

RESEARCH

Open Access



# RegulEm, an unified protocol based-app for the treatment of emotional disorders: a parallel mixed methods usability and quality study

Laura Martínez-García<sup>1,2</sup>, Alba Fadrique-Jiménez<sup>1</sup>, Vanesa-Ferreres -Galán<sup>3</sup>, Cristina Robert Flors<sup>4</sup> and Jorge Osma<sup>1,2\*</sup>

## Abstract

**Background** Interest in mental health smartphone applications has grown in recent years. Despite their effectiveness and advantages, special attention needs to be paid to two aspects to ensure app engagement: to include patients and professionals in their design and to guarantee their usability. The aim of this study was to analyse the perceived usability and quality of the preliminary version of RegulEm, an app based in the Unified Protocol, as part of the second stage of the app development.

**Methods** A parallel mixed methods study was used with 7 professionals and 4 users who were previously involved in the first stage of the development of the app. MARS, uMARS and SUS scales were used, and two focus groups were conducted. Descriptive statistical analysis and a thematic content analysis were performed in order to gather as much information as possible on RegulEm's usability and quality as well as suggestions for improvement.

**Results** RegulEm's usability was perceived through the SUS scale scores as good by users (75 points) and excellent by professionals (84.64 points), while its quality was perceived through the uMARS and MARS scales as good by both groups, with 4 and 4.14 points out of 5. Different areas regarding RegulEm's usability and suggestions for improvement were identified in both focus groups and 20% of the suggestions proposed were implemented in the refined version of RegulEm.

**Conclusion** RegulEm's usability and quality were perceived as good by users and professionals and different identified areas have contributed to its refinement. This study provides a more complete picture of RegulEm's usability and quality prior analysing its effectiveness, implementation and cost-effectiveness in Spanish public mental health units.

**Keywords** Unified protocol, Smartphone app, Usability, Quality, Mixed methods, M-health

\*Correspondence:

Jorge Osma  
osma@unizar.es

<sup>1</sup>Department of Psychology and Sociology, Faculty of Social and Human Sciences, University of Zaragoza, C/Atarazanas, 4, Teruel 44003, Spain

<sup>2</sup>Health Research Institute of Aragón, Biomedical Research Center of Aragón (CIBA), Av. San Juan Bosco, 13, Zaragoza 50009, Spain

<sup>3</sup>Mental Health Unit of the Hospital Comarcal of Vinaròs, Av. Gil d'Atrosillo, s/n, Vinaròs, Castellón 12500, Spain

<sup>4</sup>Mental Health Unit of the Font de Sant Lluís Health Center, C/Arabista Ambrosio Huici, 30, Valencia 46013, Spain



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

## Background

Interest in smartphone applications (mHealth apps) for psychological care has grown in recent years [1] due to increasing digitalization within medicine, including mobile health applications [2]. This growing interest could be driven by the fact that mHealth apps are emerging as a viable tool to overcome barriers that mental health care faces [3].

Barriers include the high cost of treatments and long waiting lists, among others [4]. At the same time, the prevalence of mental health disorders, and especially Emotional Disorders (EDs; nomenclature that includes anxiety, depression and related disorders) [5] is high, with this group of disorders being the most prevalent in general population worldwide [6]. This also applies to Spain, where approximately 21.6% of people fulfil criteria for an anxiety disorder and 18.7% for a depressive disorder [7] and patients face long waiting lists [8] and long periods between sessions [9] in public mental health units. Apps to provide psychological treatment could help to overcome these barriers by acting as a complement to face-to-face therapy, reducing the workload of the professionals and boosting the effectiveness of treatment [10]. In turn, treatment through this type of apps may seem more accessible and present in patients' daily lives [11].

While evidence is emerging on the efficacy of these apps for the improvement of EDs [1, 12], it is still inconclusive [13]. In this regard, a systematic review and meta-analysis does not recommend to use apps as standalone psychological interventions but to combine them with face-to-face therapy ("blended care"; BC) [14]. BC offers additional advantages to those of app based stand-alone or classic face-to-face psychological interventions, such as improving the transfer of content learned into daily life [15]. In addition, BC could help professionals save time and maintain or even increase face-to-face treatment outcomes [16].

BC can be implemented in different manners, such as integrated in treatment or in a sequential manner [16]. It is known that the majority of patients in Spanish public mental health units prefer individual therapy (85.4%) to group (14.2%) and online (0.4%) [17]. Therefore, BC may be a promising option to improve the state of mental health care in Spanish public mental health units, in an integrated way in which face-to-face sessions are combined with an app.

Despite the efficacy and benefits of BC, evidence from literature suggests a shortage of high quality mHealth apps [18]. Consistent with this, a scoping review identified various problems for the use of these apps, involving issues concerning validity and usability, among others [19]. In this sense, not involving users in the development process and poor usability have been outlined as two of the reasons for weak engagement with these mHealth

apps [20]. In turn, theories of technology acceptance, such as the Technology Acceptance Model (TAM), provide a promising framework for examining and ensuring the acceptability of these apps, suggesting technology characteristics as predictors of technology acceptance. In the case of the TAM, technology acceptance would be explained by the user's attitude towards use, determined by perceived usefulness and ease of use [21]. Therefore, special attention needs to be paid to these aspects to ensure optimal app engagement.

On the one hand, it is necessary to ensure that mHealth apps are designed to meet the needs of end users before they are used as part of the intervention [22]. In the Participatory Design approach, the end user participates actively in the design development process, being a key component of [23]. In turn, it has been suggested that user participation in the design phase of the app may prevent future problems related to its use that may arise in clinical practice [24].

On the other hand, the prevention of quality issues in mHealth apps is important, as these can lead to limited efficacy or potentially harm the user [25]. Therefore, it is necessary to assess the quality of these type of apps from the early stages of their development [26]. In this regard, despite various attempts to develop general criteria for defining and evaluating the quality of mHealth apps, this implies a challenge due to the wide range of functionalities and areas of application of these apps, as well as their constant evolution [26]. In an attempt, Stoyanov et al. [27] conducted a literature review of the existing criteria for assessing the quality of mHealth apps, with the aim of developing a reliable and objective scale to measure the degree to which these apps comply with quality criteria. Thus, they established participation, functionality, aesthetics and quality of information, as well as subjective quality as quality criteria for these apps. As a result of this review, they developed the Mobile Application Rating Scale (MARS), one of the most widely used tools for evaluating the quality of mHealth apps [28].

In addition, usability can be defined in terms of the degree to which a product can be utilized by specific users to reach specific goals in an effective and efficient manner, while at the same time facilitating user satisfaction in a specific context of usage [29]. Given that lack of usability can be a significant obstacle to the adoption of mhealth apps [30], another method is to apply app assessment procedures to ensure good usability [22]. Usability has been identified as a core component of best practice in apps development [31]. Thus, its evaluation is motivated by aspects such as reporting on the redesign and refinement of the interface, among others [32]. In addition, evidence indicates that it is not only important to involve end-users, but also health professionals, in the

design of mental health apps [33] and in their usability evaluation [32].

When evaluating mental health apps' usability, a mixed-methods approach is recommended [34, 35]. In this regard, the Medical Research Council framework for developing and evaluating complex interventions highlights the importance of integrating quantitative and qualitative data at various stages, from development to implementation [36]. Furthermore, considering the aspects of app integration in routine practice, as emphasized by the Normalization Process Theory [37], a mixed methods approach allows to capture the complexity of how users interact with and adapt to the application in real-world settings.

