

Acceptability, Engagement and Exploratory Outcomes of an Emotional Wellbeing App: Mixed Methods Preliminary Evaluation and Descriptive Analysis

Mia Eisenstadt 1,2

Shaun Liverpool 1,2,3

Athina- Marina Metaxa 2,4,5

Roberta Maria Ciuvat 2,4

Courtney Carlsson 2

1 Evidence Based Practice Unit, Anna Freud National Centre for Children and Families and University College London, London, United Kingdom

2 Paradym Ltd, Bloomsbury, London, United Kingdom

3 Faculty of Health, Social Care and Medicine, Edge Hill University, Ormskirk, United Kingdom

4 Division of Psychology and Language Sciences, University College London, London, United Kingdom

5 Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom

Abstract

Background: There is growing evidence suggesting that the emotional wellbeing of the general public have been negatively affected within the past year. Consequently, there have been an increase in demand for wellbeing support. Whilst there is substantial empirical support for mental health apps that target diagnosed conditions, there is less research on emotional wellbeing apps. Amongst existing wellbeing apps, few studies have been conducted on apps that are based on lived experience and those that seek to enhance user's understanding of their emotional patterns. Thus, the acceptability of these novel apps will require further evaluation before upscaling.

Objective: The primary aim of this evaluation is to describe the acceptability, engagement, and preliminary outcomes of using an app (Paradym) designed to promote emotional wellbeing and positive mental health.

Methods: This is a pre- and post -mixed methods single-arm evaluation, aggregate with digital analytics data. We anonymously collected real world data on demographics, wellbeing, usability and acceptance of the app using validated questionnaires and open-ended questions. Participants tested the app for a minimum of 2 weeks before completing follow up measures. Google Analytics was used to record levels of app engagement. Chi Square and t-tests were conducted to analyze quantitative data, and a thematic analysis approach was adopted for qualitative data.

Results: A total of 115 participants completed baseline questionnaires, of which 91 (79.13%) users downloaded the app. The sample was diverse in terms of ethnicity, including of 50 (43.5%) people self-identifying as belonging minority ethnic groups. The majority of the sample were females (78, 67.24%) and between the ages 18-25 (39, 33.62%). Thirty-four app users who completed questionnaires at baseline and follow-up providing valuable feedback to inform future directions of Paradym. Favorable themes emerged describing the app's content, functionality and underlying principles. Although usability feedback varied across items, a considerable number of participants (22, 64.7%) found that the app was easy to use. Google Analytics revealed that at least 27 (80%) persons used the app on a daily basis. Based on preliminary observations, app users experienced increased mental wellbeing. Post hoc analyses indicated that the reduction in depression scores ($t_{33} = -2.16$) and the increase on the wellbeing measures ($t_{33} = 2.87$) were statistically significant. No adverse events were reported during the evaluation period.

Conclusion: The findings of this evaluation are encouraging and documents positive preliminary evidence for the Paradym app.

Keywords: smartphone, app, wellbeing, awareness, mental health, formative

Acceptability, Engagement and Exploratory Outcomes of an Emotional Wellbeing App: Mixed Methods Preliminary Evaluation and Descriptive Analysis

Introduction

Emotional well-being, as an important predictor of human health and longevity with the potential to reduce the risk of physical and mental health disorders, is well established in the literature (Agteren & Iasiello, 2020; Bell et al., 2019; Chida & Steptoe, 2014; Huppert, 2009; Stranges et al., 2014; World Health Organization, 2001; Zhang et al., 2014). Concerningly, research suggests that emotional wellbeing has declined in the general population within the past year. Therefore, there has been an increased need for accessible support services to meet this demand. The evidence suggests that one mechanism for improving levels of emotional wellbeing is raising one's emotional self-awareness. Further, apps have been identified as a promising mode of delivery based on its accessibility, scalability and potential to provide anonymous services.

Emotional Wellbeing and Emotional Self Awareness

Emotional wellbeing is defined as “a component of mental health that includes happiness, interest in life, and satisfaction” [27]. Researchers and policy makers have made the case for the promotion of emotional wellbeing to advance human health and reduce the costs associated with poor mental wellbeing and increased risk of mental health disorders (Huppert; 2009; Van Agerteran, 2021).

Emotional awareness refers to the ability to recognise one's emotions as they are being experienced (Veirman et al., 2011) and communicate them (Agnoli et al, 2019). Low levels of emotional awareness have been associated with poor emotional wellbeing and increased risk of mental health disorders, such as anxiety (Kauer, 2012). High levels of emotional awareness have been associated with a reduction in depressive symptoms and increased positive affect and emotional regulation (Beckner et al., 2010; Chambers et al., 2008; Ciesla & Roberts, 2007; Gratz & Roemer, 2004). There is also evidence suggesting that emotional awareness can enable flexibility of behavioural responses to difficult emotions and contribute to improved relationships with others (Gross & Munoz, 1995; Lane, 1987; Sutton, 2016; Sutton et al., 2015).

Emotional Wellbeing Apps

Based on the positive associations between emotional self-awareness and emotional wellbeing, there have been a growing number of digital interventions adopting this approach (Mckay et al.,

2019; Monge, Roffarello & De Russis, 2019). These apps offer a route to support users to improve their emotional wellbeing through their personal digital devices (e.g., Smartphones and Ipad). Therefore, wellbeing apps can be used at any time, independent of location, with the added bonus that users can use the apps privately (Bakker et al. 2016). To this end, there has been both a growing demand for, and development of, wellbeing apps that has accelerated in recent years (Coyle et al., 2014; Neary & Schueller, 2018). A systematic review of 53 apps revealed that within the existing range of wellbeing apps, apps have aimed to increase users' management of emotions through primarily, mindfulness, cognitive behavioural and mood tracking approaches (Eisenstadt et al). For instance, popular meditation and breathing apps, such as HeadSpace and Calm, have encouraged users to become aware of emotions and manage thoughts and emotions through daily meditation and breathing exercises (Champion et al., 2018; Mani, et al., 2015; Sieverdes et al., 2017).

Other apps have provided strategies from Cognitive Behavioural Therapy, such as challenging negative thoughts, catastrophizing and facilitating positive reappraisal (Bakker et al., 2018; Nave, 2017). Further, some apps use mood tracking to enable users to identify emotions or moods and record and monitor these moods as one approach to emotional awareness (McClould et al., 2020; Throuvala et al., 2020; Bakker et al., 2018; Ponzio et al., 2020). In response to particular moods, some apps provide suggested activities to manage moods or emotional responses (Bakker, et al., 2018). Although less frequently available, some wellbeing apps, such as Mood Hacker, has been informed by principles of positive psychology which encourages users to identify their strengths, and increase their mindful self-awareness (Birney et al., 2016).

Of the current range of wellbeing apps, it was noted that apps less frequently drew on lived experiences to explore wellbeing topics. Similarly, storytelling as a medium to convey psychoeducational concepts were less studied within the literature (Eisenstadt et al., 2021). The closest examples to this approach used fictional characters to guide users through the app (Litvin et al., 2020), or real-life testimonials of other people who overcame difficult circumstances (Daugherty et al. 2018). Further, to best of our knowledge, there are no studied apps that encourage users to go beyond the exploration of their current emotional states, to inquire into their emotional patterns as a means of understanding their personality traits and tendencies (Eisenstadt et al., 2021). Given the rapid development of apps, and relatively fewer studies on apps that use storytelling, lived experience, and or emotional patterns, there appears to be a significant gap in this area that could benefit from further exploration. This is also important, given the calls to explore wider techniques that could impact levels of engagement with digital interventions (Eisenstadt et al., 2021).

The App Under Evaluation

The Paradym app was developed to support users to increase their emotional awareness through learning about their emotional patterns to contribute to enhanced emotional wellbeing, increased self-awareness and improved life satisfaction. The app uses lived experience and storytelling as a starting point to introduce users to psychoeducational content that covers key areas of a person's life (i.e., love and relationships, body image, work and success, and identity). The app also incorporates concepts of lived experience which refers to the first-hand, personal involvement of the user and the meaning that their past experience brings to current situations (Lindseth & Norberg, 2004). As a starting point, the app was designed for young adults aged 18 and above. Before further development of the app, and to ensure successful implementation, concepts of the Technology Acceptance Model was applied (TAM). The model proposes that the assessment of perceived usefulness and perceived ease of use can determine whether users will engage with the new digital intervention. It is also recommended that user anxiety be monitored as previous studies identified anxiety as an external variable to the Technology Acceptance Model.

