frontiers | Research Topics

Designing and evaluating digital health interventions

Edited by
Kim Cornelia Marie-Louise Bul, Nikki Holliday and Edith Talina Luhanga

Published in
Frontiers in Digital Health

ISSN 1664-8714
ISBN 978-2-8325-6630-5
DOI 10.3389/978-2-8325-6630-5

Generative AI statement
Any alternative text (Alt text) provided alongside figures in the articles in this ebook has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Designing and evaluating digital health interventions

Topic editors

Kim Cornelia Marie-Louise Bul — Coventry University, United Kingdom
Nikki Holliday — Coventry University, United Kingdom
Edith Talina Luhanga — Carnegie Mellon University Africa, Rwanda

Citation

Bul, K. C. M.-L., Holliday, N., Luhanga, E. T., eds. (2025). *Designing and evaluating digital health interventions*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-6630-5

Table of contents

frontiers | Frontiers in Digital Health

TYPE Editorial
PUBLISHED 08 May 2025
DOI 10.3389/fdgth.2025.1612380

OPEN ACCESS

EDITED AND REVIEWED BY
Toshiyo Tamura,
Waseda University, Japan

*CORRESPONDENCE
Kim Bul
ac2658@coventry.ac.uk

RECEIVED 15 April 2025
ACCEPTED 23 April 2025
PUBLISHED 08 May 2025

CITATION
Bul K, Holliday N and Luhanga ET (2025) Editorial: Designing and evaluating digital health interventions.
Front. Digit. Health 7:1612380.
doi: 10.3389/fdgth.2025.1612380

Editorial: Designing and evaluating digital health interventions

Kim Bul[1]*, Nikki Holliday[2] and Edith Talina Luhanga[3]

[1]Centre for Intelligent Healthcare, Coventry University, Coventry, United Kingdom, [2]Centre for Healthcare and Communities, Coventry University, Coventry, United Kingdom, [3]Carnegie Mellon University Africa, Kigali, Rwanda

KEYWORDS

digital, health, development, evaluation, interventions, technologies

Editorial on the Research Topic
Designing and evaluating digital health interventions

Background

The World Health Organization (1) defines digital health as "*the systematic application of information and communications technologies, computer science, and data to support informed decision-making by individuals, the health workforce, and health systems, to strengthen resilience to disease and improve health and wellness*". Although the COVID-19 pandemic accelerated development of DHIs internationally, there remains a divide between high-income countries showing a wealth of personalized and immersive health platforms while initiatives across low-and-middle income countries (LMICs) are still limited given their priority on mobile technologies and wireless connectivity due to limited infrastructure, resources and focus on basic healthcare needs (2). Hence, the academic literature is often split between examining effectiveness of these interventions in high-income countries vs. LMICs. Additionally, there is consensus that there is disparity in DHI access across groups such as women, migrants and older people (3). Despite a greater focus on women's health (e.g., perinatal mental health) and gender-specific DHIs, their healthcare is hindered due to unequal access and usage (4, 5). The Medical Research Council Framework for Complex Interventions stresses that well-designed interventions should consider theoretical frameworks of behaviour change, human-centred design through patient/public involvement and engagement (PPIE) as well as co-design, usability, feasibility and robust evaluation through trials to prove effectiveness in improving people's health (6). In response to these global disparities and methodological challenges, the current Research Topic aims to advance the field by showcasing rigorous and inclusive approaches to the design and evaluation of DHIs.

Aims and objectives

The aim of the current Research Topic was to establish a collection of high-quality diverse manuscripts representing the dynamic academic field of DHIs. Scientific

evidence on DHIs is continuously evolving due to its importance and relevance while there often is a lack of funding, methodological guidance and inclusivity of the digital healthcare landscape. With two systematic reviews, eight original articles, two brief research reports, one study protocol, and one curriculum, instruction and pedagogy article the current collection gives an up-to-date, multidisciplinary and widespread contribution to the field of DHIs, specifically in the area of intervention design and evaluation.

The current research topic

Synthesizing the results of 30 RCTs, small to moderate effects of DHIs on mental health were demonstrated by Morello et al. However, due to the complexity of interventions it was not clear what the effective ingredients (e.g., enhancing cognitive reappraisal) of the interventions were. To promote better reporting of complex interventions and to support identification of working mechanisms it is recommended that interventions are described using the TIDieR checklist (7). Furthermore, to design a theoretically sound and effective DHI it is important to incorporate behaviour change strategies (e.g., self-monitoring). The brief research report from Ghantasala et al. explores tailored motivation messages as a strategy to increase engagement and physical activity. The report indicates that personalized messages adjusted to mood, self-efficacy and progress are perceived as more motivating compared to generic messages. Lakha et al. used a case study to demonstrate how a digital scrapbook format can be a feasible and accessible format for patients with chronic pain and illness. Increasing engagement across DHIs is important, as dropouts can be high [up to 82% (8)]. Gamified approaches are used as an engaging way to increase knowledge and health outcomes. In the case of the pilot study reported by Seaver et al. a game-based learning approach was used to educate people about sleep hygiene and improve outcomes such as sleep quality and anxiety. The digital therapeutics mobile app from Jeong et al. demonstrated a gamified approach to facilitate breath control. This included machine learning and biofeedback visualization to improve engagement.

Given research funding is often limited it is important to examine cost-effective ways to design and evaluate DHIs in a research context and potentially collaborate with industry initiatives. Instead of developing DHIs from scratch there are opportunities to repurpose existing interventions to improve health outcomes (9). The study from Chen et al. examined a mobile health intervention called MyTrack + - paired with existing commercials apps and weight scales—to support