In this sense, our team has developed an app based on a transdiagnostic Cognitive Behavioural Therapy (CBT) intervention, the Unified Protocol for Transdiagnostic Treatment of Emotional Disorders (UP) [38]. Transdiagnostic interventions such as UP address common mechanisms underpinning a wide range of disorders [39], enabling treatments to be designed for a wider group of disorders instead of for specific disorders, thus making it possible to treat individuals with comorbidity [40]. Given that EDs are characterized by difficulties in emotion regulation [41], the goal of the UP is to train in adaptive emotional regulation skills through eight modules [38]. Regarding its efficacy, to date, six systematic reviews have been published, five of them meta-analyses [42–47], demonstrating its utility with statistically superior effects to waiting list conditions and comparable to or lightly superior to those obtained by disorder-specific

CBT [45, 47]. A preliminary version of the app, named RegulEm, has been developed as a result of the first stage of a participatory process involving users and professionals of Spanish public mental health units familiar with the UP [48].

The aim of the current study was to analyse the perceived usability and quality of the preliminary version of RegulEm, as the need for this type of testing prior to assessing apps effectiveness has been emphasized [49]. A parallel mixed methods study was used with users and professionals who were also involved in the first stage of the participatory process. In this way, we seek to ensure that the app is appropriately designed and targeted to the end-users' needs. The information collected will be used to refine the preliminary version of the app that will be included in a pilot study and a later randomized controlled trial (RCT) that will analyse the effectiveness, implementation, and cost-effectiveness of UP in blended format in Spanish public mental health units.

## Methods

### Participants

The present study used a convenience sample which consisted of professionals and users who had collaborated in a focus group study prior to the development of the app [48]. Both groups, professionals and users, were familiar with the UP because they had applied or received the UP in group format within a multicenter RCT focused on analysing its effectiveness and cost-effectiveness in Spanish public mental health units [48].

In a first attempt, 7 professionals and 9 users were contacted to participate via email. Of the users, 3 failed to respond and 2 were unable to attend the sessions due to scheduling difficulties. Consequently, a total of 11 participants were enrolled in this study. Sociodemographic characteristics of professionals and users are depicted in Table 1.

### Measures

Sociodemographic information was collected through questions included in a Google Forms survey of users and professionals in order to obtain a more detailed description of the characteristics of the sample.

System Usability Scale (SUS) [50] in its Spanish version [51] was used to evaluate the usability of the app. The SUS scale is a widely used 10-item questionnaire scored on a 5-point Likert scale ranging from 1 (strong disagreement) to 5 (strong agreement).

Mobile Application Rating Scale (MARS) [27] and its user version (uMARS) [52], both in their Spanish versions [53, 54], were used to evaluate the quality of the app. The MARS provides an objective and subjective score of app quality and four scores from four objective quality subscales (engagement, functionality, aesthetics

**Table 1** Sociodemographic characteristics of participants (n = 11)

Variable	Users (n = 4)	Professionals (n = 7)
	Mean (SD)	Mean (SD)
Age	43.25 (5.38)	47.86 (9.58)
	Proportion (%)	Proportion (%)
Gender		
Men	-	14.3 (n = 1)
Women	100 (n = 4)	85.7 (n = 6)
User's ED diagnosis*		
Generalized anxiety disorder	25 (n = 1)	-
Adjustment disorder	25 (n = 1)	-
Obsessive-compulsive disorder	25 (n = 1)	-
Non-specified anxiety disorder	25 (n = 1)	-
	Mean (SD)	Mean (SD)
Professional's experience		
Years of clinical experience	-	16.57 (10.11)
Hours of training received in UP	-	42.00 (30.27)
Patients to whom UP has been applied	-	53.57 (31.45)

\*Users were no longer meeting diagnostic criteria and had been discharged at the time of the study

and information) across 23 items rated on a 5-point Likert scale ranging from 1 (“poor”) to 5 (“excellent”), except 5 items (14, 15, 16, 17 and 19) which also include a “not applicable” option. Furthermore, it provides a perceived impact score through 6 items rated on a 5-point Likert scale ranging from 1 (“strongly disagree”) to 5 (“strongly agree”). The uMARS also provides an objective and subjective score of app quality and four scores from the same four objective quality subscales across 20 items rated on a 5-point Likert scale ranging from 1 (“poor”) to 5 (“excellent”), except 4 items [13–16] which also include a “not applicable” option. In turn, it also provides a perceived impact score from 6 items rated in the same manner as in the MARS.

Finally, a semi-structured interview guide developed for this purpose by the focus group moderator was used to collect as much qualitative information as possible regarding different specific aspects of the app’s perceived usability (entertainment, ease of use, content, aesthetics and graphics, facilitators and barriers). The moderator of the focus groups elaborated the questions taking into account the general research aim of the study [55]. To this end, and with the aim of facilitating the integration of quantitative and quantitative results, the interview guide questions were prepared taking into account the items of the quantitative questionnaires used [27, 50–54], including questions on similar issues (entertainment, ease of use, engagement, aesthetics) and expanding with questions on aspects other than those collected in the quantitative methods (app facilitators and barriers). A first draft of the interview guide was prepared, reviewed and approved by the entire research team. The guide interview can be found in Table 2.

**Table 2** Questions from the focus group interviews with professionals and users

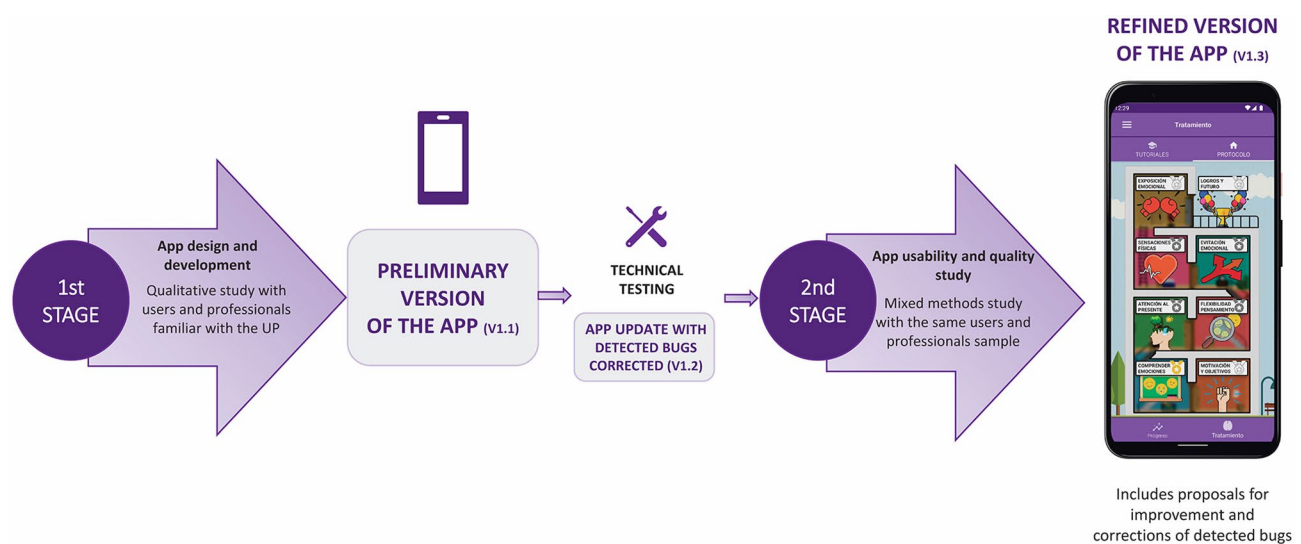
- 1) Have you found RegulEm an entertaining app?
- 2) In general, did you find RegulEm easy to use?
- 3) Do you consider that RegulEm has provided you with information of interest?
- 4) How would you rate the aesthetics and graphic features of RegulEm?
- 5) In general, to what extent do you consider that the different components included in RegulEm (e.g., videos, audios, exercises, True and False exercises, testimonials, emergency button and weekly assessments of emotions and motivation) have been useful to you?
- 6) What aspects would you highlight as facilitators when using RegulEm? That is, what things or aspects of the app do you consider could help when carrying out the intervention (both related to the content and the design of the app)?
- 7) What aspects would you highlight as barriers when using RegulEm? That is, what things or aspects of the app do you consider might not help or hinder you when carrying out the intervention (both related to the content and the design of the app)?

