Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

TETI, an LLM-Augmented Telegram Chatbot for WHO Self-Help Plus (SH+): Single-arm mixed-methods feasibility study protocol

Stefano Fait
Fondazione Bruno Kessler
Leonardo Sanna
Fondazione Bruno Kessler
Silvia Rizzi
srizzi@fbk.eu

Fondazione Bruno Kessler

Study protocol

Keywords: Chatbot, Large language models (LLM), Self-Help Plus (SH+), Acceptance and Commitment Therapy (ACT), Workplace mental health, User experience (UX), User engagement (UE), Feasibility, Mixed-methods

Posted Date: November 10th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8592161/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

TETI, an LLM-Augmented Telegram Chatbot for WHO Self-Help Plus (SH+): Single-arm mixed-methods feasibility study protocol

Stefano Fait[1]; Leonardo Sanna[2] and Silvia Rizzi[1,*]

[1] Digital Health Research, Centre for Digital Health & Wellbeing, Fondazione Bruno Kessler, Via Sommarive 18, 38123 Trento, Italy

[2] Intelligent Digital Agents, Centre for Digital Health & Wellbeing, Fondazione Bruno Kessler, Via Sommarive 18, 38123 Trento, Italy

[*] Corresponding Author:

Name Surname: Silvia Rizzi

Affiliation: Digital Health Research, Centre for Digital Health & Wellbeing, Fondazione Bruno Kessler,

Address: Via Sommarive 18, 38123 Trento, Italy

Email: srizzi@fbk.eu

Phone: 0461312415

ABSTRACT

Background: Workplaces expose employees to psychosocial risks that can undermine mental well-being and contribute to stress and common mental health problems. Scalable digital conversational agents can reduce access barriers and provide brief, on-demand psychoeducation aligned with stepped-care models. TETI will be a Telegram chatbot delivering WHO Self-Help Plus (SH+) content, an ACT-based program, augmented with large language model (LLM) modules to enhance personalisation and engagement. This protocol primarily aims to explore User Experience (UX) and Eser Engagement (UE) with TETI in a workplace

setting, involving staff from the Bruno Kessler Foundation, Italy. Secondary aims are to assess acceptability and feasibility, and to explore proximal changes in distress, well-being, psychological flexibility, and knowledge retention.

Methods: This mixed-methods feasibility study will involve stakeholders in a workplace context, including domain experts in psychology, usability and potential end-users, who will interact with the chatbot TETI on alternate days over two weeks. Participants will receive SH+-consistent psychoeducational content (text, audio, visuals) via Telegram, with LLM-supported conversational flexibility under predefined guardrails. The primary outcomes will focus on UX and UE, which will be assessed using a Semantic Differential scale, an adapted version of the uMARS, and the CES-LCC. Qualitative post-use semi-structured interviews will be conducted to enrich the findings and provide further qualitative insights. Secondary outcomes will also be evaluated, including measures of distress and well-being, psychological flexibility, and a brief knowledge quiz assessing understanding of the intervention's content. To determine the feasibility of the intervention, metrics such as recruitment, retention, interaction with the chatbot, and overall user satisfaction will be assessed. This study will involve assessments at three time points: baseline (T0), post-intervention (T1, after two weeks of chatbot use), and follow-up (T2, two weeks after T1).

Discussion: This study will provide early evidence on the usability, acceptability and feasibility, along with early signs of change of a digital version of SH+, which will deliver psychoeducational content via a chatbot-based messaging system augmented with LLM support. Therefore, findings will help shed light on the strengths, limitations, and future directions for the digital adaptation of the WHO SH+ intervention, including the use of an LLM-augmented chatbot.

Trial registration: In accordance with Italian Law No. 3/2018, art. 2, paragraphs 7 and 10, this study does not require ethics committee approval because: i) it focuses on the usability of a prototype digital intervention, ii) no patients or vulnerable populations will be involved, iii) no clinical treatment or collection of sensitive health data is foreseen. Moreover, clear adherence to

ethical principles is ensured, including voluntary informed consent and protection of participants' privacy.

Keywords: Chatbot; Large language models (LLM); Self-Help Plus (SH+); Acceptance and Commitment Therapy (ACT); Workplace mental health; User experience (UX); User engagement (UE); Feasibility; Mixed-methods.

Background

Work environments expose many workers to psychosocial risks (excessive workloads, lack of decision-making autonomy, job insecurity, problematic organizational cultures) that undermine mental well-being and fuel stress, anxiety, and other disorders [1]. It is estimated that 15% of working-age adults experience mental health issues at any given time, with an economic impact of 12 billion lost workdays and approximately $1 trillion per year, making it urgent to strengthen psychological support in the workplace [2]. International institutions recommend a systemic approach that integrates risk prevention, promotion of well-being, and support for individuals experiencing mental health issues, aiming to enhance participation, safety, and performance at work [3]. Recommended interventions include organizational measures (e.g., flexibility, workload management, combating violence and harassment), training managers and workers in mental health skills and literacy, and individual interventions to manage stress [4]. The World Health Organization's (WHO) overall strategy for promoting mental health is based on principles of inclusion and scalability. In this context, particular attention is paid to the potential role of digital technologies in promoting psychological well-being, as they can represent an optimal solution; people can access them from anywhere and at any time, thus remaining flexible to their needs and reducing barriers related to service availability, social stigma, and asking for help [5].