The Process of Evaluation

The current evaluation was further guided by the Medical Research Council (MRC) guidelines, which suggests addressing possible uncertainties identified during development of new interventions [30]. Therefore, in this evaluation, we aimed to describe the acceptability, engagement and preliminary outcomes of using an app (Paradym) designed to promote emotional wellbeing by adopting novel approaches. At this formative stage, the paper also aimed to record and report any negative consequences of using the app. To address these aims, the following questions were addressed:

Engagement

1. What are the characteristics of users accessing and using the app?
2. What are the users' levels of engagement with the app?
3. Are there differences between completers and non-completers?

Acceptability

4. What did users like about the app?
5. How usable or useful did users find the app?

Preliminary outcomes

6. Would users' wellbeing increase after the intervention period?
7. Would users' mental health symptoms (e.g., depression) decrease after the intervention period?
8. Would participants experience severe/high levels of anxiety or other negative experiences during the intervention period?

Feasibility

9. Can users' feedback be used to inform upgrades to the app?
10. Would it be feasible to conduct further evaluations of the app using the current recruitment strategies and outcome measures?

Methods

Design

The current evaluation adopted a mixed methods study design with a pre-post single arm approach aggregate with digital analytics data. We adhered to the CONSORT guidelines [31] where applicable, and registered the user testing protocol on the Open Science Framework a priori [32]. The only change in protocol was the decision to not collect data on anxiety at baseline. As the purpose of the app was not specifically designed to address mental health symptoms (e.g. anxiety), to minimize the burden of completing lengthy questionnaires, we only collected data from the anxiety measure at follow up. However, this was partly in line with our research question to explore if the use of the app introduces high levels of anxiety.

Participants

Real world (i.e. in-the-wild) data were collected from an international pool of potential users. The user testing was advertised on Facebook campaigns and through social media over a 6-week period between May and June 2021. Anyone coming in contact with information about the user testing had the opportunity to participate in the user testing. To be eligible for the user testing, participants needed to be over the age of 18 and not have a diagnosable mental health condition.

Procedures

Participants enrolled in the user testing through a link that was provided on the web advertisement and on the Paradym website. The link provided access to online questionnaires for the baseline assessment and instructions on how to download the app from the Apple app store for iPhone users or Google Play for Android users. An automated email was sent to participants 2 weeks after completing the first battery of questionnaires for participants to complete follow up questionnaires. The questionnaires at Time 2 included measures completed at Time 1 and additional measures to capture acceptability of the app and user experience. Participants who completed questionnaires at both time points were entered into a free prize draw with the opportunity to win one of three £20 Amazon Gift Cards.

Intervention description

Paradym is an app aimed at supporting emotional wellbeing within the general population and targets adults over 18 years of age. It aims to provide users with support to develop greater emotional awareness and enhanced life satisfaction. Paradym was initially a crowdfunded project in the early stages of its development specifically to ensure that commercial interests were not taken into account in the early development of the intervention. The app provides a low-cost self-guided programme with the aim of supporting users' levels of emotional awareness and emotional wellbeing. Paradym is a standalone app but it can also be used in combination with online group coaching sessions. However, for this evaluation, participants only accessed the standalone app.

Paradym was designed drawing on evidence-based strategies and designed in response to key topics that users identified that they would like support with such as success, body, identity and love and relationships. Clinical psychologists, coaches & researchers were involved in the development of the content and the selection and review of psychoeducational content, suitable evidence-based strategies and exercises (e.g. for example, journaling and recording personal notes).

Although an integrative theoretical approach underpins Paradym, key therapeutic theories such as Acceptance and Commitment Therapy (ACT) [33], and Schema Therapy informed the lived experience and storytelling activities. Following consultation with users, further theory-informed strategies were applied to produce content that users indicated were important. Paradym is made up of the following key strategies:

1. Digital lessons

Psychoeducational digital lessons are structured into 5 pillars: Aware, Success, Love, Identity and Body. These topics were chosen based on the results of an end-user consultation conducted early in the development phase. Examples of psychoeducational content include an introduction to developing emotional awareness [38], supporting users with the identification and understanding of their emotions [39], and identifying and strengthening values [33].

Each digital lesson begins with an explanation of the psychoeducational concept based on lived experience through personal storytelling. Storytelling is expected to help to make the psychological content relatable, foster perspective taking, and support identification with the storyteller [41–43]. Storytelling has been found to boost engagement in both psychological interventions and apps [44]. Digital lessons are provided as chapters for reading, audio for listening, or video to watch speakers present the content.

2. Emotional Patterns

At the end of each digital lesson, the user is asked to identify their own emotional pattern in relation to the content, which draws on some concepts from Schema Therapy [45]. The purpose of is for the user to gain greater emotional awareness. To achieve this the app guides users through specific tasks such as to reflect where they lie on a spectrum of optimism and pessimism.

3. Reflections

Reflections are provided to the users via push notifications to prompt engagement with the app through a new daily reflection. Research shows that there are many benefits of reflection, such as supporting determination, with tasks despite a stressful context [48] and that reflection may reduce rumination that is linked with internalising difficulties and interpersonal conflict [49].

Further, notifications have been found to increase app engagement. Several studies have indicated that the use of mobile phone app interventions in psychological content must be combined with the active engagement of users [50,51].

Data collection tools

Demographics

To understand the demographic profile of users accessing and using the interventions, participants were asked to provide anonymous information about their age, gender, ethnicity and country of residence. Age was captured using age groupings of approximately 10 years (e.g., 18-25, 26-35 and 36-45). To allow inclusivity, participants were asked to enter their gender and ethnicity using free text. Participants also selected their geographic location (e.g. city, state, country) from a global list. For the purpose of analysis, ethnicity was then dichotomized as Caucasian or Minority Ethnic Groups. Age was recategorized to represent 45+ years, as there were fewer persons in each age category above 45 years. Location was then categorized as UK, USA and Other to represent the majority of responses.

Emotional wellbeing and mental health factors

The outcome measures used to assess wellbeing were the World Health Organisation's wellbeing measures (WHO-5) [53], the Satisfaction with Life Scale (SWLS)[54], the Emotional Self-awareness Scale (ESAS-R) [55], the Patient Health Questionnaire (PHQ-9)[56] and the General Anxiety Disorder test (GAD-7) [57]. The WHO-5, SWLS, ESAS and PHQ9 were administered pre- and post-

app use. The GAD-7 was administered post-app use alongside the System Usability Scale (SUS) [58] and the semi-structured acceptability questionnaire.

WHO-5

The World Health Organisation-5 (WHO-5) allows for a brief assessment (under 1 minute) of wellbeing over a 2-week period [53]. Individuals are asked to indicate for each of the 5 statements how they felt over the past 2 weeks using a 6-point Likert scale ranging from 0 = “at no time” to 5 = “all of the time”. The WHO-5 is derived from a 28-item version based on items from the Zung scales for depression, distress and anxiety as well as from the General Health Questionnaire and the Psychological General Well-Being Scale [60]. WHO-5 has been validated as a measure for depression for both adolescents and older adults, with high measurement invariance [61]. A high score indicates a high level of wellbeing.

SWLS

The Satisfaction with Life Survey (SWLS) [54] is a short 5 item scale designed to measure global cognitive assessment of satisfaction with one’s life. The estimated time for completion of the SWLS has been reported as approximately 1 minute. The SWLS has been shown to have very high construct validity, with Cronbach’s alpha = 0.85 - $\alpha = 0.87$ [62], and moderately high reliability (alpha = 0.78) [63]. A high score obtained from the SWLS indicates a high level of life satisfaction.