### Procedure

This study comprises the second stage of the participatory design and development process of RegulEm (different stages of the process can be seen in Fig. 1).

Regarding study design, in the present work a convergent parallel mixed methods study was used. In this mixed methods design, quantitative and qualitative information is collected and analyzed simultaneously but independently, and the results obtained from both approaches are then integrated for interpretation [56]. To ensure adequate reporting of the information, the Good Reporting of A Mixed Methods Study (GRAMMS) guidelines [57] for mixed methods studies have been followed (see Additional file 1).

The study was performed under the approval of the ethics committee of General University Hospital of Castellón and in line with the principles of the Declaration



**Fig. 1** RegulEm development process



of Helsinki. All participants accepted and signed the informed consent.

All participants were informed about the aim of the study and received a file of the preliminary version of the app. They were instructed to use the app for two weeks before participating in the focus groups to guarantee that they had enough time to review the modules of which the app is composed [48]. In this regard, RegulEm is based on the UP-patient manual [58], which is aimed at training in adaptive emotional regulation skills through 8 modules: [1] Setting goals and maintaining motivation; [2] Understanding your emotions; [3] Mindful emotion awareness [4] Cognitive flexibility; [5] Countering emotional behaviors; [6] Understanding and confronting physical sensations; [7] Emotion exposures, and [8] Recognizing accomplishments and looking to your future. Each module of the 8 in RegulEm mirrors the content of the corresponding UP manual module and follows a structure of content presentation, comprehension evaluation, exercises, and conclusion. The content is delivered through videos, and comprehension of the content is assessed with true/false questions that provide feedback. Regarding exercises, they are designed as instant message conversations. After completing a module, users can access the module again to review the content and practice additional exercises. Finally, the [conclusion](#) section summarizes and reinforces the module's content. For a more detailed presentation of the functionalities included in RegulEm see [48].

A total of two focus groups were conducted, each lasting an hour and a half. Both focus group were carried out online via Cisco's WebEx platform. Prior to the focus groups, all participants completed sociodemographic information as well as the SUS [50, 51] and MARS [27, 53], for professionals, or uMARS [52, 54], for users, scales through a Google Forms survey.

To conduct both focus groups, we followed an semi-structured interview guide developed for this purpose (see [Table 2](#)). They were moderated by the principal investigator of the team, a senior researcher who has prior experience in the implementation of focus groups. Only one observer was present. In turn, the focus group moderator adopted a neutral interviewing style, and he formulated the questions in an open-ended manner to stimulate spontaneous answers and discussion among participants. Both focus groups were recorded to facilitate subsequent verbatim transcription.

#### Data analysis

The sociodemographic information and quantitative information from the SUS, MARS and uMARS was analyzed through descriptive statistical analysis using SPSS software [59]. SUS results were presented separately for the two groups and both total scores and 10-item means

and standard deviations were reported. Regarding MARS and uMARS, means and standard deviations of app overall objective quality score, the four objective quality subscales and total subjective quality score and its 4 related items were presented. Also, both means and standard deviations of the 6 items related to perceived impact and the total score were reported.

The report of the qualitative part of this study was conducted according to the Consolidated Criteria for Reporting Qualitative Research (COREQ) [60] (see [Additional file 2](#)). Qualitative data was analyzed using the MAXQDA program [61]. First, recordings of the focus groups were transcribed verbatim. After, in order to determine emergent categories of analysis derived from the focus groups data, a thematic content analysis was conducted [62]. This method enabled the creation of a hierarchical coding schedule that facilitated the identification of general categories, including more specific subcategories and their corresponding areas, from the focus groups information. In this sense, using an inductive approach, a coding system was developed in which two members of the research team (PhD students both trained in qualitative analysis), grouped the main ideas extracted from the transcription of the focus groups into "areas." These areas consist of the ideas that were most repeated during the focus groups and are textual examples that emerged in the focus groups. Once these "Areas" were created, they were grouped into "Subcategories", including those areas that shared common characteristics, generating a higher order classification. Finally, these subcategories were grouped into main "Categories" based on the main information we wanted to obtain through the focus groups. In order to ensure inter-rater reliability, two researchers worked independently to extract the categories, subcategories, and corresponding areas. The two researchers then compared their code systems and if any discrepancies arose between their findings, a third researcher (senior researcher with experience in qualitative analysis) was consulted. Thus, the codes that did not match were compared, refined and reorganized to reach a consensus and obtain a first version of the code system. This first version was shared with the third researcher, which led to a further refinement and reorganization of the code system resulting in its final version. Finally, a Cohen's Kappa reliability analysis was conducted between the initial data extraction and the final version in order to evaluate the reliability of the qualitative analyses.

The qualitative information from the focus groups and the quantitative information from the questionnaires will be analysed independently and will be integrated and reported in the [discussion](#) section. The integration of the results obtained through both approaches will be performed in a manner that will explain the quantitative results from a qualitative approach [63], i.e., the

**Table 3** Descriptive statistics for SUS across professionals and users

SUS items	Professionals M <sup>A</sup> (SD <sup>B</sup> )	Users M <sup>A</sup> (SD <sup>B</sup> )
I think I would like to use the app frequently	4.43 (0.53)	4.25 (0.96)
I found the app to be unnecessarily complex	1.57 (0.53)	2.25 (0.96)
I thought the app was easy to use	4.29 (0.49)	3.25 (0.96)
I think that I would need support of a technical person to be able to use the app	1.57 (0.79)	1.00 (0.00)
I found the various functions in the app were well integrated	4.14 (0.90)	3.75 (1.26)
I thought there was too much inconsistency in the app	1.57 (1.13)	1.75 (0.96)
I would imagine that most people would learn to use the app very quickly	4.00 (0.82)	3.00 (1.15)
I found the app very cumbersome to use	1.71 (0.49)	1.50 (0.58)
I felt very confident using the app	4.57 (0.53)	3.75 (0.96)
I needed to learn a lot of things before I could get going with the app	1.14 (0.38)	1.50 (0.58)
SUS total score	84.64 (8.60)	75.00 (7.36)

Note. SUS = System Usability Scale; M<sup>A</sup> = mean; SD<sup>B</sup> = standard deviation; Responses were scored on a 5-point Likert scale ranging from 1 = strong disagreement to 5 = strong agreement

**Table 4** Descriptive statistics for MARS across professionals (n = 7)

MARS	M <sup>A</sup> (SD) <sup>B</sup>	Range
<b>App overall objective quality</b>	4.14 (0.31)	3.68–4.43
Engagement	3.83 (0.55)	2.80–4.20
Functionality	4.14 (0.24)	3.75–4.50
Aesthetics	4.14 (0.42)	3.67–4.67
Information	4.43 (0.35)	4.00–4.86
<b>Subjective quality</b>	3.40 (0.28)	3.50–4.25
Would you recommend this app to people who might benefit from it?	4.29 (0.48)	4.00–5.00
How many times do you think you would use this app in the next 12 months if it was relevant to you?	4.00 (0.00)	4.00–4.00
Would you pay for this app?	3.14 (0.69)	2.00–4.00
What is your overall (star) rating of the app?	3.86 (0.69)	3.00–5.00
<b>Perceived impact</b>	4.23 (0.48)	3.60–5.00
<i>Awareness</i> – This app is likely to increase awareness of the importance of addressing intense emotions more effectively	4.57 (0.53)	4.00–5.00
<i>Knowledge</i> – This app is likely to increase knowledge/understanding of how to deal with intense emotions more effectively	4.71 (0.76)	3.00–5.00
<i>Attitudes</i> – This app is likely to change attitudes toward addressing intense emotions more effectively	4.29 (0.48)	4.00–5.00
<i>Intention to change</i> – This app is likely to increase intentions/motivation to address intense emotions more effectively	3.86 (0.69)	3.00–5.00
<i>Help seeking</i> – Use of this app is likely to encourage further help to address intense emotions more effectively (if it's required)	4.43 (0.53)	4.00–5.00
<i>Behaviour change</i> – Use of this app is likely increase addressing emotions more effectively	3.86 (0.69)	3.00–5.00

Note MARS = Mobile Application Rating Scale; M<sup>A</sup> = mean; SD<sup>B</sup> = standard deviation; Objective quality subscales and total objective quality score: 1 = Poor and 5 = Excellent; 4 subjective quality items and total subjective quality score: 1 = Not at all / None / No / 1\* / Poor and 5 = Definitely / >50 / Yes / 5\* / Excellent; 6 perceived impact items and total perceived impact score: 1 = Strongly disagree and 5 = Strongly agree

qualitative data will be integrated with the quantitative data in order to understand the latter in greater detail.