The adoption of digital tools, particularly chatbots designed for psychological support based on artificial intelligence, can broaden access, reduce barriers, and provide short, scalable interventions for stress management and promoting well-being in the workplace. Recent evidence shows that chatbots produce small to moderate but significant reductions in depressive and anxiety symptoms in the short term, with benefits tending to emerge between 4 and 8 weeks of use [6]. A 2023 review also indicates that Artificial Intelligence (AI) conversational agents are particularly effective in reducing psychological distress, with more pronounced effects for generative, multimodal systems delivered via messaging apps, factors that increase personalization and engagement [7].

In the workplace, practical benefits include: (i) 24/7 access, anonymity, and reduced stigma, with just-in-time micro-interventions (psychoeducation, Cognitive Behavioural Therapy (CBT)/Acceptance and Commitement Therapy (ACT) techniques, guided meditations) available when needed [7]; (ii) scalability and early triage: chatbots can serve as the first level of screening and guidance, freeing up specialist resources for more complex cases, in line with scale-up models and the Organisation for Economic Co-operation and Development (OECD)-recognized need to use digital technology to increase productivity and coverage of occupational health interventions [8–10]; (iii) integration with Employee Assistance Program (EAP)/ Human Resources (HR) programs and corporate platforms, with collection of aggregated and anonymized results to assess the impact on stress, absenteeism, and well-being, while preserving privacy [9,10].

The study described in this protocol is part of the WHO's strategies. It aims to assess the acceptability and feasibility of an intervention delivered via digital tools to promote psychological well-being. The intervention will consist of structured dialogue segments of the WHO's Self Help Plus (SH+) intervention [11–14] enriched by large language model (LLM) modules to personalize content and enhance user engagement. The target population is workers.

The intervention, originally validated by the WHO itself and called SH+ [11–13], is a protocol that has been effective in reducing psychological distress and preventing the onset of mental disorders in various contexts, with results supported by RCTs and meta-analyses, and is also available in Italian [15]. There are also experiences and studies of digital adaptation (including the chatbot format) that indicate good usability and engagement potential, making the model suitable for a business context [14,16].

SH+ is based on the principles of Acceptance and Commitment Therapy (ACT), a form of third-wave cognitive-behavioral therapy [17,18], and is organized into five chapters each addressing a strategy for managing stress and promoting psychological well-being. These chapters, which are then translated into sessions, are: These chapters, which will be translated into sessions,will cover: (1) Grounding (mindfulness); (2) Detachment (defusion); (3) Acting in accordance with one's values (value-based behavioral activation); (4) Being Kind (gratitude); and (5) Making Space (acceptance).

Compared to an earlier digital version [14], Large Language Model (LLM) modules have been added to enhance user engagement. A preliminary study describes the potential benefits of incorporating these modules compared to a rule-based version. [19].

This trial will allow for a rigorous yet lightweight assessment of the feasibility and impact of SH+ in digital format, with potential scalability and integration into workplace wellness programs if the results are positive.

Aim of the Study

This study aims to test the prototype of a chatbot powered by an LLM, named TETI, to gather feedback and identify the needs of key stakeholders (eg, domain experts in psychology, usability specialists, and potential end users). The insights collected will be used to qualitatively refine the chatbot's dialogue, focusing on the intervention's usability, accessibility, and acceptability.

The study is part of the design and development phase of the Obesity-Related Behavioral

Intervention Trials (ORBIT) model [20], specifically Phase 1a. It will involve the delivery of psychoeducational materials in text, audio, and video formats by a virtual coach (i.e., the digital assistant TETI) via the Telegram platform.

Primary objective

Since the standard, face-to-face version of the SH+ intervention has been validated and tested across different contexts and on different populations [7,21,22], our primary aim is to assess its feasibility and applicability in a digital format, with a particular focus on the integration of a chatbot powered by LLM.

Therefore, the main objective of this study is to explore the User Experience (UX) and User Engagement (UE) of key stakeholders after interacting with an early prototype of the TETI chatbot delivered via Telegram.

Secondary objectives

The secondary objectives focus on assessing pre- and post-intervention psychological well-being using self-report questionnaires administered at baseline and at the end of the psychoeducational program, with particular attention to psychological distress, well-being, and psychological flexibility. In addition, we will aim to evaluate learning outcomes related to the psychoeducational content through a short learning quiz administered at the end of the program. These clinical outcomes are intended solely as exploratory and serve to assess the feasibility of the measures, not to infer efficacy.