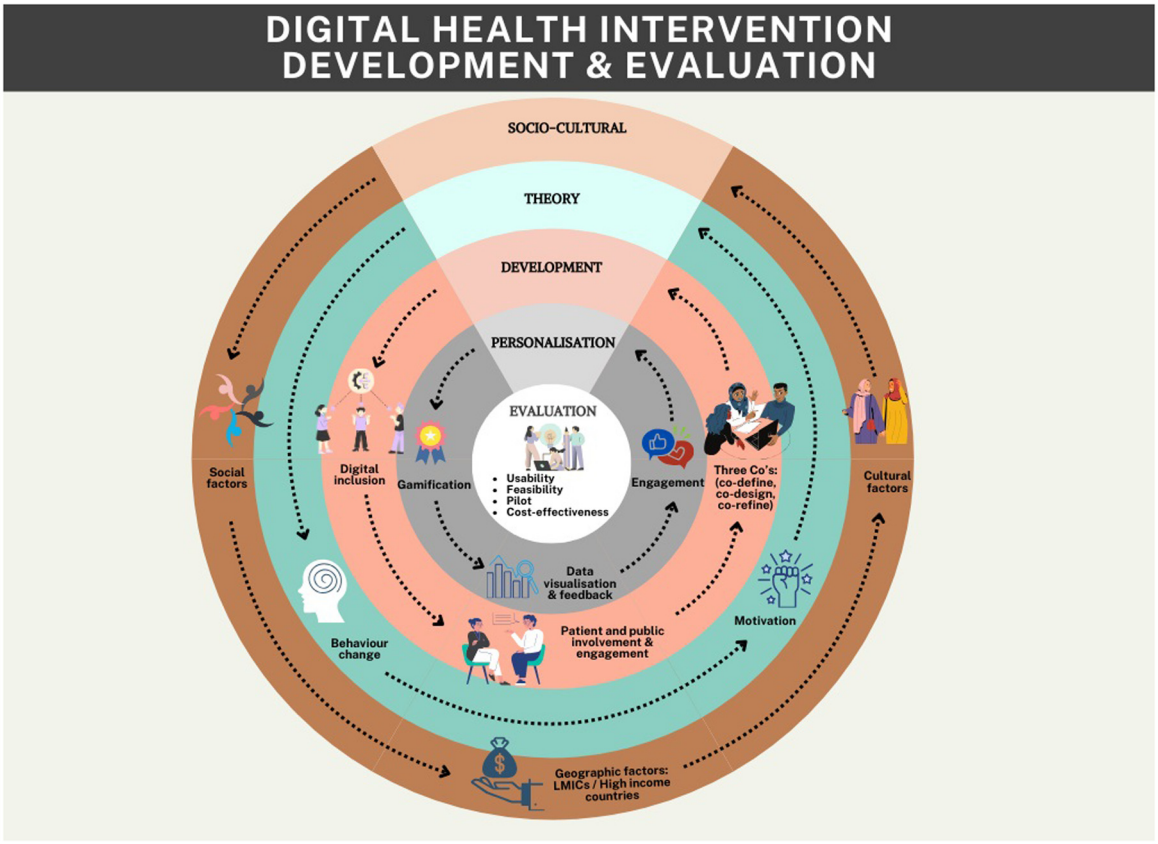


FIGURE 1
Relevant research domains in digital health intervention development and evaluation.

long-term weight loss maintenance. Small usability and pilot studies were performed to iteratively incorporate end-user feedback while keeping R&D processes efficient. For DHI design, it is important to follow established frameworks like MRC or co-design frameworks such as the three Co's framework (10, 11) to ensure the intervention fits with end-user needs. The study of Hamaguchi et al. focuses on a novel smartphone app to support patients with mild cognitive impairment and dementia. Although initial positive results were found in terms of engagement, further multi-centre randomized controlled trials are recommended to evidence improvement of other health outcomes (e.g., cognition). The study protocol developed by Castelnuovo et al. describes a clinical RCT to demonstrate efficacy of an innovative digital therapy to promote weight loss in patients with obesity by increasing their treatment adherence. Complementing trial data with process evaluation, as done by Ali et al., is crucial to gain a deeper understanding of participant experiences on how, why and when people are engaging with these tools. However, it must be noted that the field of DHIs lacks methodological guidance on evaluating its effectiveness beyond RCTs (12) and is quite segmented due to its inter-disciplinary nature and approach. Implementation science is an important aspect of DHIs and goes beyond the development of guidelines on how to implement interventions in clinical practice, for example through educating future healthcare professionals on this. The curriculum, instruction and pedagogy article from Loizou et al. uses Virtual Reality simulations for healthcare professionals to practice carer and patient interactions based on affective intelligent agents incorporated into the learning scenarios.

Healthcare access remains a global challenge and DHIs, such as telehealth, can support in streamlining provision and reducing waiting times (13). The study of Tennankore et al. describes a pilot study examining the potential of the Virtual Hallway platform in facilitating patient-focused specialist care through synchronous phone conversations. This demonstrated high acceptability and potential to reduce waitlists and unnecessary referrals. The qualitative case study approach from Sowon et al. refers to a "community of purpose" and how they affect the usage of mobile health interventions in the context of maternal care in Sub-Saharan Africa context. Although results may not be generalisable to high-income countries, it provides insights into the complexities of promoting use and adoption of DHIs across similar economic and cultural contexts. Another important topic regarding healthcare access in the field of DHIs is the topic of digital exclusion. The systematic review of Udenigwe et al. focussed on gender transformative approaches in mobile health interventions for maternal health in sub-Saharan Africa, primarily consisting of text-based messaging. It stresses the limited extent of such approaches and highlights the need for inclusivity in the digital landscape in maternal health across LMICs. Inclusivity should be considered for any DHIs across any context and research stage, including high-income countries (14). The brief research report article from Collombon et al. focuses on recruiting adults aged 50 years and older with low socioeconomic status for participation in online physical activity interventions. As stressed by the report, personal paper-based invitation letters worked best compared to social media and advertisements through a gym. It is important to use inclusive recruitment strategies to maximize diversity and ensure equal access to the research study, specifically when the end-users are hard to reach.

Future research recommendations

The field of DHIs is developing rapidly with known potential to tackle healthcare challenges. However, challenges remain in terms of accessibility, engagement and evaluation methodologies across different patient populations and economic settings. The contributions in this collection demonstrate the importance of incorporating behavioural science, human-centered design, and rigorous evaluation frameworks to enhance effectiveness. Inclusivity must be prioritized from intervention design to recruitment strategies, to ensure equitable access for underserved groups. Interdisciplinary collaborations between academia, industry and healthcare providers are crucial and future research should explore cost-effective and scalable solutions, leveraging AI, machine learning, and gamification while maintaining a focus on ethical considerations and end-user needs. The current research topic presents a global and interdisciplinary perspective, furthering our understanding of the field and shaping future research directions and recommendations (Figure 1).

Dr. Kim Bul
Assistant Professor, Research Centre for Intelligent Healthcare, Coventry University
Nikki Holliday
Design Manager, Research Centre for Healthcare and Communities, Coventry University
Dr. Edith Talina Luhanga, Carnegie Mellon University Africa

Author contributions

KB: Writing – original draft, Writing – review & editing. NH: Writing – original draft, Writing – review & editing. EL: Writing – original draft, Writing – review & editing.

Acknowledgments

We would like to acknowledge Dr. Akanksha Singh for supporting the editors in designing Figure 1 through the use of Canva software.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. World Health Organization. WHO guideline: recommendations on digital interventions for health system strengthening. (2023).

2. Makri A. Bridging the digital divide in health care. *Lancet Digit Health*. (2019) 1(5):e204–5. doi: 10.1016/S2589-7500(19)30111-6

3. Kaihlanen AM, Virtanen L, Buchert U, Safarov N, Valkonen P, Hietapakka L, et al. Towards digital health equity-a qualitative study of the challenges experienced by vulnerable groups in using digital health services in the COVID-19 era. *BMC Health Serv Res*. (2022) 22(1):188. doi: 10.1186/s12913-022-07584-4

4. Acilar A, Sæbø Ø. Towards understanding the gender digital divide: a systematic literature review. *Glob Know Mem Commun*. (2023) 72(3):233–49. doi: 10.1108/GKMC-09-2021-0147

5. Vuijk P, Bul K, Buil M, Rauws M, Curie K, Amesz C, et al. Effectiveness of a blended school-based mindfulness program for the prevention of co-rumination and internalizing problems in Dutch secondary school girls: a cluster randomized controlled trial. *Trials*. (2024) 25(1):40. doi: 10.1186/s13063-023-07885-x

6. Shahsavari H, Matourypour P, Ghiyasvandian S, Nejad MRG. Medical research council framework for development and evaluation of complex interventions: a comprehensive guidance. *J Educ Health Promot*. (2020) 9(1):88. doi: 10.4103/jehp.jehp_649_19

7. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *Br Med J*. (2014) 348:1–12. doi: 10.1136/bmj.g1687

8. Bricker J, Miao Z, Mull K, Santiago-Torres M, Vock DM. Can a single variable predict early dropout from digital health interventions? Comparison of predictive models from two large randomized trials. *J Med Internet Res*. (2023) 25:e43629. doi: 10.2196/43629

9. Bul K, Holliday N, Magee P, Wark P. From development to exploitation of digital health solutions: lessons learnt through multidisciplinary research and consultancy. *J Enab Technol*. (2020) 14(4):273–9. doi: 10.1108/JET-09-2020-0035