## Results

### Usability and quality of the app

#### Usability assessment

As can be seen in Table 3, the professionals tended to report higher usability results than those reported by users, with a difference of almost 10 points above them in the SUS total score. However, the usability of the app could be rated as “excellent” by professionals and “good” by users, according to the classification suggested by Bangor et al. [64], with means of SUS total score of 84.64 ± 08.60 and 75.00 ± 7.36 respectively. From the analysis of each of the items, it can be observed that in the case of professionals the lowest score corresponded to “I

needed to learn a lot of things before I could get going with the app”, while the highest score was for “I felt very confident using the app”. For users, the lowest score was “I think that I would need support of a technical person to be able to use the app” and the highest score was for “I think I would like to use the app frequently”. For further detail, means and standard deviations of the answers to the items and the total usability score can be observed in Table 3.

#### Quality assessment

As can be observed in Tables 4 and 5, the overall mean score for the objective quality of the app was similar in both groups but higher among professionals (4.14 ± 0.31 compared to 4.00 ± 0.27 for users) and can be interpreted

**Table 5** Descriptive statistics for uMARS across users ( $n=4$ )

uMARS	$M^A$ (SD) <sup>B</sup>	Range
<b>App overall objective quality</b>	4.00 (0.27)	3.73–4.24
Engagement	3.75 (0.34)	3.40–4.20
Functionality	3.80 (0.28)	3.60–4.20
Aesthetics	3.92 (0.74)	3.00–4.67
Information	4.56 (0.31)	4.25–5.00
<b>Subjective quality</b>	3.81 (0.62)	3.00–4.50
Would you recommend this app to people who might benefit from it?	4.75 (0.50)	4.00–5.00
How many times do you think you would use this app in the next 12 months if it was relevant to you?	4.25 (0.50)	4.00–5.00
Would you pay for this app?	2.25 (1.50)	1.00–4.00
What is your overall (star) rating of the app?	4.00 (1.41)	2.00–5.00
<b>Perceived impact</b>	4.67 (0.47)	4.00–5.00
<i>Awareness</i> – This app has increased my awareness of the importance of addressing intense emotions more effectively	4.75 (0.50)	4.00–5.00
<i>Knowledge</i> – This app has increased my knowledge/understanding of how to deal with intense emotions more effectively	5.00 (0.00)	5.00–5.00
<i>Attitudes</i> – The app has changed my attitudes toward addressing intense emotions more effectively	4.75 (0.50)	4.00–5.00
<i>Intention to change</i> – The app has increased my intentions/motivation to address intense emotions more effectively	4.75 (0.50)	4.00–5.00
<i>Help seeking</i> – This app would encourage me to seek further help to address intense emotions more effectively (if I needed it)	4.25 (0.96)	3.00–5.00
<i>Behaviour change</i> – Use of this app will increase addressing my emotions more effectively	4.50 (0.58)	4.00–5.00

Note uMARS=Mobile Application Rating Scale: user version;  $M^A$  = mean;  $SD^B$  = standard deviation; Objective quality subscales and total objective quality score: 1=Poor and 5=Excellent; 4 subjective quality items and total subjective quality score: 1=Not at all / None / No / 1\* / Poor and 5=Definitely / >50 / Yes / 5\* / Excellent; 6 perceived impact items and total perceived impact score: 1=Strongly disagree and 5=Strongly agree

as “good quality”, with 5 being the maximum rating possible.

In turn, as mentioned in the Method section, the MARS and uMARS objective quality scales contain four different subscales. In our study the information subscale got the highest mean score for both professionals ( $4.43 \pm 0.35$ ) and users ( $4.56 \pm 0.31$ ), followed by aesthetics ( $4.14 \pm 0.42$  for professionals and  $3.92 \pm 0.74$  for users) and functionality ( $4.14 \pm 0.24$  for professionals and  $3.80 \pm 0.28$  for users). The engagement subscale got the lowest mean score for both groups,  $3.83 \pm 0.55$  for professionals and  $3.75 \pm 0.34$  for users.

Finally, regarding subjective quality and perceived impact, both scores were rated higher by users, with a mean score of  $3.81 \pm 0.62$  for subjective quality compared to  $3.40 \pm 0.28$  for professionals and a mean score of  $4.67 \pm 0.47$  for perceived impact compared to  $4.23 \pm 0.48$  for professionals. The maximum rating possible was 5 for both overall scores. For further detail, means and standard deviations of objective quality subscales scores and subjective quality and perceived impact scores and items can be found in Tables 4 and 5.

#### Extracted categories, subcategories and areas

The thematic content analysis allowed for the identification of 68 areas, 11 subcategories and 3 categories for the professionals' focus group and 29 areas, 14 subcategories and 3 categories for the users focus group.

#### Information gathered in the professionals' focus group

As mentioned, three main categories were identified: Use of the app, design, and suggestions for improvement.

*Regarding the “Use of the app” category, three sub-categories were mentioned:*

**Facilitators** The professionals mentioned app characteristics and components that they found could make its use easier. Ten areas were identified, of which the most frequently mentioned were “Motivation elements”, “Dynamic” and “Progress graphs”. Some literal examples of the information mentioned are: “You have added a lot of reinforcement and these things are very good” and “Seeing progress has seemed to me to keep up the motivation quite a bit”.

**Barriers** The professionals commented on different app aspects that they consider to be obstacles when using it. On the one hand, they identified “digital gap” and “Time and effort”. On the other hand, 4 areas related to characteristics of some components of the app as videos or exercises were identified (e.g.: “Long Videos”). Some verbatim examples of the information provided are: “You have to put time and effort into it. And I think that’s a drawback” and “The videos are 21 minutes long, to do them in one go, it’s hard to keep your attention for so long”.

**Strengths** The professionals pointed out 3 areas as positive aspects of using the app. Specifically, they found it to be useful, empowering and highlighted its potential as a complement to face-to-face sessions with the therapist. Some literal examples of the mentioned information include: “I see it as a very good complement for the patient, as a follow-up, and for the therapist as well” and “We see patients every month at best... This will allow you to give continuity to the work”.

*Regarding the “Design” category, three subcategories were mentioned:*

**Information** Seven different areas related to app content were identified, some of the most frequently mentioned were: interesting, well collected, consistent and evidence-based. Some verbatim examples of the information provided are: “You integrate the different things from the first module. You can see the coherence between one module and another” and “For me it is a professional app, with an approach and a basis behind it. Yes, evidence-based and professional”.

**Aesthetics** This subcategory focuses on the visual aspect of the app, which was described as simple and correct. Some literal examples of the information mentioned are: “Correct. Simple, without stridency, but it gets to the point” and “Visually correct. I found it OK but not excessively striking”.

**App components** The professionals pointed out different characteristics of the app’s main components. The most mentioned areas were related to the UP-content videos (UP representative content and App therapist), the “Present awareness” module audios (Accessible and Varied) and the Testimonials (Real examples). Some literal examples of the mentioned information include: “The content videos where the girl appears reflect quite well the fundamental part of what is involved in the UP-sessions, the main ideas” and “The girl really conveys a lot of tranquillity. She catches my attention, how well she expresses and explains herself. That helps and makes it easier for you to maintain attention”.

