis, have been used in the design and implementation of mental health smartphone interventions. Situated approaches provide quantitative data to help understand needs, behaviors and patterns of use for a specific user group, as well as bespoke measures of particular aspects of the UX. Constructs can assist by structuring analysis of the UX around agreed definitions and robust methodologies, enabling measurement and comparison of data using validated instruments, as well as analysis of multiple aspects of the UX in parallel.

The utility of any UX data depends on the stage of development of an intervention, the interests of the user and designer, as well as the strength of the method being used. The design and implementation process for smartphone interventions is rarely linear and requires repeated adaptation to the changing needs of patients and clinicians. Both situated and construct-based UX measures can be helpful at various stages. Our findings highlight how, in the early stages of design, situated methods can assist in determining the real needs of users, and help anticipate responses to potential barriers to use. During implementation, situated approaches also enable identification of UX issues that may arise when interventions are deployed into real-world use, as well as understanding how UX needs differ according to the type of user (45). As interventions become more established, construct-based approaches using established definitions and validated

https://doi.org/10.1017/S0266462319003507 Published online by Cambridge University Press

Table 2. Summary of the pros and cons of structured clinical tools using situated and construct-based UX assessment methods

	Examples	Pros	Cons	Potential applications
Structured UX assessment tools in clinical practice	• Anxiety and Depression Association of America Web site for mental health apps • Psyberguide Web site • American Psychiatric Association App Evaluation Model	• Easy to access for time-poor clinicians, administrators or patients • Require minimal expertise in user experience to apply in real-world practice • Help users understand the need to consider user experience alongside other measures of value, such as effects on clinical outcomes	• Some public recommendations are provided by clinicians who may or may not have experience with smartphone interventions or user experience testing • Combinations of situated and construct-based measures used, but often without being made explicit or with clear rationale	• Tools for making individual decisions, such as the APA App Evaluation Model helpful in clinical practice • Governments or mental health organizations seeking flexible methods of regulating or accrediting smartphone interventions • Patients wanting to make informed choices about smartphone interventions, but have not had any face-to-face care

measurement techniques can help track how the UX is changing with time or the impact of new iterations of an intervention. These findings also emphasize how dogmatic adherence to theoretical constructs may miss identification of actionable UX issues using grounded methods, as well as how analysis driven solely by situated methods may miss deeper insights gained through broader theoretical frameworks.

Recent work (46) reveals several UX methods being used in eHealth, such as eye-tracking (47) and card sorting (48), which were not reported in this review. Notably, experimental paradigms such as split testing (49) were lacking. Additionally, there was an absence of organized use of higher-level framework methodologies, including DT, which are being used to enhance the design of smartphone interventions in other areas, such as heart failure (50). Importantly, however, this study was not an exhaustive systematic review, and aimed to identify and analyze UX methods currently in use in mental health. Additionally, studies may not always report on preliminary UX testing of interventions. Hence, not all UX methods may have been detected.

Our review also highlights how, without clear definitions and validated measurement, construct-based UX measurement can be significantly limited. For example, Povey et al. (43) did not specifically define the constructs used, making it difficult to draw useful conclusions about the UX or understand how these data could be applied to improve the apps reviewed. From the point of view of comparison, this means that, while both proclaimed to assess the acceptability of a smartphone app, they may have been assessing quite different aspects of the UX. Further, where several constructs are used at once, each must be defined and measured separately if clear conclusions are to be drawn. This can allow for the development of bespoke combinations of constructs and validated measures (38). However, subtle differences in meaning between constructs can be lost when constructs are bundled without rationale or validated measurement. For instance, some studies use satisfaction as a measure of acceptability (51), while other literature recognizes acceptability as a distinct measure of the extent of behavior change resulting from an intervention, and can differ from levels of satisfaction toward the same intervention (52). Future studies should investigate the development of a battery of constructs with established definitions and validated tools for measuring the UX of smartphone interventions to increase their utility across the literature and in practice, particularly in mental health.