10. Pearce G, Magee P. Co-creation solutions and the three co's framework for applying co-creation. *Health Educ*. (2024) 124(1/2):20–37. doi: 10.1108/HE-09-2022-0077

11. Pearce G, Holliday N, Sandhu H, Eftekhari H, Bruce J, Timms E, et al. Co-creation of a complex, multicomponent rehabilitation intervention and feasibility trial protocol for the PostUraL tachycardia syndrome exercise (PULSE) study. *Pilot Feasibil Stud*. (2023) 9:143. doi: 10.1186/s40814-023-01365-4

12. Hrynyschyn R, Prediger C, Stock C, Helmer SM. Evaluation methods applied to digital health interventions: what is being used beyond randomised controlled trials?—a scoping review. *Int J Environ Res Public Health*. (2022) 19(9):5221. doi: 10.3390/ijerph19095221

13. Caffery LJ, Farjian M, Smith AC. Telehealth interventions for reducing waiting lists and waiting times for specialist outpatient services: a scoping review. *J Telemed Telecare*. (2016) 22(8):504–12. doi: 10.1177/1357633X16670495

14. Krukowski RA, Ross KM, Western MJ, Cooper R, Busse H, Forbes C, et al. Digital health interventions for all? Examining inclusivity across all stages of the digital health intervention research process. *Trials*. (2024) 25(1):98. doi: 10.1186/s13063-024-07937-w

frontiers | Frontiers in Digital Health

TYPE Original Research
PUBLISHED 12 September 2023
DOI 10.3389/fdgth.2023.1243253

OPEN ACCESS

EDITED BY
Nikki Holliday,
Coventry University, United Kingdom

REVIEWED BY
Beenish Chaudhry,
University of Louisiana at Lafayette, United States
Martin Dyrba,
Helmholtz Association of German Research Centers (HZ), Germany

*CORRESPONDENCE
Reo Hamaguchi
reo-h@nifty.com

RECEIVED 20 June 2023
ACCEPTED 29 August 2023
PUBLISHED 12 September 2023

CITATION
Hamaguchi R, Hirokawa Y, Takahashi H, Hachiya T, Kawano H, Isotani S, Ito E, Handa N, Saito R, Horie S and Ide H (2023) Retrospective observational study of a novel smartphone app on the management of patients with mild cognitive impairment or mild dementia. Front. Digit. Health 5:1243253. doi: 10.3389/fdgth.2023.1243253

Retrospective observational study of a novel smartphone app on the management of patients with mild cognitive impairment or mild dementia

Reo Hamaguchi[1]*, Yoshihiro Hirokawa[2], Hirotsugu Takahashi[1], Tsuyoshi Hachiya[3], Haruna Kawano[4], Shuji Isotani[4], Emi Ito[1], Nobuhiro Handa[1], Ryozo Saito[5], Shigeo Horie[1] and Hisamitsu Ide[1]

[1]Department of Digital Therapeutics, Graduate School of Medicine, Juntendo University, Tokyo, Japan, [2]Hirokawa Clinic, Kyoto, Japan, [3]Department of Advanced Informatics for Genetic Disease, Graduate School of Medicine, Juntendo University, Tokyo, Japan, [4]Department of Urology, Graduate School of Medicine, Juntendo University, Tokyo, Japan, [5]Life Quest Inc., Tokyo, Japan

Introduction: In this study, we aimed to evaluate the feasibility, utility, and potential effects of LQ-M/D App, a smartphone application developed by Life Quest Inc., Tokyo, Japan, for patients with mild cognitive impairment (MCI) and mild dementia. The app incorporates cognitive and physical exercise training, lifestyle habit acquisition features, and a continuity improvement feature added in the post-update version to enhance user engagement. The continuity improvement feature includes the optimization of training content, and disease education, and enables family monitoring via a family app.
Methods: A retrospective analysis was conducted on app usage, cognitive and exercise training implementation and interruptions, questionnaire response rates, and cognitive assessments in a single institution. A total of 20 patients used the app, with 10 patients using the pre-update version without the continuity improvement feature, and the other 10 patients using the post-update version with the continuity improvement feature.
Results and Conclusion: The results demonstrated that the LQ-M/D App could be effectively used by the study population, and the continuity improvement feature positively influenced app usage in several aspects. Although a potential association between app usage and cognitive ability was suggested, the scatter in the data points warrants cautious interpretation. Limitations of the study included a small sample size, a single institution setting, and the retrospective nature of the study. In the future, a randomized controlled trial design using a larger sample size and multiple institutions to further evaluate the effectiveness of LQ-M/D App in managing MCI and mild dementia should be performed.

KEYWORDS

smartphone app, digital therapy, nonpharmacological intervention, mild cognitive impairment, dementia

Introduction

The prevalence of dementia in Japan is increasing, and the number of people in Japan with dementia in 2025 is estimated to be 6.75 million (1). Although pharmacotherapy is thought to alleviate symptoms in patients who have developed dementia, its effectiveness in halting or slowing the progression of the disease is very limited. Therefore,

interventions are recommended to be performed in patients with mild cognitive impairment (MCI) before the onset of dementia or in the early stages of dementia. The prevalence of MCI is reported to be 15%–25% in those aged 65 years and older (2–4), and the conversion rate from MCI to dementia is reported to be approximately 5%–15% per year (5). On the other hand, MCI-to-normal reversion has also been reported to be 16%–41% per year (6), and although it is controversial whether such improvement can actually be observed, it is considered important to actively address this from the time a patient has MCI, to prevent subsequent dementia. Cholinesterase inhibitors used for dementia have been reported to improve cognitive function in MCI patients; however, they have not been shown to prevent the progression to dementia (7). Non-drug therapies centered on moderate physical activity and lifestyle modification have been recommended to reduce the progression from MCI to dementia (8–10). In addition, lifestyle-associated diseases, such as hypertension, diabetes, dyslipidemia, and a history of cerebrovascular disease are known to be risk factors for the progression from MCI to dementia, and appropriate management of these risk factors is important (10, 11). Management of these risk factors requires that the patient and family members have an appropriate understanding of MCI, and they receive guidance to prepare for the progression to dementia, including understanding and improving lifestyle habits.

In recent years, the field of mobile health using smartphones and mobile phones has gained attention, and is beginning to be used for older adults (12, 13), including those with MCI (14, 15). However, treatment-focused apps designed specifically for patients with MCI or dementia are limited at present. Cognitive Assessment for Dementia, iPad version (CADi), and the Cogstate Brief Battery (CBB) are examples of apps intended for the assessment of cognitive function (16, 17). CADi can be used as a new mass screening tool that comprises memory and cognitive tasks, and has demonstrated good reliability and validity for the screening of patients with dementia (16). CBB, on the other hand, has shown promise for unsupervised at-home cognitive assessments, but it primarily focuses on evaluation, rather than intervention (17). Exercise-based interventions are expected to slow cognitive decline as non-pharmacological therapies (18), and it was reported that a smartphone app called "HealtheBrain" can be used without any problems in older adults with or without MCI (19). This app, designed to provide a visuospatial memory exercise called the square stepping exercise, aims to improve cognitive function in older adults. In addition, cognitive training, traditionally conducted on paper but now often conducted using computers, has been suggested to be effective for cognitive function (20). A digital calendar with activity/event reminders via short messages that can be used on smartphones or mobile phones, which can be set by family members, professionals, or other supporters, has been used by older adults aged 65 years and older (including older adults who are aware of memory loss), and it was reported that the use of this system can be learned easily (21).