Finally, five subcategories were mentioned in relation to the “**Suggestions for improvement**” category: Adherence, videos, exercises, emergency button and intervention format. Some verbatim examples of the information suggested are: “Notifications would be interesting to maintain motivation” and “I have missed being allowed to add information once sent. When you write in a paper record you can... but here it’s more complicated”. For further information on the areas identified and suggestions for improvement implemented, see 3.3 section.

For each category, the different subcategories and areas identified from the information gathered in the professionals’ focus group, as well as textual examples of each area, can be found in Additional file 3.

#### **Information gathered in the users’ focus group**

As mentioned, three main categories emerged: Strengths, barriers, and suggestions for improvement.

*Regarding the “Strengths” category, five subcategories were mentioned:*

**App therapist** The app’s therapist (i.e., the woman who stars in the UP-content videos, leading the explanation and guiding the treatment through the different modules) was highly appreciated by users. Five areas were identified in this regard, being “Facilitates focused attention” and “Peaceful” the most frequently mentioned. Some literal examples of the information mentioned are: “You feel so close to her. I think this girl conveys very much a sense of closeness” and “She catches your attention and makes you focus”.

**Testimonials** The users mentioned this app component was useful and interesting as well as being able to identify with the stories included. Some verbatim examples of the information provided are: “I identify with them” and “They have been very good for me. I answered the exercise and then I looked at it and redid it. That’s what I liked the most”.

**UP-content videos** Users mentioned that the content of the videos was interesting and well collected and explained and the supporting texts that appeared while the therapist was explaining were useful. Some literal examples of the mentioned information include: “It seemed to me that you have condensed all the information very well” and “Very well summarized, very well explained, very concise, very much to the point”.

**Exercises** Two areas were identified by users within this subcategory. They mentioned the exercises were specific and the one thing that helps while doing them is that they are very guided. Some literal examples of the information mentioned are: “Very guided. That helps a lot. That means you can get things, don’t get lost and don’t waste time” and “The questions are very specific. It’s very focused so that you answer what you have to answer and don’t ramble”.

**Utility** Users mentioned that they find the app very useful, so they think that many people could benefit from using it. Some verbatim examples of the information provided are: “I would support everyone doing it” and “There



are many people who would benefit a lot from this type of app”.

*Regarding the “Barriers” category, four subcategories were mentioned:*

**Commitment** Users commented on some aspects that they consider could threaten the commitment to the app. The most repeated area was “Concentration required”. In addition, “Effort” and “Time” areas were also mentioned. Some literal examples of the mentioned information include: “It has been an effort for me at certain times” and “When you get into the app you have to be 100% on it, you can’t be doing other things. You have to be very focused”.

**Technical aspects** This subcategory focuses on technical bugs that may happen while downloading and using the app. In this regard, one user mentioned problems downloading the application because of limited phone space, as can be observed in this literal example of the information provided: “I didn’t have enough space on the phone, so I had to download it on another mobile...”

**Use of the App** Two areas were identified: Two patients mentioned having difficulties in finding the testimonials and another patient in finding the drop-down button used to answer some of the exercises. Some literal examples of the information mentioned are: “I think the testimonials are very hidden” and “There are things that I was not very clear about. Above all, the drop-down button: for some unknown reason, I have not seen the sign meant to unfold and I have had to see the explanation several times”.

**App usage instructions video** The app usage instructions video (i.e., video that appears on the first screen of the app, prior to the start of treatment, giving instructions on how to move around the app) emerged as a subcategory, as 2 areas related mainly to the audio of the video were identified: “The sound quality is like very homemade. There’s an echo and it makes it look like you’ve made the video at home”.

Finally, five subcategories were mentioned in relation to the “**Suggestions for improvement**” category: Adherence, videos, aesthetics of “the UP house” included in the App, emergency button, and weekly emotional assessment. Some verbatim examples of the information suggested are: “I would prefer a more curvilinear design rather than so straight, the curves are relaxing” and “I miss the option of a relaxation technique on this button. If I had a crisis and pressed the button, I would not, for example, be satisfied with the text “at the next visit with the therapist you will discuss it”, I would like to have a solution right now”. For further information on the

areas identified and suggestions for improvement implemented, see 3.3 section.

For each category, the different subcategories and areas identified from the information gathered in the users’ focus group, as well as verbatim examples of each area, can be found in Additional file 3.

### **App integration of suggestions for improvement and content-related bug correction**

As can be seen in Additional file 3, a total of 20 different areas related to suggestions for improvement were collected, 14 proposed by professionals and 7 by users, with both groups agreeing on one of them. In total, 20% of the suggestions for improvement have been implemented in the refined version of the app. One of the suggestions implemented was notifications, proposed by both groups and which were programmed to be sent as a reminder to log into the app after 5 days without logging into it and to complete the weekly emotional assessment. Another suggestion implemented, only proposed by professionals, was for the app to resume playing a video from where the user had left off the previous time they had watched it. Finally, the remaining suggestions included were proposed by users and consisted of improving the audio quality of the app usage instruction video and the possibility of adding and/or modifying emotions to be assessed weekly.

Of the 16 suggestions not included, 14 related to the videos, exercises, and emergency button could not be implemented due to lack of budget but a note of them was made for future versions. As for the remaining 2, therapist-patient contact was not considered because it would increase the workload for the professionals and the complementary use in group therapy did not correspond to an improvement of the app itself but to the intervention format in which it could be used. Finally, in both focus groups some syntactic bugs were detected, which were corrected in the refined version of the app.

### **Data extraction reliability**

As observed in Table 6, the agreement index between the first data extraction and the final version of the extraction ranged from 83.52 to 100%, which indicates a moderate-high agreement between the first data extraction and the final version of the extraction.

### **Discussion**

The aim of this study was to analyse the perceived usability and quality of the preliminary version of ReguEm by users and professionals that had been involved in its design from the first stage of the process. In terms of the findings obtained, results highlight important key areas and data that can contribute to the ongoing

**Table 6** Data extraction reliability

FG session	Areas		Subcategories		Categories	
	%	K	%	K	%	K
Patients	94.87	0.85, almost perfect	100	1.00, perfect	100	1.00, perfect
Professionals	83.52	0.42, moderate	92.31	0.63, substantial	80.00	0.54, moderate

Note FG: Focus group; %: Percentage of agreement between the draft version and the final version of data extraction; K: Cohen's kappa.

research and practice in mHealth field and helped to refine this preliminary version of RegulEm.

Regarding SUS results, the usability of RegulEm was perceived as good by users and excellent by professionals. On the one hand, regarding users, the scale item with the highest mean score was the one related to the belief that they would use the app frequently, which can be explained by the different strengths of the app collected in the focus group, such as the testimonials section, in line with previous work on mental health apps that include this component as a way to enhance users' motivation [65]. Another identified strength that could explain this item to be the highest scored could be the exercises, of which the "guidance" and "concrete" areas were highlighted, in line with previous studies in which users of mental health apps indicated a preference for apps with simple and clear exercises guidelines [33]. On the other hand, in relation to the professionals, the item with the highest mean score was the one about feeling confident using the app, which could be explained by the qualitative information collected in relation to the information contained in the app, which they valued as professional and evidence-based, in line with previous work [66]. In turn, the item related to believing that they would use the app frequently also obtained a high average score, which could be understood given the strengths regarding the use of the app collected in the focus group of professionals, such as the fact that they consider it useful, a good complementary element to face-to-face therapy and empowering. Furthermore, in relation to empowerment, given the characteristics of BC, in which the user must take part of the management of the treatment through the work with the app, empowerment, related to self-efficacy [67], could be an important aspect. Bandura's social learning theory suggests that therapeutic change would derive from self-efficacy, defined as "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" [68]. In this sense, an empowering effect of using the app could be considered to influence the favorable outcome of the intervention. In this regard, previous work has found a partially supported mediating effect of self-efficacy, along with other cognitive variables, on the effect of Internet-based interventions for the treatment of depression [69] and posttraumatic stress disorder [70].