The findings of our review can help in navigating emerging evidence-based tools for evaluating the UX of smartphone interventions in clinical practice. Commonly used tools, such as star rating systems and user comments, provide limited UX insights, and can be difficult to generalize beyond one individual's experience. One option is to deploy either situated or construct-based methods to evaluate each candidate product in turn, but this is often not feasible on skills or resource grounds. A promising recent development has been the creation of simplified assessment methods that provide heuristic or streamlined approaches to detect UX issues (without necessarily requiring in-depth evaluation of each product). The pros and cons of these tools are summarized in Table 2. Many use a combination of situated and construct-based measures. For instance, the American Psychiatric Association (APA) provides a publicly available online 5-step "App Evaluation Model" to help clinicians and patients decide on adopting smartphone apps (53). The model includes a step on evaluating the UX, specifically focusing on "ease of use." Rather than recommending specific construct-based measures, the model suggests UX issues for consideration, such as whether the app can be customized or is easy to use on a long-term basis. Alternatively, stakeholders can rely on UX judgments made by clearing houses. Psyberguide, a nonprofit web site that publishes evaluations of mental health apps, includes an assessment of UX using The Mobile App Rating Scale (MARS), a 23-item questionnaire assessing UX constructs, including engagement, functionality, and aesthetics, as well as quality of information, subjective quality, and perceived impact (54). MARS gives an overall score to reflect "app quality," and has been shown to have high levels of inter-rater reliability and internal consistency (54). Such tools utilize a combination of situated and construct-based approaches. Given the emergence of such evidence-based tools, clinicians and researchers should be skilled in navigating the merits and challenges associated with the included situated and construct-based measures, and ensure the rationale behind their use is justified. Additionally, similar to other areas of medicine, where organizations recommend the use of certain tools, clinicians and patients should expect the evidence behind using certain UX assessment approaches to be provided and peer-reviewed.

Our review highlighted that there were no studies examining the potential relationships between clinical outcomes and the UX, though this remains an important area for future research.

https://doi.org/10.1017/S0266462319003507 Published online by Cambridge University Press

As Leigh highlights, it is difficult to imagine how governments would increase support for improving app development and implementation based on evidence explicitly focusing on usability without clear links to improved patient reported outcomes (55). However, some studies in our review, such as Attwood et al. (29) suggest that situated UX data can provide insight into behavioral aspects of illness, which through subgroup analysis can help improve understandings of target populations, as well as the kinds of interventions most likely to provide benefit. For instance, future research involving subgroup analysis of UX data from smartphone apps tracking manic behaviors in people with bipolar disorder could help clarify triggers for relapse in certain groups, and guide the timing of interventions for maximum benefit.

In addition, with the development of machine learning and novel human–machine interfaces, such as conversational agents and "Internet of things" devices, there is potential for even greater intervention complexity in digital mental health. Future interventions will have broader scope for personalization and adaptive intervention design, with clinicians and designers potentially able to modify the format and type of care provided according to patients' needs and response to therapy, such as in Tess, the psychological artificial intelligence chatbot built by X2AI Inc. (56). UX assessment is being conducted in relation to these interventions (56), and could grow to become particularly helpful in identifying which specific aspects of interventions help facilitate the best clinical outcomes. Future studies could utilize combinations of situated and construct-based measures to explore the potential effects of certain design and implementation modifications on clinical outcomes, such as whether sleep monitoring affects the timing of access to care in patients becoming unwell.

Conclusions

The spectrum of methods examined in this review highlights how informed trade-offs can guide approaches to UX assessment of mental health smartphone interventions with reference to resources, experiences, and service demands. Simple, bespoke measures may help detect easily fixable UX issues, while complex decisions such as comparison of iterations or competitive selection of one intervention over another require robust construct-based methods. Moving forward, more sophisticated UX assessment tools should be used for mental health smartphone interventions, and the relationship between UX data and clinical outcomes further explored.