As described above, it has been reported that providing exercises, scheduling, and obtaining activity records via smartphones can be used without problems by older patients, including those with cognitive impairment. In addition, computer-based cognitive training can be substituted by mobile devices, such as smartphones. A smartphone application for patients with MCI or mild dementia called "LQ-M/D App" was developed by Life Quest Inc. (Tokyo, Japan), which includes the following functions: (i) training content (cognitive training and exercise training), and (ii) content for the acquisition of lifestyle habits (diet, sleep, and walking time). At present, there are a limited number of treatment-focused apps specifically tailored for patients with MCI or dementia. The LQ-M/D App is specifically designed for older adults with cognitive impairments to provide a non-pharmacological intervention. The purpose of this study was to investigate the feasibility and utility of the LQ-M/D App, and we conducted a retrospective observational study to analyze the clinical and application data of patients with MCI or mild dementia who used LQ-M/D App.

Patients and methods

Study design

This observational study was retrospectively conducted to investigate the effects of LQ-M/D App on patients with MCI or mild dementia treated at Hirokawa Clinic, Kyoto, Japan, from September 1, 2020, to August 31, 2021, using the clinic's medical records and data from the app. All patients were provided with smartphones containing the pre-downloaded LQ-M/D App, and used the LQ-M/D App, as well as received standard treatments for MCI or mild dementia. All procedures were performed in accordance with the ethical principles described in the 1995 Declaration of Helsinki. Written informed consent was obtained from each patient. The study was approved by the Institutional Review Board of Juntendo University and registered in UMIN Clinical Trials (UMIN000047077). In this study, external monitoring by the Japan Organization for Research and Treatment of Cancer (JORTC) verified data accuracy by comparing original sources with analysis data for a random subset of cases, with no discrepancies identified, confirming data integrity.

LQ-M/D App

In this study, the smartphone application "LQ-M/D App" developed by Life Quest Inc. was used. LQ-M/D App has two main features: (i) provision of training content, and (ii) content for the acquisition of lifestyle habits (**Figure 1**).

The training content consists of cognitive function training and physical exercise training. The cognitive training consists of multiple menus targeting various cognitive functions, including memory function, language function, judgment function, calculation function, executive function, and visuospatial cognitive function, with a total of 21 different tasks. The physical exercise training comprises dual-task training and other types of training, and includes 38 dual-task exercises and 24 other types of physical activities. The lifestyle habits acquisition feature collects

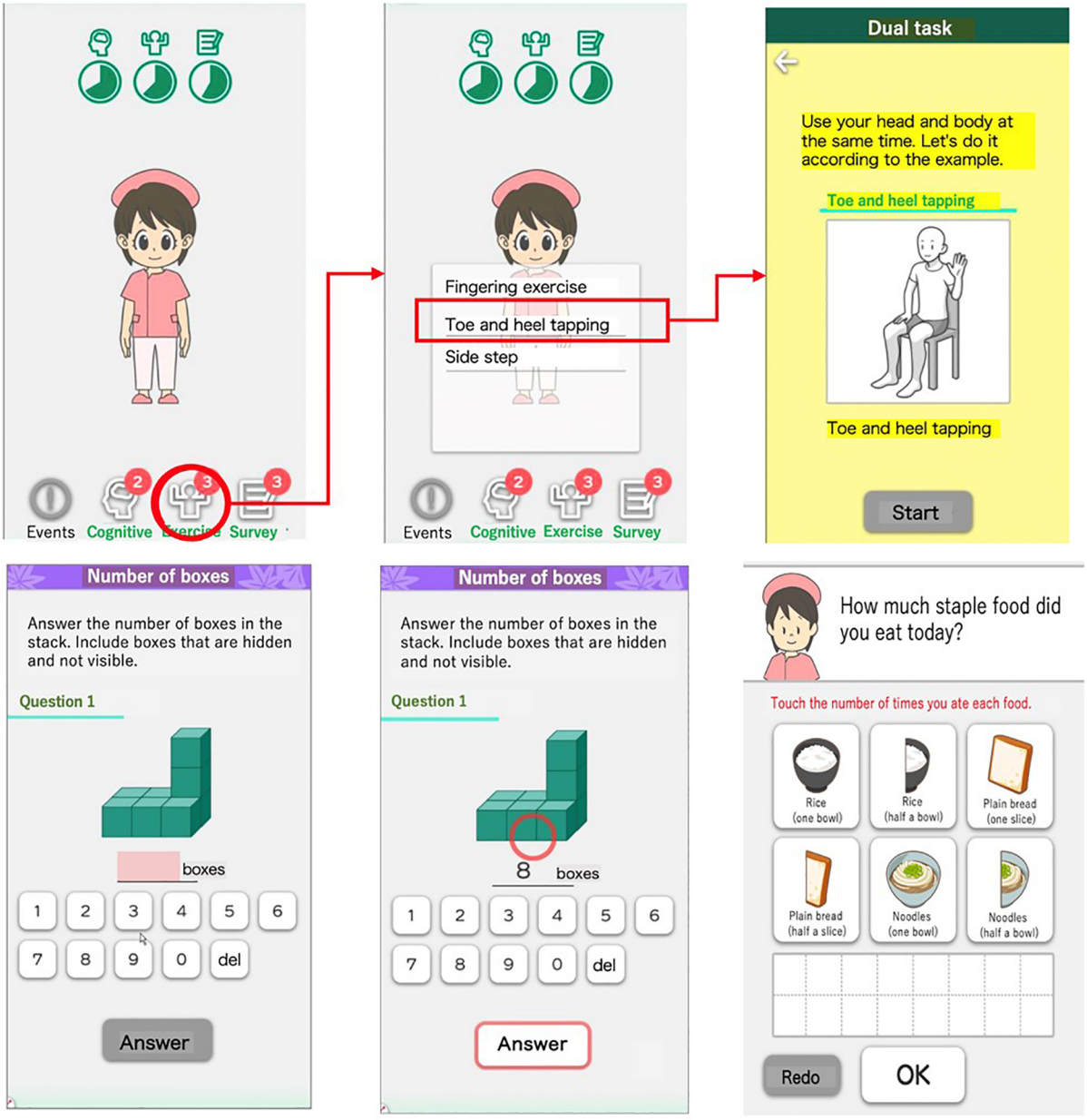


FIGURE 1
Overview of LQ-M/D App main features. Screenshots of LQ-M/D App, showcasing its main features: (i) training content, and (ii) acquisition of lifestyle habits. Images depict examples of exercise training (top-right), cognitive training (bottom-left and center-bottom), and the dietary questionnaire (bottom-right).

information on diet, sleep, and walking time through questionnaires conducted within the app. However, at present, the app does not have a feature for physicians to review records and provide feedback.

On December 21, 2020, LQ-M/D App was updated to include a continuity improvement feature designed to enhance sustained user engagement. This feature consists of optimizing training content, providing disease education, and enabling family monitoring via a family app (Figure 2). The optimization of training content adjusts and reconfigures the content based on usage, with varying intervention frequencies by an avatar. Disease education offers valuable information and resources to help patients and their families better understand and manage the condition. The family app enables monitoring of the patient's app usage, and enables simple message exchanges between patients and their family members. As a result, families can better understand the patient's situation, and an improvement in the patient's app usage can be expected. A summary of the features of the LQ-M/D App is shown in Table 1.