Methods

Study design

This is a mixed-methods feasibility study with a co-design orientation, conducted in a workplace setting (Fondazione Bruno Kessler (FBK) in Trento). Primary outcomes focus on UX and UE to be measured post-use with an adapted version of the user-version Mobile App Rating Scale (uMARS) [23,24], the Comprehensive Evaluation Scale for LLM-Powered Counseling Chatbots (CES-LCC) [25], and a bespoke Semantic Differential scale [26] covering communication, structure, materials, and content variables. Secondary exploratory outcomes include changes in psychological distress and well-being, measured with the Patient Health Questionnaire-4 (PHQ-4) [27] and the World Health Organization-Five Well-Being Index (WHO-5) [28], and psychological flexibility measured with the Acceptance and Action Questionnaire version 2 (AAQ-II) [29].

Stakeholders comprise domain experts (psychology/usability) and potential end-users recruited via convenience sampling; eligible participants are adults able to provide informed consent. Participants will interact with TETI within a defined, brief exposure window embedded in their routine context; questionnaires will be administered at baseline and post-use, and post-use interviews will elicit perceived usefulness, clarity, accessibility, and improvement priorities. Quantitative analyses will be descriptive with exploratory within-participant pre–post comparisons (reporting effect sizes and confidence intervals), while qualitative data will undergo thematic analysis. Mixed-methods triangulation will integrate findings to inform iterative refinement, strengthen fidelity safeguards to SH+, and determine readiness for subsequent evaluative studies.

Participants

Given that this protocol describes a single-arm study with no planned between-group comparisons, a formal power analysis is not required. We plan to recruit a total of 22 stakeholders: 8 domain experts (4 psychologists; 4 usability experts) and 14 potential end-users (10 FBK employees; 4 HR/welfare experts).

This sample size is justified based on our previous experience with a similar methodology used in the development of ALBA [16], a chatbot prototype that guided participants through the SH+ intervention but, unlike TETI, was not based on large language model (LLM) technologies and followed a predefined interaction flow regardless of the responses provided. In that study, a sample of 28 participants proved more than sufficient to achieve data saturation and to provide comprehensive insights for intervention refinement. Moreover, in the field of digital health and usability testing, it is well established that relatively small, heterogeneous groups can identify a substantial portion of usability issues and feasibility barriers. Classic usability research indicates that testing with as few as five users can reveal up to 85% of usability problems, depending on task complexity [30]. Subsequent research further supports the idea that larger groups (e.g., 10–15 users) significantly increase the likelihood of detecting problems while reducing the risk of overlooking critical issues [31]. Furthermore, as the sample reaches 20 participants, a very high level of certainty is achieved, ensuring that the vast majority of existing design flaws have been identified. By involving 22 participants, our study exceeded these established benchmarks, ensuring a broad range of perspectives and strengthening the feasibility assessment of the SH+ digitalization.

All data will be collected confidentially, following participants' informed consent, and will subsequently be analyzed and processed using statistical software. We will invite participants to take part in the study through word of mouth and personal acquaintances. Participation will be entirely voluntary, and no compensation will be provided. The objectives of the study will be clearly explained from the outset. Only participants who provide informed consent on a

voluntary basis will be recruited. All adult volunteers are eligible to participate, as there are no additional exclusion criteria.

Procedure and Timeline

The experimental procedure is organized as follows

1. First, participants will be provided with the privacy policy and will be asked to give informed consent before starting the study;
2. Participants will complete pretest questionnaires to collect baseline data on psychological distress, well-being and psychological flexibility (T0);
3. Participants will interact with the TETI chatbot modules via Telegram on their phones every other day for two weeks;
4. Once participants have finished interacting with the TETI chatbot, they will complete questionnaires regarding their overall experience with the prototype (focusing on usability, engagement, and the perceived comprehensibility and usefulness of the content) as well as a short learning quiz assessing their acquisition of the psychoeducational material (T1);
5. Finally, two weeks after the end of the interaction a brief semi-structured interview will be conducted, and the questionnaires assessing psychological distress, well-being and psychological flexibility will be collected again at follow-up (T2).

Only participants who provide informed consent and agree to the privacy policy will take part in the experimental phase. Data collection will be conducted confidentially and anonymously: participants will be assured that their information will remain confidential, and during the analysis phase, they will be identified solely by an alphanumeric code. During the data collection phase, semi-structured interviews will be conducted and recorded, and the recordings will later be transcribed and analysed. Participants who are interested in the results of the research in

which they took part may, at the end of the study, contact the principal investigator, who will make the findings of the data analysis available to them.

During phase 2 of the study, participants will complete three questionnaires to collect baseline measures of psychological distress, well-being, and psychological flexibility. Specifically, the instruments employed include the Patient Health Questionnaire-4 (PHQ-4) [27], the World Health Organization-Five Well-Being Index (WHO-5) [28] and the Acceptance and Action Questionnaire II (AAQ-II) [29].