Financial support. This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Conflicts of interest. The opinions expressed in this paper are those of the authors and do not necessarily represent the decisions, policies, or views of the WHO.

References

1. **World Health Organization** (2016) Global diffusion of eHealth: Making universal health coverage achievable. Report of the third global survey on eHealth Global Observatory for eHealth. World Health Organization: Geneva.
2. **Gulliver A, Griffiths KM, Christensen H** (2010) Perceived barriers and facilitators to mental health help-seeking in young people: A systematic review. *BMC Psychiatry* **10**. https://doi.org/10.1186/1471-244X-10-113.
3. **Dorsey ER, Chan YF, Mcconnell MV, *et al.*** (2017) The use of smartphones for health research. *Acad Med* **92**, 157–160.
4. **Sabesan S, Kelly J** (2015) Implementing telehealth as core business in health services. *Med J Aust* **202**, 231–233.
5. **Estai M, Kanagasingam Y, Xiao D, *et al.*** (2017) End-user acceptance of a cloud-based teledentistry system and Android phone app for remote screening for oral diseases. *J Telemed Telecare* **21**, 44–52.
6. **Zaliani S, Gilani MS, Nikbin D, *et al.*** (2014) Determinants of telemedicine acceptance in selected public hospitals in Malaysia: A clinical perspective. *J Med Syst* **38**, 111.
7. **George D, Hassali MA, HSS A-S** (2018) Usability testing of a mobile app to report medication errors anonymously: Mixed-methods approach. *JMIR Hum Factors* **5**, e12232.
8. **Knowles SE, Toms G, Sanders C, *et al.*** (2014) Qualitative meta-synthesis of user experience of computerised therapy for depression and anxiety. *PLoS ONE* **9**. https://doi.org/10.1371/journal.pone.0084323.
9. **Smith MS, Lawrence V, Sadler E, *et al.*** (2019) Barriers to accessing mental health services for women with perinatal mental illness: Systematic review and meta-synthesis of qualitative studies in the UK. *BMJ Open* **9**, e024803.
10. **Knaak S, Mantler E, Szeto A** (2017) Mental illness-related stigma in healthcare: Barriers to access and care and evidence-based solutions. *Healthc Manag Forum* **30**, 111–116.
11. **Gronholm PC, Henderson C, Deb T, *et al.*** (2017) Interventions to reduce discrimination and stigma: The state of the art. *Soc Psychiatry Psychiatr Epidemiol* **52**, 249–258.
12. **Bhugra D, Tasman A, Pathare S, *et al.*** (2017) The WPA-lancet psychiatry commission on the future of psychiatry. *Lancet Psychiatry* **4**, 775–818.
13. **Razzouk R, Shute V** (2012) What is design thinking and why is it important? *Rev Educ Res* **82**, 330–348.
14. **Scholten H, Granic I** (2019) Use of the principles of design thinking to address limitations of digital mental health interventions for youth: Viewpoint. *J Med Internet Res* **21**(1), e11528.
15. **Yardley L, Morrison L, Bradbury K, *et al.*** (2015) The person-based approach to intervention development: Application to digital health-related behavior change interventions. *J Med Internet Res* **17**, e30.
16. **Kylberg M, Haak M, Iwarsson S** (2018) Research with and about user participation: Potentials and challenges. *Aging Clin Exp Res* **30**, 105–108.
17. **Bakker D, Kazantzis N, Rickwood D, *et al.*** (2016) Mental health smartphone apps: Review and evidence-based recommendations for future developments. *JMIR Ment Health* **3**, e7.
18. **Torous J, Firth J, Mueller N, *et al.*** (2017) Methodology and reporting of mobile heath and smartphone application studies for schizophrenia. *Harv Rev Psychiatry* **25**, 146–154.
19. **Feather JS, Howson M, Ritchie L, *et al.*** (2016) Evaluation methods for assessing users' psychological experiences of web-based psychosocial interventions: A systematic review. *J Med Internet Res* **18**(6), e181.
20. **Nicholas J, Fogarty AS, Boydell K, *et al.*** (2017) The reviews are in: A qualitative content analysis of consumer perspectives on apps for bipolar disorder. *J Med Internet Res* **19**, e105.
21. **Torous J, Nicholas J, Larsen ME, *et al.*** (2018) Clinical review of user engagement with mental health smartphone apps: Evidence, theory and improvements. *Evid Based Ment Health* **21**, 116–119.
22. **Torous J, Andersson G, Bertagnoli A, *et al.*** (2019) Towards a consensus around standards for smartphone apps and digital mental health. *World Psychiatry* **18**, 97–98.
23. **Abrahão S, Bordeleau F, Cheng B, *et al.*** User experience for model-driven engineering: Challenges and future directions. In: Gray J, Kulkarni, eds. *Proceedings ACM/IEEE 20th international conference on model driven engineering languages and systems*. Piscataway, NJ, USA: IEEE Institute of Electrical and Electronics Engineers Inc., 229–236.
24. **Law ELC, Roto V, Hassenzahl M, *et al.*** (2009) Understanding, scoping and defining user experience: A survey approach. In: Olsen DR and Arthur RB, eds. *Proceedings conference on human factors in computing systems*. New York, NY, USA: ACM Press, 719–728.
25. **Baumel A, Birnbaum ML, Sucala M** (2017) A systematic review and taxonomy of published quality criteria related to the evaluation of user-facing eHealth programs. *J Med Syst* **41**(128). doi:10.1007/s10916-017-0776-6.