In this study, we used both the pre-update version without the continuity improvement feature, and the post-update version with the continuity improvement feature, each for a duration of 6 months, to compare and investigate their respective effects. Patients and caregivers were provided with comprehensive instructions on how to use the app. During regular monthly

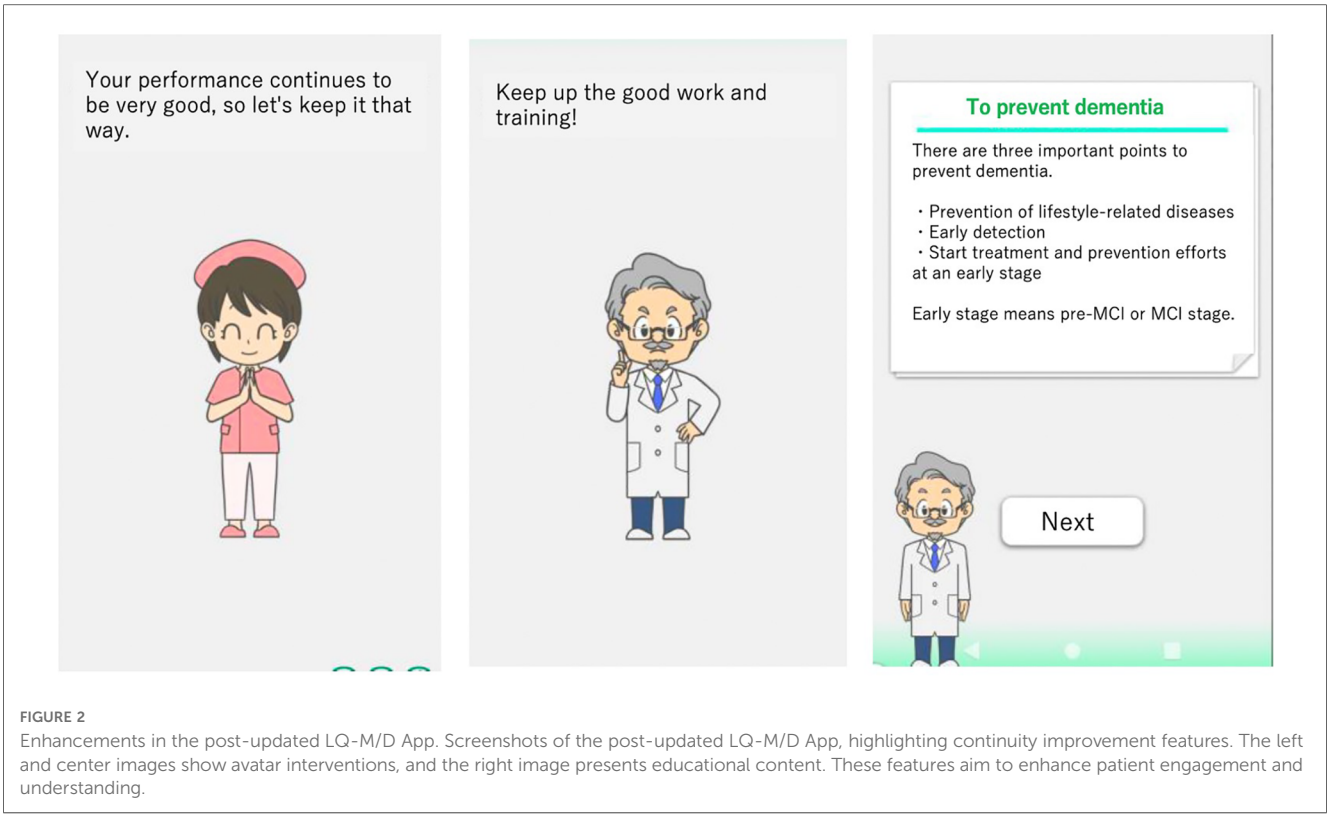


FIGURE 2
Enhancements in the post-updated LQ-M/D App. Screenshots of the post-updated LQ-M/D App, highlighting continuity improvement features. The left and center images show avatar interventions, and the right image presents educational content. These features aim to enhance patient engagement and understanding.

TABLE 1 Summary of features of the LQ-M/D App.

Feature	Pre-update version	Post-update version
Training content		
Cognitive function training[a]	○	○
Physical exercise training[a]	○	○
Lifestyle habits acquisition		
Collection of information on diet, sleep, and daily walking time[a]	○	○
Continuity improvement feature		
Optimizing training content[a]	×	○
Providing disease education[a]	×	○
Family monitoring via family app[b]	×	○

[a]All features except for the family monitoring feature are utilized by patients through the LQ-M/D App.
[b]The family monitoring feature is utilized by family members through the family app.

check-ups, healthcare providers asked about app usage to evaluate its use by the patients and caregivers. However, the actual usage of the app was ultimately at the discretion of the users themselves.

Assessment procedures

In this study, data from medical records were retrospectively collected, including patient background, MCI and dementia treatment courses, and scores for Mini-Mental State Examination (MMSE), Hasegawa's Dementia Scale-Revised (HDS-R), Clinical Dementia Rating (CDR), and Alzheimer's Disease Assessment Scale-Cognitive Subscale Japanese version (ADAS-jcog). Additionally, information on app malfunctions occurring during use, daily app usage duration (in minutes), the daily number of cognitive training and exercise training sessions performed, daily number of interruptions of cognitive training and exercise training, and the presence or absence of daily questionnaire responses were obtained from the app data. The data were compared between the pre-update version without the continuity improvement feature and the post-update version with the continuity improvement feature. Furthermore, MMSE scores of 24–27 points were classified as MCI, and those of 20–23 points as mild dementia. For these patients, the association between daily app usage duration (in minutes), the daily number of cognitive training and exercise training sessions, and changes in MMSE, HDS-R, and ADAS-jcog scores from the start of app use to 6 months after the end of app usage were examined.

Statistical analyses

For both the pre-update version without the continuity improvement feature and the post-update version with the continuity improvement feature, unpaired *t*-tests were conducted for daily app usage duration (in minutes), daily number of cognitive training and exercise training sessions, and daily number of interruptions for cognitive training and exercise training. Additionally, the Mann-Whitney *U*-test was performed to compare the presence or absence of daily questionnaire responses between the pre-update and post-update versions.

10.3389/fdgth.2023.1243253

Pearson's correlation coefficients were measured between the changes in MMSE, HDS-R, and ADAS-jcog scores (6-month score—initial score) and daily app usage duration, and daily number of cognitive training and exercise training sessions. Standard deviations of mean dataset values were calculated. All *p*-values were two-sided, and a *p*-value of less than 0.05 was considered to indicate a statistically significant difference between two groups. All statistical analyses were performed using Easy R software (version 1.61; Saitama Medical Center, Jichi Medical University, Saitama, Japan), a graphical user interface that is a modified version of R (The R Foundation for Statistical Computing, Vienna, Austria) (22).

Results

Patient characteristics

A total of 20 patients (7 men and 13 women) used LQ-M/D App, with a mean age of 72.3 (range: 50–82) years at the start of app usage. The MMSE, HDS-R, CDR, and ADAS-jcog scores at the start of app usage are shown in Table 2. The mean daily app usage duration (minutes), mean daily number of cognitive training and exercise training sessions, mean daily number of interruptions, and the proportion of questionnaire responses are presented in Table 3 for the entire study period as well as from the first month to the sixth month.

TABLE 2 Patient characteristics.

Number of patients	20
Age, years	
Mean (range)	72.3 (50–82)
Sex	
Men	7
Women	13
MMSE score	24.7 ± 3.0
HDS-R score	24.4 ± 3.7
CDR score	0.45 ± 0.15
CDR category (number of patients)	
0	2
0.5	18
1	0
2	0
ADAS-jcog score	10.4 ± 3.8
Medication (anti-dementia drugs)	
Yes	17
No	3

MMSE, mini-mental state examination; HDS-R, Hasegawa's dementia scale-revised; CDR, clinical dementia rating; ADAS-jcog, Alzheimer's disease assessment scale-cognitive subscale Japanese version.