The PHQ-4 [27] is an ultra-short 4-item questionnaire used to screen for anxiety and depression. It consists of two questions on anxiety and two on depression. Responses are on a 4-point scale, ranging from "not at all" to "nearly every day."

The WHO-5 [28] is a self-assessment tool that measures mental well-being. It consists of five statements relating to the past two weeks. Each statement is rated on a 6-point scale (0–5), with higher scores indicating greater mental well-being. The tool has been translated into over 30 languages, including Italian.

The AAQ-II [29] is a 10-item questionnaire using a 7-point scale (1 = never true, 7 = always true) designed to assess a person's psychological flexibility and their ability to remain in contact with their emotions. Higher scores indicate greater psychological flexibility. The questionnaire has been translated and adapted for the Italian context. [32].

During phases 4 and 5, participants will be asked to evaluate their experience with the prototype by completing three questionnaires and participating in a semi-structured interview. For the questionnaire-based evaluation, we will employ the Semantic Differential scale[26], an adapted version of the uMARS [23] (details regarding the adapted version are provided below) and the CES-LCC [25].

The Semantic Differential scale consists of a set of bipolar adjective pairs (e.g., judgmental – welcoming, boring – engaging, stressful – relaxing), each rated on a 5-point rating scale . The adjective pairs were chosen in accordance with the study's objectives. The questionnaire is organized into four dimensions, corresponding to the main variables under investigation: (1)

Communication (empathy and listening, smoothness and fluidity, chatbot interaction and vocabulary); (2) Module Structure (session duration, notifications and reminders, external links); (3) Materials (diary, audio tracks, infographics, and videos); and (4) Content (content evaluation, clarity, and related aspects).

The uMARS [23,24] questionnaire evaluates mobile applications by covering four objective dimensions and one subjective dimension. Given that only some sections of the uMARS are relevant to the aims of this study, and to minimize the load on participants, we will employ a shortened version of the uMARS questionnaire, incorporating only the relevant subscales: Engagement (5 items), Information (4 items) and Subjective Quality (4 items) along with the Perceived Impact subscale (6 items), which assesses users' perceptions of the app's usefulness.

In addition, we will employ the CES-LCC [25] to assess the quality and perceived effectiveness of the chatbot interaction, highlighting the LLM's role as a core component of the intervention. The scale features 27 items grouped into nine dimensions: Understanding Requests, Providing Helpful Information, Clarity and Relevance of Responses, Language Quality, Trust, Emotional Support, Guidance and Direction, Memory, and Overall Satisfaction.

To further examine the intervention, we will conduct brief semi-structured interviews, audio-recorded, to gather additional qualitative feedback on the user experience and participants' suggestions.

At the end of the interaction with the chatbot, participants will complete a short learning quiz to assess their acquisition of the psychoeducational material presented during phase 3. The quiz will consist of two questions per chapter, for a total of 10 multiple-choice items. Each question will offer four possible answers, of which only one is correct.

Table 1 summarises the instruments used and the times at which they will be administered.

Table 1. Summary of the assessment instruments and their administration timing.

Objective	T0 (Week 0)	T1 (Week 2)	T2 - follow up (2 weeks later)

Primary	-	Semantic Differential Scale; Adapted uMARS; CES-LCC	Semi-structured interview
Secondary	PHQ-4; WHO-5; AAQ-II	Learning quiz	PHQ-4; WHO-5; AAQ-II

Study intervention and technological tool

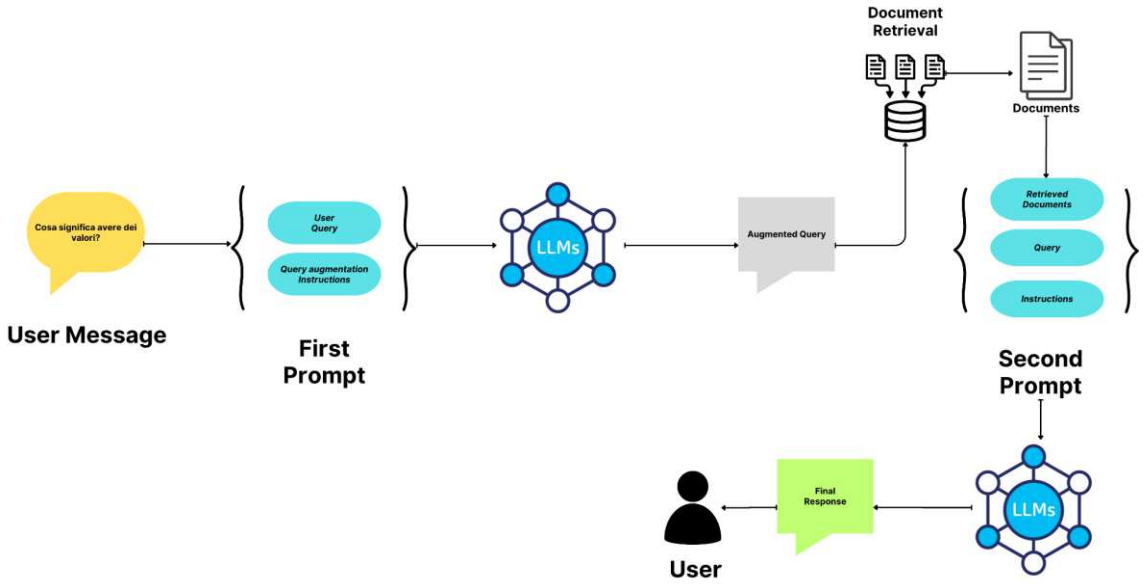
SH+ is structured in five sessions based on ACT: (1) Grounding, (2) Unhooking, (3) Acting on your values, (4) Being kind, (5) "Making space." The in-person group intervention has been adapted to individual delivery via chatbot on an app (TETI) [16].