ESAS-R

The Emotional Self-Awareness Scale-Revised (ESAS-R) [55] is a 30 item scale and all items are on a 5-point Likert scale ranging from 0 to 5 (0 = Never, 1 = Very Little, 2 = Sometimes, 3 = Often, 4 = A lot). Subscales range from 0 to 20. Total scale ranges from 0 – 132. Subscales included: Recognition, Identification, Communication and Contextualisation. ESAS-R has been shown to have high validity and reliability, (alpha = .83 - .90) [64]. A high score from this measure is indicative of a high level of emotional self-awareness.

PHQ-9

The Patient Health Questionnaire (PHQ-9) is a depression scale which scores each of the 9 Diagnostic Statistical Manual -IV criteria as 0 (not at all to 3 (nearly every day). The PHQ-9 has been validated for use in primary care [44]. It is not a screening tool for depression but can monitor severity of symptoms and response to treatment. Scores over 10 have good sensitivity (88%) and specificity

(88%) for diagnosis of major depression by interview. It has high internal reliability, Cronbach's $\alpha = 0.89$ [65]. The construct validity of the PHQ-9 is also high with community and clinical samples [66]. A high score obtained from PHQ-9 is indicative of a severe level of depression.

GAD-7

The 7-item Generalized Anxiety Disorder Scale (GAD-7) is a brief clinical measure that assesses the presence and severity of Generalized Anxiety Disorder. The self-report scale asks how often during the last 2 weeks individuals experienced symptoms of GAD. Total scores range from 0–21 with cut-off scores of 5, 10, and 15 being indicative of mild, moderate, and severe anxiety respectively. Increasing scores on the GAD-7 are strongly associated with greater functional impairment in real-world settings. Sample items are rated from 0 (not at all) to 3 (nearly every day). Scores over 10 have good sensitivity (89%) and specificity (82%) for diagnosis of GAD by interview, and the scale has high internal reliability, as suggested by a Cronbach's α of 0.92 [57]. The scale has been widely used and considered a valid and reliable screening tool in previous research, presenting good reliability, factorial and concurrent validity [67]. A high score on the GAD-7 measure is an indicator of severe anxiety.

Acceptability and usability

System Usability Scale

The Systems Usability Scale (SUS) is a 10-item scale with all items measured on a 5-point Likert scale ranging from 1 to 5 (1 = strongly disagree and 5 = strongly agree) with total scores ranging from 0 to 100. This widely used scale in technology gives an overview of subjective assessments of the usability [58]. The SUS has been shown to have excellent construct validity and reliability (0.81 - 0.94) [68,69]. A high score on the SUS indicates a high level of usability with a score above 68 considered above average.

Semi-structured questions

In addition to the SUS, we aimed to capture information about acceptability of the app by identifying facilitators and barriers to using the app. For the open-ended questions the following questions were adapted from a study by Kern et al., (2018) [59] to fit the current evaluation:

1. Why would you use [Paradym]?
2. Why would you not use [Paradym]?

3. Explain why [or why not] you would prefer to use [Paradym] to seeing a mental health professional?

Adverse events

To capture any side effects of using the app, participants were provided with the email address for the research team and requested to contact a member of the research team if they experienced any distress or discomfort or experienced any other issue as a result of their usage. Adverse events (i.e. side effects) were identified and recorded as any untoward medical/clinical occurrence which did not necessarily have a causal relationship with the intervention. Any adverse events arising during the evaluation period were assessed for severity, causality, seriousness, and expectedness (i.e., relating to the information provided by the app).

Data analysis

Descriptive statistics were calculated for participant characteristics at baseline (pre-intervention), and Google Analytics estimates were used to report engagement of the app. To address the aims of the evaluation, the main focus was on descriptive data. However, some exploratory significance tests were conducted on within-group mean differences at the 2 time points (i.e., baseline and follow-up) on the mental health and wellbeing measures. Paired samples t-tests were conducted on continuous data and Chi-square tests were used for categorical variables in SPSS. Qualitative data from the semi-structured open ended questions were captured in questionnaire format and coded in categories and analyzed using the steps outlined for thematic analysis [70]. The data was coded line by line and then clustered into provisional themes. Each candidate theme was reviewed by going back to the coded data, to check that the meaning was accurate. Next, the uncoded datasets were revised, to ensure that data were relevant to the research question. Finally, the provisional theme titles and the codes which they contained were checked to confirm that they remained relevant and accurate. Any discrepancies that arose were discussed at team meetings and only included after a consensus was reached.

Ethical considerations

We were guided by the research ethical principles for ensuring the rights, safety, dignity and wellbeing of the participants in this evaluation were protected [71]. All data were collected remotely from a non-clinical population, and participants voluntarily provided informed consent for their anonymized deidentifiable data to be used for research purposes. The research team's email address was available for participants to contact us if they experienced any discomfort as a result of using the app. The researchers were prepared to signpost users to relevant wellbeing support and resources if required.

Owing to the formative nature of the evaluation, findings were not expected to be generalizable or transferable in any way. We also consulted other researchers affiliated with academic institutions and practitioners known to the study team who reviewed the evaluation protocol and provided oversight with respect to research ethical guidance. In light of the above, this evaluation was viewed as user testing/service evaluation and therefore was exempt from formal ethical approvals by the Health Research Authority (Appendix 1).

No formal ethical approval was required in similar acceptability studies [64,72,73]. As explained by Ahtinen et al. (2013): “Ethics committee approval was not acquired as the study was deemed to involve minimal risk and the focus was on studying mainly user experiences” [74] . Additionally, the present evaluation does not involve a clinical population, like that of Ly et al. (2017), according to whom; “Since this pilot trial involved a non-clinical population, it was considered exempt from registration in a public trials registry.” [75] .

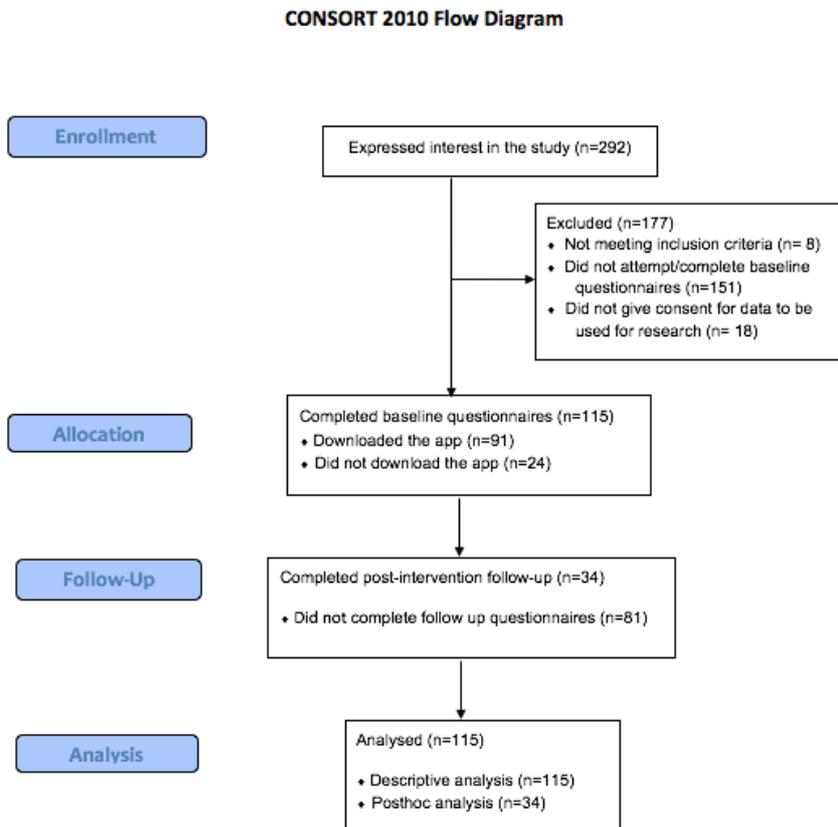
Results

Questions around engagement, acceptability and preliminary outcomes (1-8) are addressed here in the results section, and questions on feasibility (9,10) are addressed in the Future Directions subsection of the Discussion.