https://doi.org/10.1017/S0266462319003507 Published online by Cambridge University Press

26. **Ben-Zeev D, Wang R, Abdullah S, *et al.*** (2016) Mobile behavioral sensing in outpatients and inpatients with schizophrenia. *Psychiatr Serv* **67**, 558–561.
27. **Mackintosh M-A, Niehaus J, Taft CT, *et al.*** (2017) Using a mobile application in the treatment of dysregulated anger among veterans. *Mil Med* **182**, e1941–e1949. https://doi.org/10.7205/MILMED-D-17-00063.
28. **Orlowski S, Lawn S, Matthews B, *et al.*** (2016) The promise and the reality: A mental health workforce perspective on technology-enhanced youth mental health service delivery. *BMC Health Serv Res* **16**. doi:10.1186/s12913-016-1790-y.
29. **Attwood S, Parke H, Larsen J, *et al.*** (2017) Using a mobile health application to reduce alcohol consumption: A mixed- methods evaluation of the drinkaware track & calculate units application. *BMC Public Health* **17**, 1–21.
30. **ISO 9241-11:2018** (2018) Ergonomics of human-system interaction—Part 11: Usability: Definitions and Concepts: Geneva.
31. **Agnisarman SO, Madathil KC, Smith K, *et al.*** (2017) Lessons learned from the usability assessment of home-based telemedicine systems. *Appl Ergon* **58**, 424–434.
32. **Sheehan B, Lee Y, Rodriguez M, *et al.*** (2012) A comparison of usability factors of four mobile devices for accessing healthcare information by adolescents. *Appl Clin Inform* **3**, 356–366.
33. **Vilardaga R, Rizo J, Kientz JA, *et al.*** (2016) User experience evaluation of a smoking cessation app in people with serious mental illness. *Nicotine Tob Res* **18**, 1032–1038.
34. **Bangor A, Kortum PT, Miller JT** (2009) Determining what individual SUS scores mean: Adding an adjective rating scale. *J Usability Stud* **4**, 114–123.
35. **Dubad M, Winsper C, Meyer C, *et al.*** (2018) A systematic review of the psychometric properties, usability and clinical impacts of mobile mood-monitoring applications in young people. *Psychol Med* **48**, 208–228.
36. **ISO/IEC 9126-1** (2001) Software engineering—Product quality—Part 1: Quality model. International Organization for Standardization: Geneva.
37. **Kortum PT, Bangor A** (2012) Usability ratings for everyday products measured with the system usability scale. *Int J Hum Comput Interact* **29**, 67–76.
38. **Ben-Zeev D, Brenner CJ, Begale M, *et al.*** (2014) Feasibility, acceptability, and preliminary efficacy of a smartphone intervention for schizophrenia. *Schizophr Bull* **40**, 1244–1253.
39. **Brooke J** (1996) SUS: A "quick and dirty" usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, eds. *Usability evaluation in industry*. London: Taylor & Francis, 189–194.
40. **Lewis JR** (1992) Psychometric evaluation of the poststudy system usability questionnaire: The PSSUQ. *Proc Hum Factor Soc Annu Meet* **36**, 1259.