App malfunctions

During usage of the app, 11 patients experienced malfunctions. In 4 cases of patients using the pre-updated version, the training feature of the app was unstable for 36, 34, 26, and 8 days, respectively. In 6 cases of patients using the post-updated version, the app could not be launched owing to an inability to communicate with the server for 14, 13, 12, 8, 6, 5, and 4 days, respectively.

Comparison between the pre-update version and the post-update version

In this analysis, the pre-update version without the continuity improvement feature ($n = 10$) was compared with the post-update version with the continuity improvement feature ($n = 10$). For each version, the mean daily app usage time, mean number of cognitive training and exercise training sessions, mean number of interruptions, and presence or absence of questionnaire responses were shown for the entire period and from the first to the sixth month, in Figures 3–6.

Over the entire 6-month observation period, there was no statistically significant difference in mean daily app usage time between the pre-update and post-update versions. Examining the monthly results, the mean usage time tended to decrease over time in the pre-update version, whereas the post-update version did not show a clear decline in mean usage time, and at 6 months, the mean daily app usage time was significantly longer for the post-update version (Figure 3).

Regarding the mean number of daily cognitive training sessions, there was a tendency for the pre-update version to have more sessions, with a significantly higher number of sessions for

TABLE 3 Summary of App usage, training sessions, interruptions, and response rates.

	App usage duration (minutes/day)	Cognitive training		Exercise training		Questionnaire
		Sessions (times/day)	Interruptions (times/day)	Sessions (times/day)	Interruptions (times/day)	Response rate per day (%)
Observation period (6 months)	38.2 ± 41	7.6 ± 10.7	0.125 ± 0.420	3.8 ± 5.3	0.159 ± 0.517	21.7 ± 42.1
1st month	44.9 ± 39.5	8.2 ± 9.3	0.152 ± 0.489	4.0 ± 4.7	0.249 ± 0.784	22.4 ± 44.1
2nd month	39.1 ± 45.7	6.6 ± 9.5	0.118 ± 0.382	4.5 ± 7.0	0.200 ± 0.589	23.8 ± 42.9
3rd month	42.1 ± 46.8	8.9 ± 13.1	0.134 ± 0.390	4.5 ± 5.6	0.158 ± 0.420	23.4 ± 43.9
4th month	34.9 ± 40.4	7.6 ± 12.3	0.135 ± 0.446	3.2 ± 4.8	0.123 ± 0.373	19.6 ± 41.5
5th month	35.7 ± 35.9	7.8 ± 10.0	0.100 ± 0.386	3.5 ± 4.8	0.113 ± 0.354	21.6 ± 40.8
6th month	29.3 ± 31.1	6.3 ± 8.5	0.100 ± 0.400	2.8 ± 4.0	0.077 ± 0.292	17.9 ± 37.0

Data are shown as mean ± standard deviation.

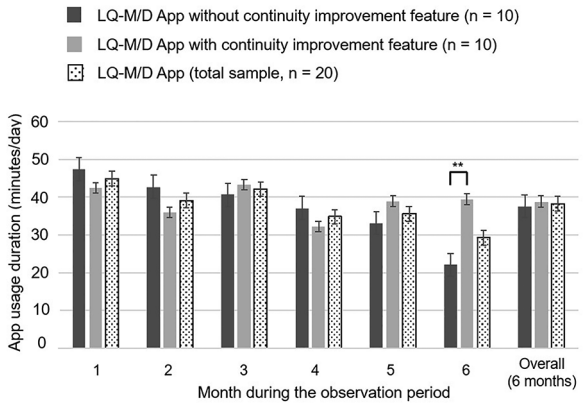


FIGURE 3
LQ-M/D App usage duration. Usage duration of LQ-M/D App ($n = 20$), LQ-M/D App without the continuity improvement feature ($n = 10$), and LQ-M/D App with the continuity improvement feature ($n = 10$), from the first to the sixth month, as well as the entire observation period (6 months). The graph shows the mean values for app usage duration in each group, with error bars representing the standard error. *$p < 0.05$, **$p < 0.005$.

the overall observation period (6 months), as well as for the first and fourth months. On the other hand, at 6 months, the number of sessions decreased in the pre-update version, and a significant increase in the number of sessions was observed in the post-update version (Figure 4A). For the mean number of interruptions, a statistically significant difference was observed between the pre-update and post-update versions for the overall observation period (6 months) and in each month, with more interruptions in the pre-update version (Figure 5A). The mean number of daily exercise training sessions was significantly higher for the post-update version during the overall observation period of 6 months and in each month, except for the second month (Figure 4B). Additionally, the mean number of interruptions was lower for the post-update version, with a statistically significant difference observed for the overall observation period of 6 months and each month, except for the first and 6 months (Figure 5B).

The mean daily questionnaire response rate tended to be lower for the post-update version, except for the sixth month. A significantly higher mean questionnaire response rate was observed for the pre-update version during the overall 6-month observation period, as well as at the first, second, and fourth months (Figure 6). However, it is difficult to compare the data, as the number of questionnaires differed between the two versions; i.e., in the pre-update version, questionnaires were conducted after breakfast, lunch, and dinner, whereas in the post-update version, questionnaires were additionally conducted upon waking and before bedtime.

Association between scores for MMSE, HDS-R, and ADAS-jcog, and app usage

In this study, patients with MMSE scores of 24–27 were classified as having mild MCI, and those with scores of 20–23 were classified as having mild dementia. Four patients with MMSE scores of 28 or higher and 1 patient with a score of 19 or lower were excluded. The remaining sample comprised 11 patients with MCI and 4 with mild dementia, totaling 15 patients. The association between daily app usage duration (in minutes), the daily number of cognitive training and exercise training sessions, and changes in MMSE, HDS-R, and ADAS-jcog scores from the start of app use to 6 months after the end of app use was investigated for these 15 patients using Pearson's correlation coefficients. Scatterplots of the associations are shown in Figure 7. The results indicate positive correlations between MMSE and HDS-R cognitive assessments and factors, such as

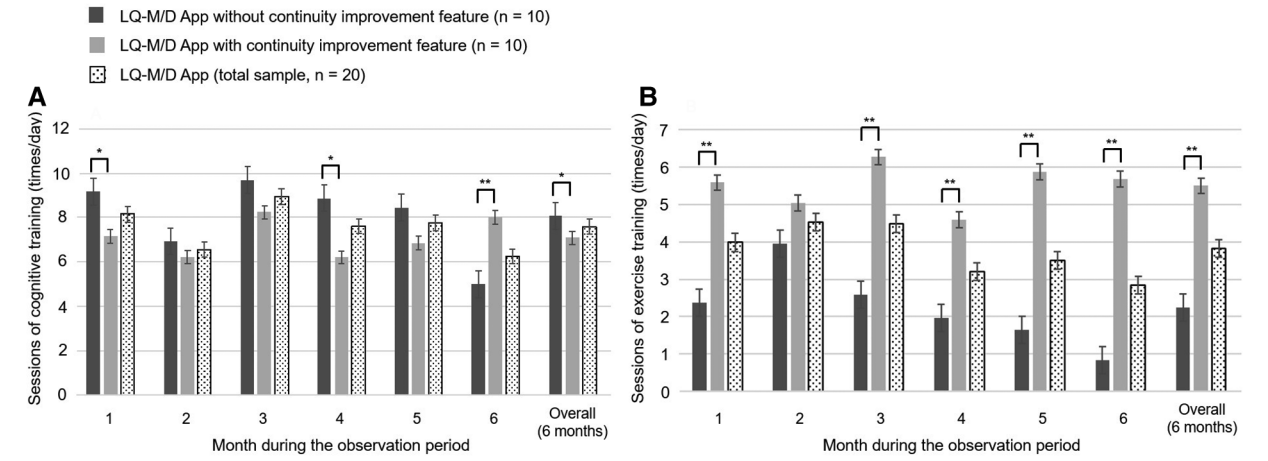


FIGURE 4
Number of cognitive and exercise training sessions on LQ-M/D App. Number of sessions of cognitive and exercise training performed by patients who used LQ-M/D App ($n = 20$), LQ-M/D App without the continuity improvement feature ($n = 10$), and LQ-M/D App with the continuity improvement feature ($n = 10$), from the first to the sixth month, as well as the entire observation period (6 months). (A) Graph showing the mean values for sessions of cognitive training performed by each group, with error bars representing the standard error. (B) Graph showing the mean values for sessions of exercise training performed by each group, with error bars representing the standard error. *$p < 0.05$, **$p < 0.005$.