Descriptive data

During the 2 weeks the user testing was advertised, 292 participants expressed interest to be part of the preliminary evaluation. 151 individuals either did not attempt or did not complete the baseline questionnaires and therefore did not reach the relevant page to be able to download the app. However, data from 18 participants were not considered as they did not provide consent to process their information for research purposes. Data from 8 participants were excluded as they indicated they were not eligible for the evaluation based on having mental health diagnoses. Consequently, data from 115 participants (i.e. completing baseline questionnaires and providing consent) were included in the evaluation. Of these, 91 participants downloaded the app and 34 of the participants completed questionnaires at both baseline and follow up and were included in the final analysis. Figure 1 depicts the flow of participants throughout the evaluation period.

Figure 1 Flowchart of participants progress through the evaluation period



Participant characteristics

At baseline (n=115), the participants were a self-identified diverse sample with n = 50 (43.48%) belonging to minority ethnic groups, with the majority female (n = 78, 67.24%). The majority of the sample (n = 80, 68.97%) were from different cities across the USA. Participants' ages ranged from 18-65+ years with the mode being 18-25 years. Table 1 provides further details of the sample. As for mental health and wellbeing characteristics, the participants were generally “dissatisfied” with their life (M = 20.16, SD = 6.24) and had “moderate” symptoms of low mood or depression (PHQ-9: M = 10.35, SD = 6.65). Participants also reported “average” scores on emotional self-awareness (ESAS: M = 56.95, SD = 13.94) and general wellbeing (WHO-5: M = 17.22, SD = 4.96).

Table 1.

Characteristics	Baseline (N=115)	Follow up (N=34)
Gender, n (%)		
Male	29 (25,00)	9 (26,47)
Female	79 (67,24)	23 (67,65)
Transgender	2 (1,72)	0 (0)
Other/prefer not to answer	7 (6,03)	2 (5,88)
Age, n (%)		
18-25	39 (33,62)	13 (38,24)
26-30	21 (18,10)	3 (8,82)
31-35	15 (12,93)	0 (0)
36-40	15 (12,93)	5 (14,71)
41-45	7 (6,03)	4 (11,76)
46-50	4 (3,45)	0 (0)
51+	15 (12,93)	9 (27,47)
Ethnicity/race,		
White	66 (56,90)	24 (70,59)
Black	12 (10,34)	2 (5,88)
Hispanic	4 (3,44)	0 (0)
Asian	20 (17,24)	6 (17,65)
Indian	1 (0,86)	0 (0)
AAPI (Asian Americans and	1 (0,86)	1 (2,94)
Jewish	2 (1,72)	0 (0)
Mixed	10 (8,62)	1 (2,94)
Country, n (%)		
USA	80 (68,97)	23 (67,65)
UK	32 (27,59)	10 (29,41)
Other	1 (0,86)	1 (2,94)
Did not answer	3 (2,59)	0 (0)

Engagement with the intervention

App usage data were made anonymous to comply with the General Data Protection Regulation and research ethical guidelines. 91 users downloaded the app and therefore it was not possible to separate the data of users that used the app only and those that used the app and completed the final battery of measures. During the evaluation period, all users (100%) logged into the app at least once. The average usage across all participants during the evaluation period was 23.53 minutes. Reflections received 215 views. Users returned to the app an average of 4.5 times with each session lasting an average of 5.22 minutes. Overall, the Daily Active Use Monthly Active Use Ratio (i.e., the proportion of users that engage with the app in a single day calculated by dividing the number of daily users by the number of monthly users) was 80% according to Google Analytics.

Attrition/dropout

In terms of attrition, 81 (70.4%) of users did not complete follow up questionnaires. The remaining 34 (29.57%) of the 115 participants completed questionnaires at the 2 required time points. Overall there were no significant differences in demographic or mental health and wellbeing characteristics. However, based on descriptive data, participants who completed the evaluation had slightly lower mean scores **at baseline** on all outcome measures including SWLS (18.94 vs. 20.67), WHO5 (16.18 vs. 17.65), ESAS (55.95 vs. 57.37) and PHQ-9 (10.12 vs. 10.45).

Acceptability

Thematic analysis of the qualitative data led to the development of 3 key themes. Participants provided favorable feedback describing acceptance of “the App’s content”, “the App’s functionality” and “the App’s underpinning principle”. Due to the interaction of these themes, users’ feedback indicated that there is an element of app versatility, with users describing use of the app at different times of the day. Some described it as part of their morning routine or used the Paradym app in conjunction with other wellbeing and health apps.

The app’s content

In this first theme, users commented on the content of the app, expressing positive evaluations of the app's chapters, reflections and exercises – and reported that listening to or reading the content helped

them with their understanding of themselves and their mental health. For example, one user expressed: “[*Because of the app, I am*] studying myself more closely which has benefited my mental health”. (Female, 51+)

Other users recognized the value of Paradym as a form of additional support separate to that of a professional. They said that Paradym could assist with self-help while their mental health professional could help with the application of the in-app content. For example, one user expressed: “*Paradym allows me to individually learn and apply it onto myself while training myself also. Seeing a mental health professional could allow them to help me directly.*” (Female, 18-25)

The app’s functionality

In this second theme, users reported that the combination of videos and text as aspects of the app’s functionality were helpful when working through the exercises as a way of building knowledge. For example, one user said: “*The videos were actually pretty engaging when I got around to watching them, and I did feel like I gleaned some insights and self-awareness from that.*” (Gender not reported, 18-25)

Others reported valuing the app as a chance to reflect and understand emotions and found the daily reflection notifications useful. For example: “*[A chance to] reflect on my emotions and life. Works best with the daily prompts.*” (Female, 18-25)

The app’s underpinning principle

In the third theme, users commented on the app’s underpinning principle. Participants reported benefits from building increased awareness of their emotional and behavior patterns, moods and having an opportunity to explore and record these. For example, one participant said: “[*The app allowed...*] building consistency and having a place to collect my thoughts about my mood and behavioral patterns.” (Female, 26-30)

Similarly, another user reported that developing self-awareness was important and linked to their personal aspirations. For example, one participant reported: “[*I would use Paradym...*] to discover bits about myself and become the person I knew I was capable of being.” (Male, 41-45)

Usability analysis

The mean usability score was almost 60 (M = 59,77 (SD= 23,65)) out of a possible 100. Participant responses varied across the items on the SUS. On the learnability subscale over 50% of the sample were neutral or disagreed that they would need help to use the app (Item 4), and over 50% were

neutral or agreed that they could learn to use the app quickly (Item 7). In the same vein, 50% of users reported feeling confident enough to use the app (Item 9). Table 2 provides further details on participants' usability experience.

Table 2

Systems Usability Scale self-report of items post intervention

Item	Agreement (N%)	Neutral (N%)	Disagreement (N%)	Missing data
1. I think that I would like to use Paradym frequently	14 (41%)	9 (26%)	11 (33%)	0
2. I found Paradym unnecessarily complex	12 (35%)	6 (18%)	16 (47%)	0
3. I thought Paradym was easy to use	14 (41%)	11 (33%)	9 (26%)	0
4. I think that I would need the support of a technical person to be able to use Paradym	3 (9%)	6 (18%)	25 (73%)	0
5. I found the various functions in this app were well integrated	15 (42%)	12 (35%)	7 (23%)	0
6. I thought there was too much inconsistency in Paradym	7 (20.6%)	9 (26,5%)	17 (50%)	1 (2.94%)
7. I would imagine that most people would learn to use Paradym very quickly	22 (64.7%)	6 (17.6%)	5 (14.7%)	1 (2.94%)
8. I found Paradym very cumbersome to use	12 (35.3%)	7 (20.6%)	14 (41.2%)	1 (2.94%)
9. I felt very confident using Paradym	17 (50%)	8 (23.5%)	9 (26.5%)	0
10. I needed to learn a lot of things before I could get going with Paradym	7 (20.6%)	8 (23.5%)	17 (50%)	2 (5.88%)

Adverse events

Regarding possible side effects or adverse events, no reports from participants were received during the evaluation period. Therefore, the research team is unaware of any harmful outcomes that could result from using the app.