Our study involves the integration of LLM modules to increase user engagement by generating personalized messages. This approach has been proven to be effective in a preliminary study [19]. The LLM modules are incorporated by constructing four different types of prompts that provide the necessary context to guide message generation within the intervention. These types of prompts are:

- Base: used in most interactions. It should be applied when there is only one standard response in the previous conversation, and the response to be generated does not require the use of other data previously entered by the user and/or explanations to be provided to the user.
- Focus: simplified version of the "basic" interaction, to be used when you want to focus the generated text on specific aspects, using elements previously saved in the conversation, so that the response is more accurate and includes elements of "memory" with respect to the information acquired.
- Explain: simplified version of the "basic" interaction, used to explain a concept that has already been addressed. It is also employed dynamically in Retrieval-Augmented Generation interactions (currently under development), where there is no predefined standard response (e.g., general question "Is anything unclear about what has been done

so far?"). In that case, the answer is retrieved, followingthe user query, from the manual using retrieval modules. In the retrieval modules (see Figure 1), the tool has access to the Self-Help+ manual in Italian, indexed in a vector database. When the user asks a question, the system reformulates it in the most consistent way possible with the manual's language and content. It then searches the database for the most relevant passages similar to the reformulated question. Based on the retrieved content, the tool generates a final answer and also indicates which sections of the manual the information was taken from.

Figure1. The Retrieval-Augmented Generation interactions



- Multiple standard answers: used when the original chatbot offered multiple standard responses depending on the user's sentiment, selected via buttons. In this case, the buttons are removed, and the LLM determines, based on context, which tone of voice is most appropriate.

The intervention is typically delivered in six weekly sessions. However, in this study, for timing reasons, participants will interact with the TETI chatbot every other day over two weeks.

Data analysis

Statistical analysis of the collected data will be performed using statistical software. Categorical variables will be summarised using absolute frequency distributions and percentages, while quantitative variables will be summarised using specific indices of centrality and variability.

Regarding the analysis of semi-structured interviews, a text mining approach will be used to extract the responses that appear most frequently in the interviews on user experience.

A convergent design will integrate quantitative and qualitative strands at interpretation using joint displays that align UX/UE domain scores with thematic findings (convergence, complementarity, divergence). Meta-inferences will prioritize actionable design changes, fidelity safeguards to SH+, and implementation implications. Integrated results will inform a ranked backlog for iteration and readiness criteria for subsequent evaluative studies (e.g., pilot RCT or implementation trial).

Discussion

Expected results

The primary objective of this protocol is to assess the impact of integrating Large Language Model (LLM) modules into the TETI chatbot by comparing usability outcomes with those obtained from testing a deterministic, pre-scripted conversational flow implemented in the previously developed ALBA chatbot [16]. Progression to a definitive trial will depend on whether predefined benchmarks for recruitment, retention, and usability are met.

Firstly, as regards sections A (Engagement) and D (Information) of the uMARS, we expect to achieve mean scores equal or superior to 3.55 (0.43) and 4.20 (0.52), respectively, as reported after testing ALBA. Particular attention will be paid to an expected increase in the 'Engagement' subscale, driven by the LLM's conversational flexibility. Secondly, we expect to achieve Semantic

Differential scores equal or superior to those previously observed across the dimensions of Communication, mean score equal or superior to 3.89 (0.63), on a 5-point scale. In particular, we hypothesize a significant improvement in the Communication dimension, as the LLM modules are expected to enhance conversational flow and perceived empathy.

Furthermore, we integrated the CES-LCC scale to evaluate AI-specific features that were not present in ALBA. This will help clarify whether the integration of LLM modules leads to a more effective psychoeducational experience or if it introduces conversational ambiguities that could confuse the user. The gathered data will determine whether TETI is ready for a large-scale randomized controlled trial aimed at assessing the clinical outcomes expected from the SH+ intervention, specifically focusing on reducing psychological distress (PHQ-4), enhancing mental well-being (WHO-5), and increasing psychological flexibility (AAQ-II).