Data regarding usability is encouraging, since usability issues have been identified as the most prevalent barrier to engagement with mental health apps [65]. However, the total perceived usability scores varied approximately 10 points between both groups, being higher the perceived usability of the professionals. More specifically, the two items of the SUS scale in which a greater difference in scores was observed between the two groups, with a mean response of one point more in the case of professionals, were "I thought the app was easy to use" and "I would imagine that most people would learn to use the app very quickly". A possible explanation of that could be that the group of professionals, having a greater knowledge and command of each of the modules that compose the intervention (UP), perceive the app as easier to use and faster to learn to handle than the group of users, who despite being familiar with the intervention for having received it previously, they know it to a different degree or in a different manner. User's point of view includes being aware of their difficulties of tolerating intense emotions, something that is not easy from them to learn. However, despite the difference between the perceived usability scores of the two groups, we consider it important to highlight as a strength that both overall usability scores were good. Finally, it is important to point out that part of the future steps will be aimed at implementing the suggestions for improvement resulting from this study that could not be included in the app, which could lead to a better perceived usability of the app and a possible alignment of the usability scores of both groups.

Furthermore, the quality of RegulEm was rated as good by both users and professionals according to the uMARS and MARS results respectively. In both groups, the information subscale got the highest mean score. Regarding professionals, the different areas identified from the focus group regarding the information contained in the app, such as well-collected information, useful, interesting and evidence-based information could account for this being the most highly rated subscale. While in the case of users, information subscale being the highest scored could be explained by the strengths identified about the video content included in the app, such as the well-gathered and interesting content or the support texts. This data is promising, since due to the long periods between sessions in Spanish public mental health units [9], much of the focus of the

BC will be on the information provided by RegulEm. In addition, the engagement subscale got the lowest mean score in comparison with the remaining subscales in both groups. This data is meaningful if we observe that both users and professionals agreed in highlighting notifications as a suggestion for improvement, in line with previous work that mentions notifications as a necessary feature in mental health apps [71, 72]. In this regard, special attention should be paid to this aspect, as engagement has been suggested as a mechanism of action for clinical outcomes of mental health apps [73] and can be affected by factors such as usability, competing priorities and time [74].

On the other hand, it is important to mention that the qualitative analysis identified different barriers to the use of the app mentioned by users and professionals. In the case of users, technical problems or time, effort and concentration required were mentioned, among others. While in the case of professionals, barriers such as the digital divide along with others related to the app content, such as the length of the videos or the complexity of the emotional exposure exercises were identified. These findings are consistent with those supported by a review focused on identifying barriers to telemedicine use that are associated with reduced acceptance, which informs that most of these barriers are user-related, such as technical literacy, followed by those related to the intervention, such as technical problems [75]. Therefore, special attention should be paid to the barriers identified in the present study, since they could be associated with lower acceptance of the app by users and, therefore, with interruption in its use or non-use, affecting its future implementation. In this sense, 20% of the app improvement suggestions identified in this study were implemented in the app, some of them directly related to the mentioned barriers. However, it is important to take into account the rest of the improvement suggestions that could not yet be included in order to implement them in future versions of the app.

Results from this study should be interpreted in light of some limitations. First, convenience sampling was used. In addition, the sample size was small and data saturation was not possible. In turn, we consider it important to highlight as a limitation that some of the questions in the focus group interview guide were closed, which could have limited the conversation or given rise to ambiguous answers, and others were formulated in a suggestive manner, which could have reflected the moderator's expectation of a response and could have incited it [76]. At the same time, focus group interactions may lead to answers that are socially expected [55]. Finally, conclusions of this study should be interpreted within the context of Spanish public mental health units and may not be generalizable to those provided in other health contexts.

While acknowledging the aforementioned limitations, this study offers several strengths. Even though special

attention needs to be paid to the usability of mHealth apps before analysing their effectiveness [22], the number of published studies reporting usability assessment results is small and has decreased slightly compared to the rapidly growing number of this type of apps [35]. Therefore, this study reinforces the need to gather more evidence regarding usability assessment within this field.

In addition, after usability assessment, studies suggest to include the preliminary version of the app in a pilot study for final refinement before it is released in a larger trial [34]. Thus, it is worth noting that this study helps to ensure that RegulEm is properly designed and oriented to the needs of end users [35] before it is used as part of a BC intervention in a pilot study and a subsequent RCT that will analyse the intervention effectiveness, implementation, and cost-effectiveness [48].

It is also noteworthy that a parallel mixed methods design was followed. In this study, quantitative and qualitative methods are used independently, analysed separately and then their results are integrated into the overall interpretation [56]. Although qualitative methods may be an optimal way to collect details about user experiences and behaviors that cannot be captured with quantitative ones, findings indicate that not enough detailed information is obtained with these techniques [77]. In this manner, as discussed in previous literature, a mixed-method approach may be more useful [34, 35]. Thus, by following this approach we ensure to obtain a more complete picture of RegulEm's usability and quality as well as suggestions for improvement. Finally, another strength is that health professionals have been involved. In this regard, there is a lack of involvement of health professionals in usability evaluation of mental health apps, despite it being known that including them helps to ensure the medical quality of the app [32].

## Conclusion

In conclusion, this study, using a parallel mixed method, has analyzed the perceived usability and quality of the preliminary version of RegulEm by users and professionals who had been involved in its participatory design and development since the first stage of the process. The usability of RegulEm was perceived as good by users and excellent by professionals, while the quality was perceived as good by both groups. In turn, the different areas identified in both focus groups provide relevant information on the usability of RegulEm and have contributed to its refinement. Thus, future work will focus on analyzing RegulEm feasibility and preliminary effectiveness in combination with face-to-face UP sessions through a pilot study in Spanish public mental health units and a subsequent RCT that will analyze its effectiveness, implementation and cost-effectiveness in the same context. In addition, upcoming work should also focus on

further improving the app by integrating suggestions for improvement that have not been included so far, as well as exploring the opinions and experiences of people who do not respond to treatment or drop out.

#### Abbreviations

Apps	Smartphone applications
BC	Blended Care
CBT	Cognitive Behavioral Therapy
COREQ	Consolidated Criteria for Reporting Qualitative Research
EDs	Emotional Disorders
GRAMMS	Good Reporting of A Mixed Methods Study
MARS	Mobile Application Rating Scale
RCT	Randomized Controlled Trial
TAM	Technology Acceptance Model
SUS	System Usability Scale
uMARS	User Version of the Mobile Application Rating Scale
UP	Unified Protocol for Transdiagnostic Treatment of Emotional Disorders

#### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12911-024-02679-w>.

Supplementary Material 1: Additional file 1. Good Reporting of A Mixed Methods Study (GRAMMS). GRAMMS checklist items with the page in the manuscript where the relevant information for that item can be found. (DOCX 13 kb).

Supplementary Material 2: Additional file 2. Consolidated criteria for reporting qualitative research (COREQ). COREQ checklist items with the page in the manuscript where the relevant information for that item can be found. (DOCX 481 kb).

Supplementary Material 3: Additional file 3. Categories, subcategories and areas identified from the information gathered in the users' and professional's focus groups, as well as textual examples of each area. (DOCX 37 kb).

#### Acknowledgements

We would like to thank all the psychologists and patients who participated in RegulEm's development process for making this study possible.

#### Author contributions

LMG and JO were responsible for conception and design; LMG, AFJ, VFG, CRF and JO were involved in data acquisition; LMG, AFJ and JO did formal data analysis; JO, VFG and CRF provided the necessary resources to carry out the study; LMG was responsible for writing and original draft preparation; LMG, AFJ, VFG, CRF and JO carried out critical review. All authors have read and agreed to the final version of the manuscript.

#### Funding

This study was funded by the Ministerio de Industria, Economía y Competitividad ISCI (PI20/00697) and co-funded by European Union through FEDER Funds "A way to make Europe". This work was supported by the Spanish Ministry of Universities through the "Formación de Profesorado Universitario" Program [FPU20/03796] and by the Gobierno de Aragón [Grant Number Research team S31\_23R]. Funders were not involved in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

#### Data availability

Data is provided within the manuscript and supplementary information files.