Preliminary efficacy

Based on the available data from the participants who completed the evaluation, a descriptive analysis showed that mean scores slightly improved on the SWLS (+1.56), the WHO-5 (+2.03), and the ESAS (+1.75). However, the increase was statistically significant only for the WHO-5 scores ($t(33) = 2.87$). In addition, participants' depression scores also decreased over time (-1.56) and were found to be statistically significant ($t(33) = -2.16$). Participants' mean anxiety levels were within the mild category (8.24 (SD = 6.59)) during the intervention period. Table 3 provides further descriptions of the sample (N=34) included in the final analyses.

Table 3

Descriptive statistics for each outcome measure at baseline and follow-up

Outcome	Baseline	Follow-up	Coefficient	Cohen's d	95% CI	P value
SWLS ^a	18.94 (5.76)	20.50 (7.03)	0.73	0.37	[-0.11;	0.07
WHO-5 ^b	16.18 (4.70)	18.21 (4.94)	0.64	0.51	[0.02;	0.007
ESAS-R ^c	55.95 (12.50)	57.70 (14.27)	0.50	0.14	[-0.34;	0.57
PHQ-9 ^d	10.12 (6.25)	8.56 (6.35)	0.78	-0.38	[-0.85;	0.038
GAD-7 ^e	NA	8.24 (6.59)	NA	NA	NA	NA

a SWLS = Satisfaction with Life Survey; b WHO-5 = World Health Organisation Wellbeing Measure; c ESAS-R Emotional Self Awareness Scale Revised, d PHQ-9 = The Patient Health Questionnaire, e GAD-7= General Anxiety Disorder test.

Discussion

Key findings and interpretations

The aim of this user testing and initial service evaluation was to describe the acceptability, engagement, and preliminary outcomes of using a new version of a wellbeing app. The findings indicated that users that downloaded the app (n=91) generally accessed and used the app during the

intervention period for an average of 5.22 minutes with an average of 4.5 sessions per user over the 2-week period. A fair number of participants that completed the post assessment (n=10, 29.41%) engaged with the app content in all 5 pillars. Despite the observed level of engagement (average = 23.53 minutes), there was high attrition (70.4%) during the evaluation period. Although there were no significant differences between the users who completed the evaluation and those who dropped out, mean scores for mental health and wellbeing measures were higher in those that dropped out. It is possible that people who experienced more psychosocial difficulties found it difficult to engage with the app or to respond to the outcome measures [76]. Mental health apps generally have high dropout rates and some studies have found that once users learn a skill or the knowledge from a particular app, they stop using it [77]. It is also conceivable that, for users with high life satisfaction, they may have had less need for the knowledge provided in the app or felt they had obtained what they needed before the end of the 2-week period.

Participants who completed the evaluation provided favorable feedback in terms of the app's content, functionality and underlying principles. Although usability feedback varied across items of the SUS, a high percentage of participants found that it was easy to use the app. This can be attributed to evidence suggesting the importance of specific features that could influence users' experiences. For example, younger digital users are more likely to engage with interventions that have features such as videos, limited text, ability to personalize, ability to connect with others and options to receive text message reminders [78]. Also, the promotion of increased self awareness is becoming popular among young people in their early twenties [79]. Therefore, since emotional awareness is one of the main purposes of Paradym, and a fair amount of content is presented via videos or audio, users could have responded particularly favorably. Similarly, Paradym includes customization features, exercises and text material. Contrary to Liverpool et al., 2020, users in this evaluation responded favorably to the text-based content (as per the digital lessons). Further research is needed to see whether specific features may have more or less impact on usage between various age groups.

In addition, based on descriptive data, all mental health and wellbeing scores improved. The change in wellbeing (WHO-5) and depression (PHQ-9) were statistically significant, indicating that the use of the app could have potentially improved some symptoms related to poor wellbeing. However, as this was not a controlled study it is not possible to make any causal claims or explicitly attribute the findings to the use of the app. Nonetheless, this is the first study to evaluate Paradym and therefore these preliminary findings can be viewed as positive and warrant further investigations. The other outcome measures did not yield statistically significant results. This could be as a result of the small sample size (n=34) or other methodological issues such as the 2-week intervention period. It could be that factors associated with life satisfaction and emotional awareness requires longer usage periods. It was also noted that anxiety levels of the participants were mild during the intervention period, which

at a minimum could indicate that the Paradym app did not induce any unnecessarily high levels of anxiety.

Comparison to other findings

In terms of the demographic profile of the sample, the evaluation captured data from a diverse sample (43.5% belong to minority ethnic groups). This is generally uncommon in mental health app studies [80] and therefore provides valuable insights from usually underserved and underrepresented populations. Although we hoped to collect data from a larger sample, the total sample size and high attrition observed in this evaluation appears to be common when evaluating digital interventions [81]. Despite these challenges, the findings are in line with the existing literature indicating the potential usefulness of apps to support positive mental health and wellbeing. Our findings suggested a significant increase in wellbeing which was also observed in similar studies [16,25]. A decrease in depression symptoms have also been reported in other studies evaluating mental health app usage [82–85]. Similar themes and statistics were also reported for user engagement and acceptability [21,86]. However, in terms of study design, this evaluation obtained both quantitative and qualitative data, including objective engagement data, to fully capture users' feedback about the app. This evaluation also demonstrated that it was possible to conduct an app evaluation within a 6-week period, while previous studies reported slightly longer evaluation periods of 8 weeks [87]) or more [88]. When compared to recent systematic reviews and meta-analyses the effect sizes in our evaluation fell within the previously observed range (e.g. 0.10–0.57) for depression (k=12, [10]; k=18, [89]). Furthermore, for wellbeing, the effect size in this evaluation (0.51) was above the previously observed range (k=5, 0.14 - 0.45) [9].

In terms of the app itself, based on available descriptions for similar apps [22], Paradym, to the best of our knowledge, may be one of the first apps which focuses on improving emotional awareness and emotional wellbeing for the general population through the medium of identifying emotional patterns, and integrates a number of psychological theories. The module-based content is similar to that of other apps [90–92], but the focus on users identifying and selecting emotional patterns to build up an overall profile appears to be unique to Paradym. This is an important finding, as preliminary feedback suggests that some users responded positively to this feature.

Strengths and limitations

The main strength of this evaluation is its ability to collect and analyze data from a diverse sample of app users. At this preliminary stage of the app's evaluation, the findings already suggest some

evidence of promise. These include significant impacts on mental health and wellbeing and no reports of adverse events, in addition to positive feedback on usage and engagement. Another strength is its ability to recruit an acceptable sample size and carry out the evaluation within a 6-week period. This could indicate a need on behalf of the users, and also demonstrate to other app developers and researchers that it is feasible to conduct ongoing evaluations during rapid prototyping phases.

Despite notable strengths, this evaluation also acknowledges some limitations. Obvious limitations include the single arm pre-/post- evaluation design which limits the ability to perform comparative analyses. The moderate sample size (n=34) has limited potential to carry out subgroup analyses (e.g. males). Studies with larger sample sizes and subgroup analysis could provide valuable information for future versions of the app. The high attrition rate (70.4%), whilst in line with results reported by other studies, is a point of concern and the future evaluations of the app would take this into account during the recruitment period. This limitation could be attributed to data collection via a link outside of the app platform since the engagement data showed higher numbers of people (n=91) continuing to use the app without completing the outcome measures. Owing to the anonymous nature of the evaluation, it was not possible to make comparisons between participants who downloaded the app and those who did not. Another limitation is the 2-week intervention period. Although this can be viewed as a strength for this preliminary evaluation, there may be benefits – such as observations about sustained usage and further impact on wellbeing – if a longer app usage period was included. Another limitation lies in the brief responses provided by users for the semi-structured questions. This limited the amount of data available to provide deeper insights during thematic analysis. These limitations and other key learnings will be considered for future evaluations of the app.