A favourable outcome could support the development and dissemination of a web app that is easily accessible and usable at any time, and that promotes psychological flexibility in the general population. This aspect is particularly relevant in the workplace, as work represents a central component of adult life and work-related stress poses a significant threat to individual health and well-being, as well as incurring high economic costs for organisations, especially in terms of absenteeism due to illness [33]. While technological progress can contribute to increased stress, for example, through increased perceived workload and the feeling of constant connection [34], it also offers advantages such as flexibility [35] and work autonomy [36]. In this context, an innovation such as the digitisation of this type of intervention can represent a valuable additional resource for stress prevention, particularly secondary stress linked to cognitive assessment and individual responses to primary stressors, such as direct environmental or organisational factors (e.g., inadequate workspaces or role ambiguity). In particular, strengthening individual resources and enhancing coping strategies can help reduce the risk of developing work-related illnesses and accidents at work [37].

The analysis of the data collected in this study will be mainly descriptive and qualitative, aimed at assessing the achievement of the study objectives.

Conclusion

This protocol describes a mixed-methods feasibility study aimed at exploring the UX and UE of TETI, a Telegram chatbot that delivers content from the WHO Self-Help Plus (SH+) programme based on ACT principles, enriched with LLM modules for personalisation and engagement. In a real-world setting (FBK), expert stakeholders and potential users will interact with TETI every other day for two weeks, with quantitative and qualitative assessments at three timepoints (T0, T1, T2). In addition to UX/UE, the study will explore the acceptability and feasibility of the intervention, as well as preliminary signals on distress, well-being, psychological flexibility, and content comprehension.

The study has distinctive strengths: (a) anchoring to an evidence-based intervention (SH+) already validated in different contexts and transferred to chatbot format; (b) evaluation focused on multiple stakeholders (psychologists, usability experts, employees, HR) in a natural context of use; (c) multimodal outcome measures (UX/UE, psychological signals, learning, interviews) and convergent methodological integration; (d) LLM architecture designed to promote personalisation, transparency and knowledge retrieval from reliable sources; (e) low cost for participants and potential organisational scalability.

However, there are limitations that may affect the interpretation of the results: small and convenience sample (n=22) in a single institution; short duration of exposure (two weeks) with limited follow-up; absence of a control group and, therefore, inability to make causal inferences about clinical outcomes; prevalence of self-reported measures and possible desirability bias; novelty/expectation effect and risk of selection bias; generalisability limited to other sectors or organisational cultures; dependence on a messaging app and potential technical constraints; risks inherent in the use of LLM (e.g., hallucinations) mitigated but not eliminated by guardrails.

The practical implications, if feasibility and acceptability are favourable, will include the possibility of integrating TETI into corporate welfare and EAP programmes as a 'light' and scalable level of psychoeducational support, consistent with stepped-care models: on-demand micro-interventions, literacy and early triage, with potential benefits for perceived stress,

engagement and organisational climate. The use of standardised metrics and qualitative feedback will guide a roadmap for improvement, with a focus on clarity of materials, session duration, notification management and responsible personalisation.

In summary, TETI represents a pragmatic step towards the digital and scalable transfer of a validated psychoeducational intervention. This feasibility study aims to produce initial evidence on usability, engagement and acceptability in the workplace, laying the methodological and operational foundations for more rigorous evaluations of effectiveness and responsible adoption within companies.

List of abbreviations

WHO - World Health Organization

AI - Artificial Intelligence

CBT - Cognitive Behavioural Therapy

ACT - Acceptance and Commitment Therapy

OECD - Organisation for Economic Co-operation and Development

EAP - Employee Assistance Program

HR - Human Resources

SH+ - Self Help Plus

LLM - Large Language Model

ORBIT - Obesity-Related Behavioral Intervention Trials

UX - User Experience

UE - User Engagement

FBK - Fondazione Bruno Kessler

uMARS - user version of the Mobile Application Rating Scale

CES-LCC - Comprehensive Evaluation Scale for LLM-Powered Counselling Chatbots

PHQ-4 - Patient Health Questionnaire-4

WHO-5 - World Health Organization-Five Well-Being Index

AAQ-II - Acceptance and Action Questionnaire version 2

Ethics approval and consent to participate

The study did not require ethics committee approval in line with Italian Law No. 3/2018, art. 2, paragraphs 7 and 10, as i) it focused on the usability of a prototype digital intervention, ii) there is no involvement of patients nor vulnerable groups, iii) no clinical treatment nor sensitive health data collection are foreseen. At the same time, a clear adherence to ethical principles is ensured (voluntary informed consent, privacy protection).