#### Declarations

##### Ethics approval and consent to participate

The study was performed under the approval of the ethics committee of General University Hospital of Castellón, with number V3\_22\_04\_2021, and in line with the principles of the Declaration of Helsinki. All participants accepted and signed the informed consent.

##### Consent for publication

Not applicable.

##### Competing interests

The authors declare no competing interests.

Received: 13 June 2024 / Accepted: 11 September 2024

Published online: 27 September 2024

#### References

- Lecomte T, Potvin S, Corbière M, Guay S, Samson C, Cloutier B, et al. Mobile apps for Mental Health issues: Meta-Review of Meta-analyses. *JMIR mHealth uHealth*. 2020;8(5):e17458.
- World Health Organization. *mHealth: new horizons for health through mobile technologies: second global survey on eHealth*. 2011.
- Chandrasekar P. Do mental health mobile apps work: evidence and recommendations for designing high-efficacy mental health mobile apps. *mHealth*. 2018;4:6–6.
- Andrade LH, Alonso J, Mneimneh Z, Wells JE, Al-Hamzawi A, Borges G, et al. Barriers to mental health treatment: results from the WHO World Mental Health surveys. *Psychol Med*. 2014;44(6):1303–17. <https://www.cambridge.org/core/article/barriers-to-mental-health-treatment-results-from-the-who-world-mental-health-surveys/8779313B29B9F3950A0A1154949E0D21>.
- Bullis JR, Boettcher H, Sauer-Zavala S, Farchione TJ, Barlow DH. What is an emotional disorder? A transdiagnostic mechanistic definition with implications for assessment, treatment, and prevention. *Clin Psychol Sci Pract*. 2019;26(2):e12278.
- World Health Organization. *Mental health atlas 2020*. Geneva; 2021.
- OECD. Tackling the Mental Health Impact of the COVID-19 Crisis: an Integrated, whole-of-society response. OECD; 2021. [https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-2021\\_ae3016b9-en](https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-2021_ae3016b9-en).
- World Health Organization. *Mental Health Atlas 2017*. Geneva; 2018.
- Peris-Baquero O, Osma J. Unified protocol for the transdiagnostic treatment of emotional disorders in group format in Spain: results of a noninferiority randomized controlled trial at 15 Months after treatment onset. *Depress Anxiety*. 2023;1981377.
- Miralles J, Granel C, Díaz-Sanahuja L, van Woensel W, Bretón-López J, Mira A, et al. Smartphone apps for the treatment of mental disorders: systematic review. Volume 8. *JMIR mHealth and uHealth*; 2020.
- Hoa K, Janni E, Wrede R, Sedem M, Donker T, Carlbring P, et al. Experiences of a guided smartphone-based behavioral activation therapy for depression: a qualitative study. *Internet Interv*. 2015;2(1):60–8.
- Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. *World Psychiatry*. 2019;18(3):325–36.
- Lu S, Xu M, Wang M, Hardi A, Cheng AL. Effectiveness and minimum effective dose of app-based Mobile Health interventions for anxiety and Depression Symptom reduction: systematic review and Meta-analysis. *JMIR Ment Health*. 2022;9.
- Weisel KK. Standalone smartphone apps for mental health — a systematic review and meta-analysis. *NPJ Digit Med*. 2018;1–10.
- Kooistra LC, Ruwaard J, Wiersma JE, Oppen P, Van, Vaart R, Van Der Gemert-pijnenJEW, et al. Development and initial evaluation of blended cognitive behavioural treatment for major depression in routine specialized mental health care. *Internet Interv*. 2016;4:61–71.
- Erbe D, Eichert HC, Riper H, Ebert DD. Blending face-to-face and internet-based interventions for the Treatment of Mental Disorders in adults: systematic review. *J Med Internet Res*. 2017;19(9):e306.
- Osma J, Suso-Ribera C, Peris-Baquero Ó, Gil-Lacruz M, Pérez-Ayerra L, Ferreres-Galan V, Torres-Alfosea MÁ, López-Escriche M, Domínguez O. What format of