Future directions

App upgrades

Participants provided valuable feedback that has already been absorbed to inform upgrades to the app. Based on these findings, the following upgrades will be prioritized before further evaluations of the app are conducted. Firstly, a proportion of participants expressed that they would like the app to be more interactive. A guided process has been added as an adjunct to the app so users can join weekly sessions as part of a group to discuss sections of the app and reflect on concepts of self-awareness with an assigned coach. Second, the feedback indicated that some users forgot about the app and needed reminders. More in-app notifications have been included in newer versions of the app and are expected to further improve levels of engagements [12,18]. Third, based on recent findings for increasing engagement [25,93] and the positive perception of the app's content and functionality,

features such as personalized check-ins will be considered for inclusion in the app to further enhance the users' experience.

Research implications

This evaluation provided a strong foundation for future evaluations of the app. The current recruitment, intervention and data collection period (in total 6 weeks) may be too ambitious to gather a large enough sample size to meaningfully test statistical significance on multiple variables. Based on our findings and the experience of carrying out this evaluation, a total of 12 weeks may be more feasible – allowing at least 1 month for each stage of the study. In this way, participants will have a longer period to engage with the app and therefore complete all modules within the app. The outcome measures used appear to be appropriate with almost 100% completion (for the n=34 users) on all items. Future evaluations of the app could maintain the current data collection tools and include the anxiety measure (i.e., GAD-7) at baseline to ensure it is accurately capturing any direct influence of the app on participants' levels of anxiety. This approach therefore may go beyond the research question addressed in this evaluation and provide data on any relationship the app may have with lowering anxiety. This may prove to be valuable information that could substantiate the findings we highlighted on the relationship between the app and depression scores. It may be just as useful to enhance the evaluation by adding a comparative group such as a blended intervention group and randomizing participants to ensure our findings are not occurring by chance. In the future evaluations, it may be just as important for us to collect qualitative data from participants who drop out to fully capture if their decision to discontinue is based on app usage or lack of time to complete outcome measures, as suggested by other researchers [94].

Practical implications

Given the range of wellbeing apps, and the preliminary positive implications of this evaluation, improving wellbeing through regular use of an emotional wellbeing app that focuses on emotional awareness and knowledge of emotional patterns, could have potential promise as a low-cost approach to increase wellbeing and related outcomes, such as life satisfaction. Sustaining user engagement in apps over time is not straightforward. The results of this evaluation are preliminary due to the moderate sample size across data collection points, however, there may be benefits to apps that draw on lived experience and storytelling, are multi-modal, and promote increased emotional awareness and understanding of emotional patterns. This evaluation recruited a higher than average number of participants that were from minority ethnic backgrounds and further apps and research initiatives ought to appeal to non-majority app users.

Conclusion

This evaluation further highlighted the value of conducting formative research on mental health and wellbeing apps. For example, the findings suggest that users are capable of providing valid feedback that could be used to inform future app upgrades. Users generally engaged with the app and provided favorable feedback regarding the app's content, functionality and underlying principles. Notably, participants in this evaluation experienced statistically significant lower levels of depression scores post the intervention period as well as increased wellbeing. The findings of this evaluation are encouraging and shows positive preliminary evidence that suggests scope for further research with underrepresented groups, such as ethnic minority populations.

Abbreviations

SWLS – Satisfaction with Life Survey

WHO-5 - World Health Organisation Wellbeing Measure

ESAS-R - Emotional Self Awareness Scale Revised

PHQ-9 - The Patient Health Questionnaire

GAD-7 - General Anxiety Disorder Test

Acknowledgements

The authors thank the participants who took the time to participate in this study. We thank Rachel Gomes for support with study design and collecting objective app usage data. We thank Isabella Vainieri, Fiby Saad and Elisa Infanti for discussions in the early stages of the research and Nick Toner for his support when editing early versions of the manuscript.

Conflicts of interest

The authors declare no direct financial gains from carrying out this study. However, it must be noted that CC is the founder and CEO of Paradym Ltd, and ME is employed as the Research Lead. To ensure researcher bias was not introduced, the data analysis was conducted independently by at least 2 members of the researcher team. The other authors have no conflict of interest.

Role of funders

This study was funded by Paradym Ltd. Apart from commissioning a call for evidence in this area, CC had no role in the data collection, data analysis and data interpretation. However, CC was

responsible for the oversight of the project and was involved in the study design and editing the drafts of the manuscript.

Authors' contributions

ME, CC and SL conceptualized the study and contributed to the study design. All authors contributed to the running of the user testing. To reduce the risk of bias in favor of the intervention, SL and RC independently analyzed the data and interpreted the output with input from ME and AM. All authors contributed to the writing and editing of the original manuscript. All authors were involved in the reviewing, editing and the final approval of the manuscript before submission.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Shah SMA, Mohammad D, Qureshi MFH, Abbas MZ, Aleem S. Prevalence, Psychological Responses and Associated Correlates of Depression, Anxiety and Stress in a Global Population, During the Coronavirus Disease (COVID-19) Pandemic. *Community Ment Health J.* 2021;57(1):101–110.
2. Ozamiz-Etxebarria N, Idoiaga Mondragon N, Dosil Santamaría M, Picaza Gorrotxategi M. Corrigendum: Psychological Symptoms During the Two Stages of Lockdown in Response to the COVID-19 Outbreak: An Investigation in a Sample of Citizens in Northern Spain [Internet]. *Front Psychol* . 2020. p. 2116.
3. Holmes EA, O'Connor RC, Perry VH, Tracey I, Wessely S, Arseneault L, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. *The Lancet Psychiatry.* 2020;7(6):547–560.
4. Moreno C, Wykes T, Galderisi S, Nordentoft M, Crossley N, Jones N, et al. How mental health care should change as a consequence of the COVID-19 pandemic. *The Lancet Psychiatry.* Elsevier Ltd; 2020. p. 813–824. PMID: 32682460
5. Alexopoulos A, Hudson J, Otenigbagbe O. The Use of Digital Applications and COVID-19. *Community Ment Health J.* 2020;56.
6. Peek N, Sujan M, Scott P. Digital health and care in pandemic times: impact of COVID-19. *BMJ Heal & Care Informatics.* 2020;27(1):e100166.
7. Aitkin, M., Clancy, B. & Nass D. The Growing Value of Digital Health [Internet]. 2017.
8. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental Health Smartphone Apps: Review and Evidence-Based Recommendations for Future Developments. *JMIR Ment Heal.* 2016;3(1):e7.
9. Gál É, Ștefan S, Cristea IA. The efficacy of mindfulness meditation apps in enhancing users' well-being and mental health related outcomes: a meta-analysis of randomized controlled trials. *J Affect Disord.* 2021;279:131–142.
10. Weisel KK, Fuhrmann LM, Berking M, Baumeister H, Cuijpers P, Ebert DD. Standalone smartphone apps for mental health-a systematic review and meta-analysis. *NPJ Digit Med.* 2019;2:118.
11. Sucala M, Cuijpers P, Muench F, Cardoso R, Soflaur R, Dobrean A, et al. Anxiety: There is an app for that. A systematic review of anxiety apps. *Depress Anxiety.* John Wiley & Sons, Ltd; 2017;34(6):518–525.
12. Wasil, A. R., Venturo-Conerly, K. E., Shingleton, R., & Weisz JR(. A Review of