Consent for publication

Not applicable

Availability of data and materials

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

Competing interests

The authors declare that they have no competing interests

Funding

This work is part of the research about AI ethical and societal challenges under the “Hub Life Science – Digital Health (LSH-DH) PNC-E3-2022-23683267 - Project DHEAL-COM – CUP: D43C22005040001”, funded by the Italian Ministry of Health within the context of National Complementary Plan for the Innovative Ecosystem of Health - Unique investment code: PNC-E.3. The ministry assumes no responsibility for the contents of this publication.

Authors' contributions

The manuscript was collaboratively developed by SF, LS, and SR. SF authored the Aim of the Study (within the Background) and the Methods sections; LS contributed to the Study Intervention and Technological Tool subsection of the Methods; and SR drafted the remaining portions of the Background and the Discussion sections. All authors read and approved the final version of the manuscript.

Acknowledgements

The authors extend their gratitude to all those who contributed to the development of this protocol and its planned intervention, including written content, video resources, graphic elements, and the design of prompts intended for integration into the LLM-based chatbot described in this protocol. Their expertise and continued support were crucial in shaping the proposed digital protocol.

Bibliography

[1] Mental health at work n.d. https://www.who.int/news-room/fact-sheets/detail/mental-health-at-work (accessed November 12, 2025).

[2] Guidelines on mental health at work n.d. https://www.who.int/publications/i/item/9789240053052 (accessed November 12, 2025).

[3] Mental Health at Work: Policy Brief. 1st ed. Geneva: World Health Organization; 2022.

[4] Ballard DW, Lodge GC, Pike KM. Mental health at work: a practical framework for employers. Front Public Health 2025;13. https://doi.org/10.3389/fpubh.2025.1552981.

[5] WHO highlights urgent need to transform mental health and mental health care n.d. https://www.who.int/news/item/17-06-2022-who-highlights-urgent-need-to-transform-mental-health-and-mental-health-care (accessed November 26, 2025).

[6] Zhong W, Luo J, Zhang H. The therapeutic effectiveness of artificial intelligence-based chatbots in alleviation of depressive and anxiety symptoms in short-course treatments: A systematic review and meta-analysis. J Affect Disord 2024;356:459–69. https://doi.org/10.1016/j.jad.2024.04.057.

[7] Li H, Zhang R, Lee Y-C, Kraut RE, Mohr DC. Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. Npj Digit Med 2023;6:236. https://doi.org/10.1038/s41746-023-00979-5.

[8] Health at a Glance: Europe 2024. OECD 2024. https://www.oecd.org/en/publications/health-at-a-glance-europe-2024_b3704e14-en.html (accessed November 12, 2025).
[9] Empowering the health workforce to make the most of the digital revolution. OECD 2021. https://www.oecd.org/en/publications/empowering-the-health-workforce-to-make-the-most-of-the-digital-revolution_37ff0eaa-en.html (accessed November 12, 2025).
[10] Digital and innovative tools for better health and productivity at the workplace. OECD 2024. https://www.oecd.org/en/publications/digital-and-innovative-tools-for-better-health-and-productivity-at-the-workplace_2ce05eb6-en.html (accessed November 12, 2025).
[11] Purgato M, Carswell K, Tedeschi F, Acarturk C, Anttila M, Au T, et al. Effectiveness of Self-Help Plus in Preventing Mental Disorders in Refugees and Asylum Seekers in Western Europe: A Multinational Randomized Controlled Trial. Psychother Psychosom 2021;90:403–14. https://doi.org/10.1159/000517504.
[12] Purgato M, Carswell K, Acarturk C, Au T, Akbai S, Anttila M, et al. Effectiveness and cost-effectiveness of Self-Help Plus (SH+) for preventing mental disorders in refugees and asylum seekers in Europe and Turkey: study protocols for two randomised controlled trials. BMJ Open 2019;9:e030259. https://doi.org/10.1136/bmjopen-2019-030259.
[13] Lorenzo-Luaces L, Howard J, De Jesús-Romero R, Peipert A, Buss JF, Lind C, et al. Acceptability and Outcomes of Transdiagnostic Guided Self-help Bibliotherapy for Internalizing Disorder Symptoms in Adults: A Fully Remote Nationwide Open Trial. Cogn Ther Res 2023;47:195–208. https://doi.org/10.1007/s10608-022-10338-5.
[14] Fietta V, Rizzi S, Gios L, Pavesi MC, De Luca C, Gabrielli S, et al. World Health Organization Evidence-Based Self-Help Plus Intervention for Stress Management via Chatbot: Protocol for Adaptation to a Tech-Enabled Model. JMIR Res Protoc 2025;14:e69644. https://doi.org/10.2196/69644.
[15] Self Help Plus (SH+): A Group-Based Stress Management Course for Adults. Generic Field-trial Version 1. 0 2021. 1st ed. Geneva: World Health Organization; 2021.
[16] Fietta V, Rizzi S, De Luca C, Gios L, Pavesi MC, Gabrielli S, et al. A Chatbot-Based Version of the World Health Organization–Validated Self-Help Plus Intervention for Stress Management: Co-Design and Usability Testing. JMIR Hum Factors 2024;11:e64614. https://doi.org/10.2196/64614.
[17] Hayes SC, Luoma JB, Bond FW, Masuda A, Lillis J. Acceptance and Commitment Therapy: Model, processes and outcomes. Behav Res Ther 2006;44:1–25. https://doi.org/10.1016/j.brat.2005.06.006.
[18] Hayes SC, Strosahl K, Wilson KG, Strosahl K. Acceptance and commitment therapy: an experiential approach to behavior change. New York, NY London: Guilford Press; 1999.
[19] Sanna L, Bolpagni M, Fietta V, Gavioli G, Franzin M, Dragoni M, et al. LLM-Enriched Finite-State Chatbots for Mental Health Support: A Case Study on Self-Help+. In: Bellazzi R, Juarez Herrero JM, Sacchi L, Zupan B, editors. Artif. Intell. Med., vol. 15734, Cham: Springer Nature Switzerland; 2025, p. 386–96. https://doi.org/10.1007/978-3-031-95838-0_38.
[20] Czajkowski SM, Powell LH, Adler N, Naar-King S, Reynolds KD, Hunter CM, et al. From ideas to efficacy: The ORBIT model for developing behavioral treatments for chronic diseases. Health Psychol 2015;34:971–82. https://doi.org/10.1037/hea0000161.
[21] Acarturk C, Uygun E, Ilkkursun Z, Carswell K, Tedeschi F, Batu M, et al. Effectiveness of a WHO self-help psychological intervention for preventing mental disorders among Syrian refugees in Turkey: a randomized controlled trial. World Psychiatry 2022;21:88–95. https://doi.org/10.1002/wps.20939.
[22] Tol WA, Leku MR, Lakin DP, Carswell K, Augustinavicius J, Adaku A, et al. Guided self-help to reduce psychological distress in South Sudanese female refugees in Uganda: a