- treatment do patients with emotional disorders prefer and why? Implications for public mental health settings and policies. *PLoS ONE*. 2019;14(6).
18. Agnew JMR, Hanratty CE, McVeigh JG, Nugent C, Kerr DP. An investigation into the Use of mHealth in Musculoskeletal Physiotherapy: scoping review. *JMIR Rehabil Assist Technol*. 2022;9(1):e33609.
  19. Giebel GD, Speckemeier C, Abels C, Plescher F, Börchers K, Wasem J, et al. Problems and barriers related to the Use of Digital Health Applications: scoping review. *J Med Internet Res*. 2023;25:e43808.
  20. 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 Health*. 2018;21(3):116–9.
  21. Davis FD, Perceived, Usefulness, Perceived ease of Use, and user Acceptance of Information Technology. *MIS Q*. 1989;13(3):319–40.
  22. Brown W, Yen P-Y, Rojas M, Schnall R. Assessment of the Health IT usability evaluation model (Health-ITUEM) for evaluating mobile health (mHealth) technology. *J Biomed Inf*. 2013;46(6):1080–7.
  23. Sanders EBN. From user-centered to participatory design approaches. *Design and the Social Sciences: making connections*. CRC; 2002. pp. 1–8.
  24. Halje K, Timpka T, Ekberg J, Bång M, Fröberg A, Eriksson H. Towards mHealth systems for support of psychotherapeutic practice: a qualitative study of researcher-clinician collaboration in system design and evaluation. *Int J Telemed Appl*. 2016;2016:5151793.
  25. Lewis TL, Wyatt JC. mHealth and mobile medical apps: a framework to assess risk and promote safer use. *J Med Internet Res*. 2014;16(9):e210.
  26. Giebel GD, Speckemeier C, Schrader NF, Abels C, Plescher F, Hillerich V, et al. Quality assessment of mHealth apps: a scoping review. *Front Health Serv*. 2024;4:1372871.
  27. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR mHealth uHealth*. 2015;3(1).
  28. Gohari SH, Khordastan F, Fatehi F, Samzadeh H. The most used questionnaires for evaluating satisfaction, usability, acceptance, and quality outcomes of mobile health. *BMC Med Inf Decis Mak*. 2022;22(1):22.
  29. Bevan N. International standards for HCI and usability. *Int J Hum Comput Stud*. 2001;55(4):533–522.
  30. Islam MN, Karim M, Inan TT, Islam AKMN. Investigating usability of mobile health applications in Bangladesh. *BMC Med Inf Decis Mak*. 2020;3:1–13.
  31. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. *J Med Syst*. 2015;39(2):1.
  32. Inal Y, Wake JD, Guribye F, Nordgreen T. Usability Evaluations of Mobile Mental Health Technologies: systematic review. *J Med Internet Res*. 2020;22:1–19.
  33. Alqahtani F, Orji R. Insights from user reviews to improve mental health apps. *Health Inf J*. 2020;26(3):2042–66.
  34. Alwashmi MF, Hawboldt J, Davis E, Fetzters MD. The iterative Convergent Design for Mobile Health Usability Testing: mixed methods Approach. *JMIR mHealth uHealth*. 2019;7(4):e11656.
  35. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. *Int J Med Inf*. 2019;126:95–104.
  36. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, Boyd KA, Craig N, French DP, McIntosh E, Petticrew M, Rycroft-Malone J, White M, Moore L. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ*. 2021;374:n2061.
  37. Murray E, Treweek S, Pope C, MacFarlane A, Ballini L, Dowrick C, Finch T, Kennedy A, Mair F, O'Donnell C, Ong BN, Rapley T, Rogers A, May C. Normalisation process theory: a framework for developing, evaluating and implementing complex interventions. *BMC Med*. 2010;8:63.
  38. Barlow DH, Ellard KK, Fairholme CP, Farchione TJ, Boisseau CL, Allen LB, et al. Unified protocol for Transdiagnostic Treatment of Emotional disorders: Therapist Guide. 2nd ed. New York: Oxford University Press; 2019.
  39. Harvey AG, Watkins E, Mansell W. Cognitive behavioural processes across psychological disorders: a transdiagnostic approach to research and treatment. Oxford University Press; 2004.
  40. Brown TA, Campbell LA, Lehman CL, Grisham JR, Mancill RB. Current and lifetime comorbidity of the DSM-IV anxiety and mood disorders in a large clinical sample. *J Abnorm Psychol*. 2001;110(4):585–99.
  41. Brown TA, Barlow DH. A proposal for a dimensional classification system based on the shared features of the DSM-IV anxiety and mood disorders: implications for assessment and treatment. *Psychol Assess*. 2009;21(3):256–71.
  42. Ayuso-Bartol A, Gómez-Martínez MÁ, Riesco-Matías P, Yela-Bernabé JR, Crego A, Buz J. Systematic review and Meta-analysis of the efficacy and effectiveness of the Unified Protocol for Emotional disorders in Group Format for adults. *Int J Ment Health Addict*. 2024;1–27.
  43. Carlucci L, Saggino A, Balsamo M. On the efficacy of the unified protocol for transdiagnostic treatment of emotional disorders: a systematic review and meta-analysis. *Clin Psychol Rev*. 2021;87:101999.
  44. Cassiello-Robbins C, Southward MW, Tirpak JW, Sauer-Zavala S. A systematic review of Unified Protocol applications with adult populations: facilitating widespread dissemination via adaptability. *Clin Psychol Rev*. 2020;78:101852.
  45. Longley SL, Gleiser TS. Efficacy of the Unified Protocol: a systematic review and meta-analysis of randomized controlled trials. *Clin Psychol Sci Pract*. 2023;30(2):208–21.
  46. Sakiris N, Berle D. A systematic review and meta-analysis of the Unified Protocol as a transdiagnostic emotion regulation based intervention. *Clin Psychol Rev*. 2019;72:101751.
  47. Schaeuffele C, Meine LE, Schulz A, Weber MC, Moser A, Paersch C, et al. A systematic review and meta-analysis of transdiagnostic cognitive behavioural therapies for emotional disorders. *Nat Hum Behav*. 2024;8(3):493–509.
  48. Osma J, Martínez-García L, Prado-Abril J, Peris-Baquero Ó, González-Pérez A. Developing a smartphone app based on the unified protocol for the transdiagnostic treatment of emotional disorders: a qualitative analysis of users and professionals' perspectives. *Internet Interv*. 2022;30:100577.
  49. Stoll R, Pina A, Gary K, Amresh A. Usability of a Smartphone Application to support the Prevention and early intervention of anxiety in Youth. *Cogn Behav Pract*. 2018;24(4):393–404.
  50. Peris-Baquero Ó, Moreno JD, Osma J. Long-term cost-effectiveness of group unified protocol in the Spanish public mental health system. *Curr Psychol*. 2022.
  51. Del Sevilla-Gonzalez R, Loeza M, Lazaro-Carrera LM, Ramirez LS, Rodríguez BB, Peralta-Pedrero AV. Spanish version of the system usability scale for the assessment of electronic tools: development and validation. *JMIR Hum Factors*. 2020;7(4):1–7.
  52. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the mobile application rating scale (uMARS). *JMIR mHealth uHealth*. 2016;4(2).
  53. Martin Payo R, Fernandez Álvarez MM, Blanco Díaz M, Cuesta Izquierdo M, Stoyanov SR, Llaneza Suárez E. Spanish adaptation and validation of the Mobile Application Rating Scale questionnaire. *Int J Med Inf*. 2019;129:95–9.
  54. Martin-Payo R, Carrasco-Santos S, Cuesta M, Stoyanov S, Gonzalez-Mendez X, Del Fernandez-Alvarez M. Spanish adaptation and validation of the user version of the Mobile Application Rating Scale (uMARS). *J Am Med Inf Assoc*. 2021;28(12):2681–6.
  55. Bryman A. *Social Research Methods 4th edition*. Social Research Methodology. New York: Oxford University Press; 2012.
  56. Schoonenboom J, Johnson RB. How to Construct a Mixed Methods Research Design. 2017;107–31.
  57. O' Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. *J Health Serv Res Policy*. 2008;13(2):92–8.
  58. Barlow DH, Ellard KK, Fairholme CP, Farchione TJ, Boisseau CL, Allen LB, et al. Unified protocol for Transdiagnostic Treatment of Emotional disorders: Workbook. 2nd ed. New York: Oxford University Press; 2019.
  59. Corp IB. *IBM SPSS statistics for windows, version 22.0*. Armonk, NY: IBM Corp; 2013.
  60. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Heal Care*. 2007;19(6):349–57.
  61. Kuckartz U, Rädiker S. *Analyzing qualitative data with MAXQDA*. Switzerland: Springer International Publishing; 2019.
  62. Schreier M. *Qualitative content analysis in practice*. Sage; 2012.
  63. Guetterman TC, Creswell JW. Integrating Quantitative and Qualitative Results in Health Science Mixed Methods Research Through Joint Displays. 2015;554–61.
  64. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean; adding an adjective rating. *J Usability Stud*. 2009;4(3):114–23.
  65. Balaskas A, Schueller SM, Cox AL, Doherty G. Understanding users' perspectives on mobile apps for anxiety management. *Front Digit Heal*. 2022;4:1–20.
  66. Lui L, Marcus JH, Barry DK, Lui CT. Professional psychology: research and practice evidence-based apps? A review of Mental Health Mobile Applications in a psychotherapy context evidence-based apps? A review of Mental Health Mobile Applications in a psychotherapy context. *Prof Psychol Res Pract*. 2017;48(3):199–210.
  67. Cattaneo LB, Chapman AR. The process of empowerment: a model for use in research and practice. *Am Psychol*. 2010;65(7):646–59.



68. Bandura A. Self-efficacy: the exercise of control. Self-efficacy. W H Freeman/ Times Books/ Henry Holt & Co.; 1997.
69. Domhardt M, Steubl L, Boettcher J, Buntrock C, Karyotaki E, Ebert DD, et al. Mediators and mechanisms of change in internet- and mobile-based interventions for depression: a systematic review. *Clin Psychol Rev.* 2021;83:101953.
70. Steubl L, Sachser C, Baumeister H, Domhardt M. Mechanisms of change in internet- and mobile-based interventions for PTSD: a systematic review and meta-analysis. *Eur J Psychotraumatol.* 2021;12(1):1879551.
71. Oyebo O, Alqahtani F, Orji R. Using Machine Learning and Thematic Analysis Methods to Evaluate Mental Health apps based on user reviews. *IEEE Access.* 2020;8:111141–58.
72. Alqahtani F, Winn A, Orji R. Co-designing a Mobile app to improve Mental Health and Well-Being: Focus Group Study. *JMIR Form Res.* 2021;5(2):e18172.
73. Graham AK, Kwasny MJ, Lattie EG, Greene CJ, Gupta NV, Reddy M, et al. Targeting subjective engagement in experimental therapeutics for digital mental health interventions. *Internet Interv.* 2021;25:100403.
74. Nwolise CH, Carey N, Shawe J. Preconception and diabetes information (PADI) app for women with Pregestational Diabetes: a feasibility and acceptability study. *J Healthc Inf Res.* 2021;5(4):446–73.
75. Reinhardt G, Schwarz PEH, Harst L. Non-use of telemedicine: a scoping review. *Health Inf J.* 2021 Oct-Dec;27(4):14604582211043147.
76. United States Agency for International Development (USAID). A Step-By-Step Guide to Focus Group Research for Non-Governmental Organizations. 2012.
77. Yen PY, Bakken S. Review of health information technology usability study methodologies. *J Am Med Inf Assoc.* 2012;19(3):413–22.

### Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.