- Popular Smartphone Apps for Depression and Anxiety: Assessing the Inclusion of Evidence-Based Content. 2019;
13. Chandrashekar P. Do mental health mobile apps work: evidence and recommendations for designing high-efficacy mental health mobile apps. *mHealth*. AME Publishing Company; 2018;4:6.
 14. Musiat P, Goldstone P, Tarrrier N. Understanding the acceptability of e-mental health - attitudes and expectations towards computerised self-help treatments for mental health problems. *BMC Psychiatry*. 2014;14(1):109.
 15. Lyubomirsky S, King L, Diener E. The benefits of frequent positive affect: Does happiness lead to success? *Psychol Bull*. 2005;131(6):803–855. PMID: 16351326
 16. Champion L, Economides M, Chandler C. The efficacy of a brief app-based mindfulness intervention on psychosocial outcomes in healthy adults: A pilot randomised controlled trial. *PLoS One*. Public Library of Science; 2018;13(12). PMID: 30596696
 17. Huberty J, Green J, Glissmann C, Larkey L, Puzia M, Lee C. Efficacy of the Mindfulness Meditation Mobile App “Calm” to Reduce Stress Among College Students: Randomized Controlled Trial. *JMIR Mhealth Uhealth*. 2019;7(6):e14273.
 18. Wasil, A. R., Palermo, E., Lorenzo-Luaces, L., & DeRubeis R. Is there an app for that? A review of popular mental health and wellness apps. *PsyArXiv*. 2021;1–51.
 19. Clarke J, Draper S. Intermittent mindfulness practice can be beneficial, and daily practice can be harmful. An in depth, mixed methods study of the “Calm” app’s (mostly positive) effects. *Internet Interv*. 2020;19:100293.
 20. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M-R, Christensen H. Smartphones for Smarter Delivery of Mental Health Programs: A Systematic Review. *J Med Internet Res*. 2013;15(11):e247.
 21. Haeger JA, Davis CH, Levin ME. Utilizing ACT daily as a self-guided app for clients waiting for services at a college counseling center: A pilot study. *J Am Coll Heal*. Routledge; 2020;
 22. Eisenstadt, M., Infanti, E. & Liverpool S. Characteristics and effectiveness of mobile apps that promote emotion regulation and well-being in the general population: a systematic review protocol. PROSPERO 2020 CRD42020213051 [Internet].
 23. Kenny R, Dooley B, Fitzgerald A. Feasibility of “CopeSmart”: A Telemental Health App for Adolescents. *JMIR Ment Heal*. 2015;2(3):e22.
 24. Neary M, Schueller SM. State of the Field of Mental Health Apps. *Cogn Behav Pract*.

- 2018;25(4):531–537.
25. Huberty J, Green J, Puzia M, Stecher C. Evaluation of Mood Check-in Feature for Participation in Meditation Mobile App Users: Retrospective Longitudinal Analysis. *JMIR Mhealth Uhealth*. 2021;9(4):e27106.
 26. Linardon J F-TM. Attrition and adherence in smartphone-delivered interventions for mental health problems: a systematic and meta-analytic review. *J Consult Clin Psychol*. 2020;88(1):1–13.
 27. Keyes CLM. Mental Health as a Complete State: How the Salutogenic Perspective Completes the Picture. *Bridg Occup Organ Public Heal*. Dordrecht: Springer Netherlands; 2014. p. 179–192.
 28. Ryan RM, Deci EL. On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annu Rev Psychol*. Annual Reviews; 2001;52(1):141–166.
 29. Foundation MH. What is good mental health? [Internet]. 2021.
 30. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ*. 2008;a1655.
 31. Bennett JA. The Consolidated Standards of Reporting Trials (CONSORT): Guidelines for Reporting Randomized Trials. *Nurs Res*. 2005;54(2).
 32. Eisenstadt, A., & Liverpool S. User Testing Service Evaluation. [Internet]. 2021.
 33. Hayes SC, Levin ME, Plumb-Villardaga J, Villatte JL, Pistorello J. Acceptance and commitment therapy and contextual behavioral science: examining the progress of a distinctive model of behavioral and cognitive therapy. *Behav Ther*. 2013;44(2):180–198.
 34. Beck JS. *Cognitive therapy: Basics and beyond*. Cogn Ther Basics beyond. New York, NY, US: Guilford Press; 1995.
 35. Bowlby J. *Attachment and loss*. Vol. 3: Loss, sadness and depression. [Internet]. New York: Basic Books.;
 36. Vignoles VL, Schwartz SJ, Luyckx K. Introduction: Toward an integrative view of identity. *Handb identity theory Res Vols 1 2*. Vignoles, Vivian L.: School of Psychology, University of Sussex, Brighton, United Kingdom, v.l.vignoles@sussex.ac.uk: Springer Science + Business Media; 2011. p. 1–27.
 37. Lineham M. *DBT Skills Training Manual, Second Edition*. 2014.
 38. Sutton A. *Measuring the Effects of Self-Awareness: Construction of the Self-*

- Awareness Outcomes Questionnaire. *Eur J Psychol. PsychOpen*; 2016;12(4):645–658.
39. Gross JJ. Emotion regulation: personality processes and individual differences. In: O. P. John, R. W. Robins and LAP, editor. *Handb Personal Theory Res.* New York; 2008. p. 701–722.
 40. Campbell L, Stanton SC. Adult attachment and trust in romantic relationships. *Curr Opin Psychol.* 2019;25:148–151.
 41. Davis CH, Gaudiano BA, McHugh L, Levin ME. Integrating storytelling into the theory and practice of contextual behavioral science. *J Context Behav Sci.* 2021;20:155–162.
 42. Mazzocco PJ, Green MC, Sasota JA, Jones NW. This Story Is Not for Everyone: Transportability and Narrative Persuasion. *Soc Psychol Personal Sci.* SAGE Publications Inc; 2010;1(4):361–368.
 43. Murphy ST, Frank LB, Moran MB, Patnoe-Woodley P. Involved, Transported, or Emotional? Exploring the Determinants of Change in Knowledge, Attitudes, and Behavior in Entertainment-Education. *J Commun.* 2011;61(3):407–431.
 44. Bottorff JL, Sarbit G, Oliffe JL, Kelly MT, Lohan M, Stolp S, et al. “If I Were Nick”: Men’s Responses to an Interactive Video Drama Series to Support Smoking Cessation. *J Med Internet Res.* 2015;17(8):e190.
 45. Young JE. Cognitive therapy for personality disorders: A schema-focused approach. *Cogn Ther Personal Disord A schema-focused approach.* Sarasota, FL, England: Professional Resource Exchange, Inc; 1990.
 46. Zarbo C, Tasca GA, Cattafi F, Compare A. Integrative Psychotherapy Works. *Front Psychol. Frontiers Media S.A.;* 2016;6:2021.
 47. Ebert DD, Zarski A-C, Christensen H, Stikkelbroek Y, Cuijpers P, Berking M, et al. Internet and Computer-Based Cognitive Behavioral Therapy for Anxiety and Depression in Youth: A Meta-Analysis of Randomized Controlled Outcome Trials. *PLoS One. Public Library of Science;* 2015;10(3):e0119895.
 48. Feldman G, Dunn E, Stemke C, Bell K, Greeson J. Mindfulness and rumination as predictors of persistence with a distress tolerance task. *Pers Individ Dif.* 2014;56:154–158.
 49. Brinker JK, Chin ZH, Wilkinson R. Ruminative thinking style and the MMPI-2-RF. *Pers Individ Dif.* 2014;66:102–105.
 50. Silva Almodovar A, Surve S, Axon DR, Cooper D, Nahata MC. Self-Directed Engagement with a Mobile App (Sinaspri) and Its Effects on Confidence in Coping

- Skills, Depression, and Anxiety: Retrospective Longitudinal Study. *JMIR Mhealth Uhealth*. 2018;6(3):e64.
51. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. *Evid Based Ment Heal*. 2018;21(3):116 LP – 119.
 52. Morris M. *Motivating Change with Mobile: Seven Guidelines*.
 53. Staehr Johansen K. The use of well-being measures in primary health care - the DepCare project; in World Health Organization, Regional Office for Europe: *Well-Being Measures in Primary Health Care - the DepCare Project*. Geneva,; 1998.
 54. Diener E, Emmons RA, Larsen RJ, Griffin S. *The Satisfaction With Life Scale*. *J Pers Assess*. Routledge; 1985;49(1):71–75.
 55. Kauer SD. Emotional self-awareness and depressive symptoms: an investigation of an early intervention mobile phone self-monitoring program for adolescents. [Internet]. The University of Melbourne; 2012.
 56. Spitzer RL, Kroenke K, Williams JBW, Group and the PHQPCS. Validation and Utility of a Self-report Version of PRIME-MD The PHQ Primary Care Study. *JAMA*. 1999;282(18):1737–1744.
 57. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *Arch Intern Med*. 2006;166(10):1092–1097.
 58. Brooke J. SUS: A “Quick and Dirty” Usability Scale. *Usability Eval Ind*. 1996.
 59. Kern A, Hong V, Song J, Lipson SK, Eisenberg D. Mental health apps in a college setting: openness, usage, and attitudes. *mHealth*. AME Publishing Company; 2018;4:20.
 60. Bech P. *Rating Scales for Psychopathology, Health Status and Quality of Life* [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 1993.
 61. Krieger T, Zimmermann J, Huffziger S, Ubl B, Diener C, Kuehner C, et al. Measuring depression with a well-being index: Further evidence for the validity of the WHO Well-Being Index (WHO-5) as a measure of the severity of depression. *J Affect Disord*. 2014;156:240–244.
 62. van Beuningen J. The satisfaction with life scale examining construct validity. *Discuss Pap*. 2012;(201209):1–23.
 63. Vassar M. A note on the score reliability for the Satisfaction With Life Scale: an RG study. *Soc Indic Res*. 2008;86(1):47–57.