41. **Venkatesh V, Davis FD** (2000) A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Manage Sci* **46**, 186–204.
42. **Lund AM** (2001) Measuring usability with the USE questionnaire. *Usability User Exp* **8**, 3–6.
43. **Povey J, Mills PPJR, Dingwall KM, *et al.*** (2016) Acceptability of mental health apps for Aboriginal and Torres Strait Islander Australians: A qualitative study. *J Med Internet Res* **18**, e65.
44. **Shen N, Levitan M-J, Johnson A, *et al.*** (2015) Finding a depression app: A review and content analysis of the depression app marketplace. *JMIR mHealth uHealth* **16**, e16.
45. **Fuller-Tyszkiewicz M, Richardson B, Klein B, *et al.*** (2018) A mobile app-based intervention for depression: End-user and expert usability testing study. *J Med Internet Res* **5**, e54.
46. **Maramba I, Chatterjee A, Newman C** (2019) Methods of usability testing in the development of eHealth applications: A scoping review. *Int J Med Inform* **126**, 95–104.
47. **Bojko A** (2013) *Eye tracking the user experience: A practical guide to research*. New York: Rosenfeld Media.
48. **Righi C, James J, Beasley M, *et al.*** (2013) Card sort analysis best practices. *J Usability Stud* **8**, 69–89.
49. **Speicher M, Both A, Gaedke M** (2014) Ensuring web interface quality through usability-based split testing. In: Casteleyn S, Rossi G, Winckler M, eds. *Web engineering. ICWE 2014. Lecture notes in computer science*, vol. **8541**. Cham: Springer, 93–110.
50. **Woods L, Cummings E, Duff J, *et al.*** (2017) Design thinking for mHealth application co-design to support heart failure self-management. *Stud Health Technol Inform* **241**, 97–102.
51. **Naslund JA, Aschbrenner KA, Barre LK, *et al.*** (2015) Feasibility of popular m-health technologies for activity tracking among individuals with serious mental illness. *Telemed e-Health* **21**, 213–216.
52. **Long AF, Gambling T, Young RJ, *et al.*** (2005) Acceptability and satisfaction with a telecarer approach to the management of type 2 diabetes. *Diabetes Care* **28**, 283–289.
53. **Torous JB, Chan SR, Gipson SY-MT, *et al.*** (2018) A hierarchical framework for evaluation and informed decision making regarding smartphone apps for clinical care. *Psychiatr Serv* **69**, 498–500.
54. **Stoyanov SR, Hides L, Kavanagh DJ, *et al.*** (2015) Mobile app rating scale: A new tool for assessing the quality of health mobile apps. *JMIR mHealth uHealth* **3**, e27.
55. **Leigh S** (2016) Comparing applets and oranges: Barriers to evidence-based practice for app-based psychological interventions. *Evid Based Ment Health* **19**, 90–92.
56. **Fulmer R, Joerin A, Rauws M, *et al.*** (2018) Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: Randomized controlled trial. *JMIR Ment Health* **5**, e64.