cluster randomised trial. Lancet Glob Health 2020;8:e254–63. https://doi.org/10.1016/S2214-109X(19)30504-2.
[23] 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:e72. https://doi.org/10.2196/mhealth.5849.
[24] Morselli S, Sebastianelli A, Domnich A, Bucchi C, Spatafora P, Liaci A, et al. Translation and validation of the Italian version of the user version of the Mobile Application Rating Scale (uMARS). J Prev Med Hyg 2021:E243 Pages. https://doi.org/10.15167/2421-4248/JPMH2021.62.1.1894.
[25] Bolpagni M, Gabrielli S. Development of a Comprehensive Evaluation Scale for LLM-Powered Counseling Chatbots (CES-LCC) Using the eDelphi Method. Informatics 2025;12:33. https://doi.org/10.3390/informatics12010033.
[26] Osgood CE, Suci GJ, Tannenbaum PH. The Measurement of Meaning. University of Illinois Press; 1957.
[27] Kroenke K, Spitzer RL, Williams JBW, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics 2009;50:613–21. https://doi.org/10.1176/appi.psy.50.6.613.
[28] The World Health Organization-Five Well-Being Index (WHO-5) n.d. https://www.who.int/publications/m/item/WHO-UCN-MSD-MHE-2024.01 (accessed November 12, 2025).
[29] Bond FW, Hayes SC, Baer RA, Carpenter KM, Guenole N, Orcutt HK, et al. Preliminary psychometric properties of the Acceptance and Action Questionnaire-II: a revised measure of psychological inflexibility and experiential avoidance. Behav Ther 2011;42:676–88. https://doi.org/10.1016/j.beth.2011.03.007.
[30] Why You Only Need to Test with 5 Users. Nielsen Norman Group n.d. https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/ (accessed February 3, 2026).
[31] Faulkner L. Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 2003;35:379–83. https://doi.org/10.3758/BF03195514.
[32] Poster and oral presentations (In alphabetical order by first author). Psychol Health 2009;24:71–419. https://doi.org/10.1080/08870440903126421.
[33] Schoger LI. Coping with work-related stressors: does education reduce work-related stress? J Public Health 2025;33:1123–34. https://doi.org/10.1007/s10389-023-02070-5.
[34] Marsh E, Vallejos EP, Spence A. The digital workplace and its dark side: An integrative review. Comput Hum Behav 2022;128:107118. https://doi.org/10.1016/j.chb.2021.107118.
[35] Ray TK, Pana-Cryan R. Work Flexibility and Work-Related Well-Being. Int J Environ Res Public Health 2021;18:3254. https://doi.org/10.3390/ijerph18063254.
[36] Meske C, Junglas I. Investigating the elicitation of employees' support towards digital workplace transformation. Behav Inf Technol 2021;40:1120–36. https://doi.org/10.1080/0144929X.2020.1742382.
[37] Glazer S, Liu C. Work, Stress, Coping, and Stress Management. Oxf. Res. Encycl. Psychol., 2017. https://doi.org/10.1093/acrefore/9780190236557.013.30.