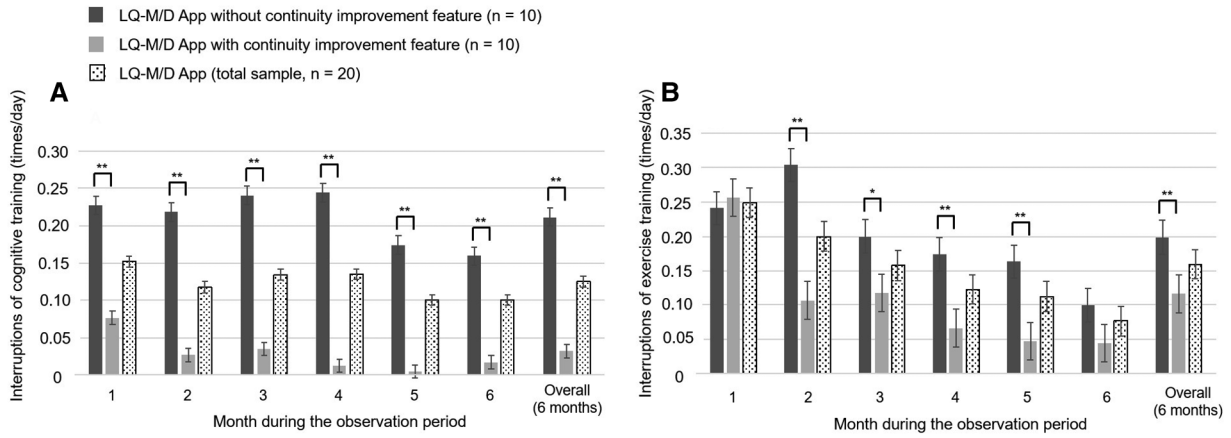


FIGURE 5
Interruptions in cognitive and exercise training on LQ-M/D App. Interruptions of cognitive and exercise training of patients who used LQ-M/D App ($n = 20$), LQ-M/D App without the continuity improvement feature ($n = 10$), and LQ-M/D App with the continuity improvement feature ($n = 10$), from the first to the sixth month, as well as for the entire observation period (6 months). (A) Graph showing the mean values for interruptions of cognitive training in each group, with error bars representing the standard error. (B) Graph showing the mean values for interruptions of exercise training in each group, with error bars representing the standard error. *$p < 0.05$, **$p < 0.005$.

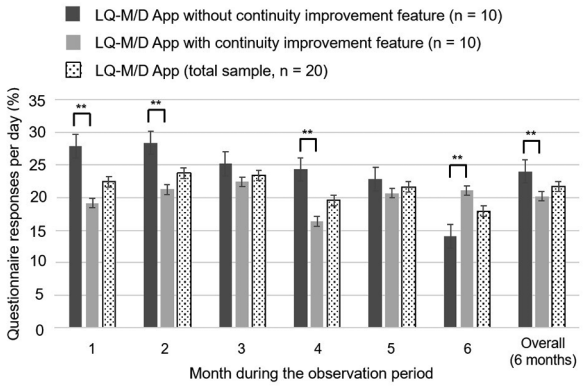


FIGURE 6
Questionnaire response rates on LQ-M/D App. Questionnaire responses for LQ-M/D App ($n = 20$), LQ-M/D App without the continuity improvement feature ($n = 10$), and LQ-M/D App with the continuity improvement feature ($n = 10$), from the first to the sixth month, as well as for the entire observation period (6 months). The graph shows the mean values for questionnaire response rates in each group, with error bars representing the standard error. *$p < 0.05$, **$p < 0.005$.

app usage duration, cognitive training sessions, and exercise training sessions, suggesting potential cognitive improvement. In contrast, a negative correlation was observed with ADAS-jcog scores, also implying cognitive enhancement. However, the *p*-values were not statistically significant, suggesting that further research is needed to confirm these findings and their effects on cognitive function.

Discussion

In this study, we investigated the feasibility and utility of LQ-M/D App, a smartphone application developed for patients with MCI or mild dementia. The results demonstrated that the app could be utilized by the study population without major issues, and that the continuity improvement feature introduced in the updated version of the app appeared to positively influence app usage in several aspects. Although some app malfunctions were observed during the study, they were within acceptable limits and did not greatly hinder the app's usability. However, although the association between app usage and cognitive ability of the patients showed a tendency for a positive correlation, the available data was not sufficient to draw definitive conclusions. Therefore, further research is needed to confirm these findings and their effects on cognitive function.

Previous studies have reported the potential benefits of nonpharmacological interventions, such as moderate physical activity and lifestyle modifications, in slowing cognitive decline and reducing the progression from MCI to dementia (23, 24). LQ-M/D App incorporates cognitive and physical exercise training, as well as lifestyle habit acquisition features, which may contribute to the management of MCI or mild dementia. The app also includes disease education and family monitoring through the family app, which may help patients and their families better understand and manage their condition, in line with recommendations for MCI management (25, 26). The comparison between the pre-update and post-update versions of LQ-M/D App in our study demonstrated that the continuity improvement feature positively influenced app usage duration, the number of cognitive training and exercise training sessions, and the number of interruptions. These findings suggest that the updated app may provide a more engaging and user-friendly experience for patients with MCI or mild dementia, which is consistent with previous reports on the usability of mobile health interventions for older adults (27–29).

Our study observed a generally low questionnaire response rate in both the pre-update and post-update versions of LQ-M/D App. A possible reason for this might be the lack of evaluation and