64. Bakker D, Rickard N. Engagement in mobile phone app for self-monitoring of emotional wellbeing predicts changes in mental health: MoodPrism. *J Affect Disord.* 2018;227:432–442.
65. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9. *J Gen Intern Med.* 2001;16(9):606–613.
66. Kocalevent R-D, Hinz A, Brähler E. Standardization of the depression screener Patient Health Questionnaire (PHQ-9) in the general population. *Gen Hosp Psychiatry.* 2013;35(5):551–555.
67. Sousa T V, Viveiros V, Chai M V, Vicente FL, Jesus G, Carnot MJ, et al. Reliability and validity of the Portuguese version of the Generalized Anxiety Disorder (GAD-7) scale. *Health Qual Life Outcomes.* 2015;13(1):50.
68. Lewis JR. The System Usability Scale: Past, Present, and Future. *Int J Human–Computer Interact.* Taylor & Francis; 2018;34(7):577–590.
69. Lewis JR, Brown J, Mayes DK. Psychometric Evaluation of the EMO and the SUS in the Context of a Large-Sample Unmoderated Usability Study. *Int J Human–Computer Interact.* Taylor & Francis; 2015;31(8):545–553.
70. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol.* Routledge; 2006;3(2):77–101.
71. Koocher, G.P. & Keith-Spiegel, P. (2008). *Ethics in Psychology and the Mental Health Professions: Standards and Cases* (3rd ed.). New York: Oxford University Press. Reviewed by Eric K. Willmarth, PhD, Saybrook Graduate School and Research Center, San Fran. *Am J Clin Hypn.* Routledge; 2009;52(1):74–75.
72. Ly KH, Dahl JA, Carlbring P, Andersson G. Development and initial evaluation of a smartphone application based on acceptance and commitment therapy. *Springerplus.* SpringerOpen; 2012;1(1):1–11.
73. Morris ME, Kathawala Q, Leen TK, Gorenstein EE, Guilak F, Labhard M, et al. Mobile Therapy: Case Study Evaluations of a Cell Phone Application for Emotional Self-Awareness. *J Med Internet Res.* 2010;12(2):e10.
74. Ahtinen A, Mattila E, Väikkynen P, Kaipainen K, Vanhala T, Ermes M, et al. Mobile Mental Wellness Training for Stress Management: Feasibility and Design Implications Based on a One-Month Field Study. *JMIR mhealth uhealth.* JMIR Publications Inc.; 2013;1(2):e11.
75. Ly KH, Ly AM, Andersson G. A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods. *Internet Interv.* Elsevier B.V.;

- 2017;10:39–46.
76. Lui JHL, Marcus DK, Barry CT. Evidence-based apps? A review of mental health mobile applications in a psychotherapy context. *Prof Psychol Res Pract*. 2017;48(3):199–210.
 77. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and Promoting Effective Engagement With Digital Behavior Change Interventions. *Am J Prev Med*. 2016;51(5):833–842.
 78. Liverpool S, Mota CP, Sales CMD, Čuš A, Carletto S, Hancheva C, et al. Engaging Children and Young People in Digital Mental Health Interventions: Systematic Review of Modes of Delivery, Facilitators, and Barriers. *J Med Internet Res*. 2020;22(6):e16317.
 79. Goodyear VA, Armour KM, Wood H. Young people learning about health: the role of apps and wearable devices. *Learn Media Technol*. Routledge; 2019;44(2):193–210.
 80. Eisenstadt, A, Liverpool, S, Infanti, E., Cuivat, R., & Carlsson C. Mobile Apps that Promote Emotion Regulation, Positive Mental Health and Well-being in the General Population: Systematic Review and Meta-analysis (Submitted for Publication). *JMIR Ment Heal*. 2021;
 81. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of Attrition and Dropout in App-Based Interventions for Chronic Disease: Systematic Review and Meta-Analysis. *J Med Internet Res*. JMIR Publications; 2020;22(9):e20283–e20283.
 82. Ha SW, Kim J. Designing a Scalable, Accessible, and Effective Mobile App Based Solution for Common Mental Health Problems. *Int J Human–Computer Interact*. Taylor & Francis; 2020;36(14):1354–1367.
 83. Howells A, Ivtzan I, Eiroa-Orosa FJ. Putting the ‘app’ in Happiness: A Randomised Controlled Trial of a Smartphone-Based Mindfulness Intervention to Enhance Wellbeing. *J Happiness Stud*. Springer Netherlands; 2016;17(1):163–185.
 84. Ponzo S, Morelli D, Kawadler JM, Hemmings NR, Bird G, Plans D. Efficacy of the digital therapeutic mobile app biobase to reduce stress and improve mental well-being among university students: Randomized controlled trial. *JMIR mHealth uHealth*. JMIR Publications Inc.; 2020;8(4). PMID: 31926063
 85. Roy MJ, Costanzo ME, Highland KB, Olsen C, Clayborne D, Law W. An App a Day Keeps the Doctor Away: Guided Education and Training via Smartphones in Subthreshold Post Traumatic Stress Disorder. *Cyberpsychology, Behav Soc Netw*.

- Mary Ann Liebert Inc.; 2017;20(8):470–478. PMID: 28737954
86. Deady M, Johnston D, Milne D, Glozier N, Peters D, Calvo R, et al. Preliminary Effectiveness of a Smartphone App to Reduce Depressive Symptoms in the Workplace: Feasibility and Acceptability Study. *JMIR Mhealth Uhealth*. 2018;6(12):e11661.
 87. Bostock S, Crosswell AD, Prather AA, Steptoe A. Mindfulness on-the-go: Effects of a mindfulness meditation app on work stress and well-being. *J Occup Health Psychol*. American Psychological Association Inc.; 2019;24(1):127–138. PMID: 29723001
 88. Firestone R, Cheng S, Dalhousie S, Hughes E, Funaki T, Henry A, et al. Exploring Pasii ka wellbeing: ndings from a large cluster randomised controlled trial of a mobile health intervention programme. *NZMJ*. 2020;133:1524.
 89. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry*. John Wiley & Sons, Ltd; 2017;16(3):287–298.
 90. Coelho CC, Tobo PR, Lacerda SS, Lima AH, Barrichello CRC, Amaro E, et al. A new mental health mobile app for well-being and stress reduction in working women: Randomized controlled trial. *J Med Internet Res*. JMIR Publications Inc.; 2019;21(11). PMID: 31697244
 91. Walsh KM, Saab BJ, Farb NA. Effects of a Mindfulness Meditation App on Subjective Well-Being: Active Randomized Controlled Trial and Experience Sampling Study. *JMIR Ment Heal*. 2019;6(1):e10844.
 92. Weber S, Lorenz C, Hemmings N. Improving Stress and Positive Mental Health at Work via an App-Based Intervention: A Large-Scale Multi-Center Randomized Control Trial. *Front Psychol*. Frontiers Media S.A.; 2019;10.
 93. Cheng VWS, Davenport T, Johnson D, Vella K, Hickie IB. Gamification in Apps and Technologies for Improving Mental Health and Well-Being: Systematic Review. *JMIR Ment Heal*. 2019;6(6):e13717.
 94. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: A systematic review and meta-analysis. *J Affect Disord*. 2020;263:413–419.