10.3389/fdgth.2023.1243253

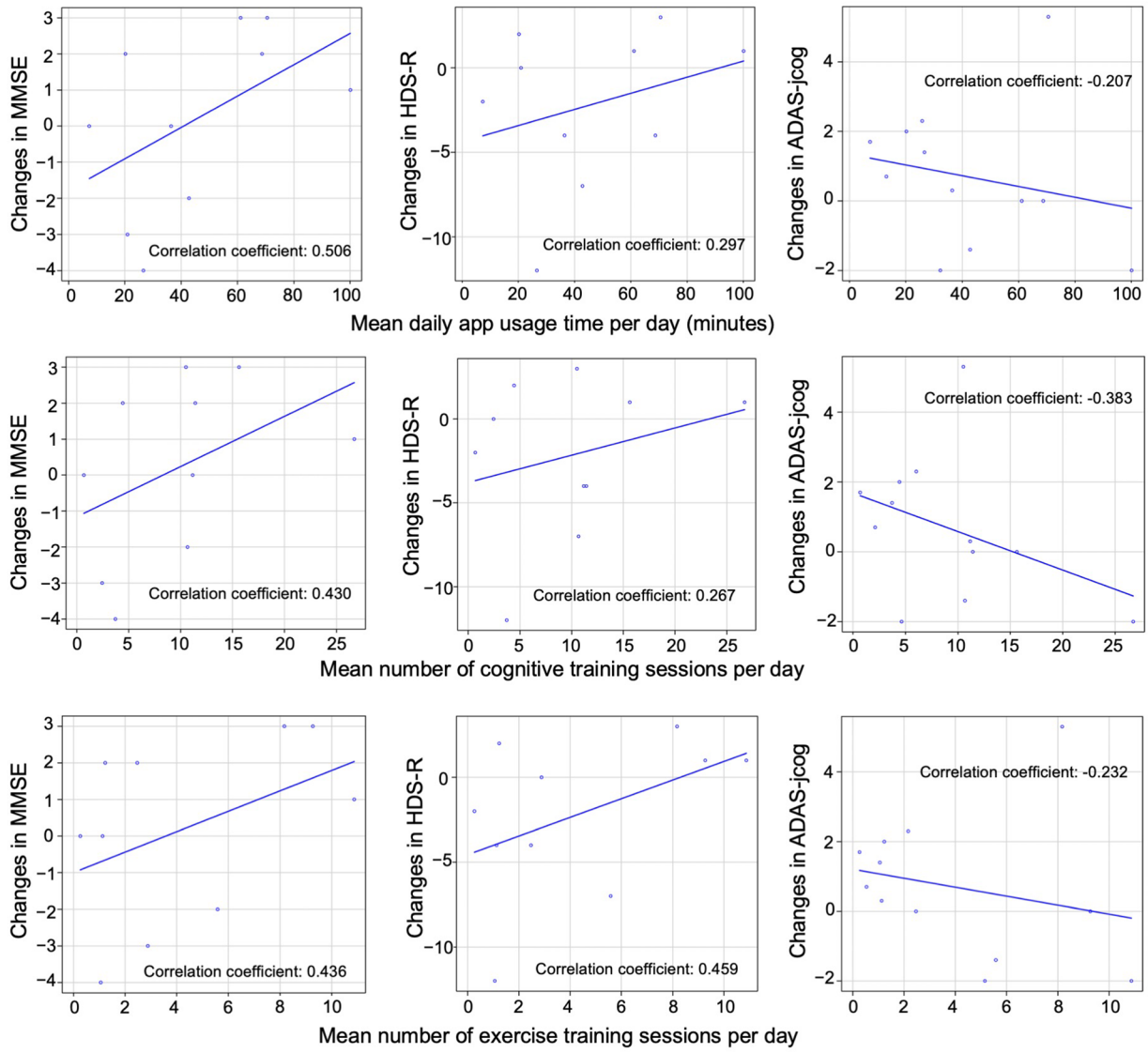


FIGURE 7
Correlation between app usage and cognitive test scores. Scatterplots illustrating the association between daily app usage duration, daily number of cognitive training sessions, and exercise training sessions with changes in MMSE, HDS-R, and ADAS-jcog scores from the start of app use to 6 months after the end of app use in 15 patients. The Pearson's correlation coefficient for each association is provided. No statistically significant differences were observed between the groups.

feedback on patients' questionnaire responses, as feedback has been shown to improve engagement and adherence in eHealth interventions (30). The post-update version showed a lower response rate compared with the pre-update version, potentially owing to response fatigue among patients resulting from the increased frequency of questionnaires in the post-update version (31, 32). Further investigation is needed to determine the optimal questionnaire frequency and the role of feedback in balancing data collection and patient engagement in electronic health interventions.

In this study, the association between app usage and cognitive improvement did not reach statistical significance, with correlation coefficients demonstrating wide-ranging confidence intervals and relatively high *p*-values. These findings suggest a potential association between app usage and cognitive improvement but should be interpreted with caution owing to the small sample size and the retrospective nature of the study. Moreover, among the 20 patients, 17 were receiving medication for dementia, and 2 patients started taking medication during the study period. Therefore, the possibility of the slight improvement in cognitive performance being a result of the effects of the medication rather than the effects of training using the App cannot be ruled out. Further research with larger sample sizes and prospective study designs is necessary to better understand the effects of the LQ-M/D App on cognitive function in patients with MCI or mild dementia, and to confirm the potential cognitive benefits of the app and to investigate the underlying mechanisms contributing to these correlations, as suggested by prior research on the effectiveness of cognitive training for cognitive function (33).

There are some limitations to our study. First, the sample size was relatively small, which may have affected the statistical power of the analyses. Second, the study was conducted in a single institution, which may limit the generalizability of the findings. Third, the study was retrospective in nature, which may have introduced biases in the data collection and interpretation. Fourth, the study lacked a control group, making it difficult to determine the specific effects of the LQ-M/D App on cognitive function compared with standard treatments alone. Future research should consider a randomized controlled trial design with larger sample sizes and multiple institutions to further evaluate the effectiveness of LQ-M/D App in managing MCI and mild dementia.

Conclusions

In conclusion, this study demonstrated the feasibility and potential utility of LQ-M/D App for patients with MCI or mild dementia. The app's continuity improvement feature appeared to positively influence app usage, and the observed trends in cognitive assessments suggested potential cognitive improvement, although further research is needed to confirm these findings. LQ-M/D App may be a promising tool for the management of MCI and mild dementia, potentially contributing to the prevention or delay of cognitive decline in this population.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Institutional Review Board of Juntendo University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

All authors conceived and designed the study. RH performed the literature review and drafted the manuscript. TH and HK independently managed the study data. SI analyzed the data and provided insights for the interpretation of the results. HI supervised the overall study execution. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by Life Quest Inc., Tokyo, Japan, which provided the app used in the research free of charge and covered the necessary expenses for conducting the research. The funder was not involved in the study design, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

Acknowledgments

The authors thank Helena Akiko Popiel of Tokyo Medical University for her editing of this article.

Conflict of interest

RH, YH, HT, NH, and SH hold executive or advisory positions within Life Quest Inc. EI is involved in development and research at Life Quest Inc., although not directly engaged in the development of the app analyzed in this study. RH, YH, HT, and SH own unlisted shares of the company. RS is the founder and a shareholder of Life Quest Inc.

The remaining authors declare that the research was conducted in the absence of any commercial relationships that could be construed as potential conflict of interest.

Publisher's note

References

1. Ministry of Health, Labour and Welfare Scientific Research Grant. Special research project for fiscal year 2014: a study on future projections of elderly population with dementia in Japan (2015). (in Japanese)

2. Petersen RC. Early diagnosis of Alzheimer's disease: is MCI too late? *Curr Alzheimer Res.* (2009) 6(4):324–30. doi: 10.2174/156720509788929237

3. Luck T, Luppa M, Briel S, Riedel-Heller SG. Incidence of mild cognitive impairment: a systematic review. *Dement Geriatr Cogn Disord.* (2010) 29(2):164–75. doi: 10.1159/000272424

4. Ward A, Arrighi HM, Michels S, Cedarbaum JM. Mild cognitive impairment: disparity of incidence and prevalence estimates. *Alzheimers Demen.* (2012) 8 (1):14–21. doi: 10.1016/j.jalz.2011.01.002

5. Mitchell AJ, Shiri-Feshki M. Rate of progression of mild cognitive impairment to dementia–meta-analysis of 41 robust inception cohort studies. *Acta Psychiatr Scand*. (2009) 119(4):252–65. doi: 10.1111/j.1600-0447.2008.01326.x

6. Roberts R, Knopman DS. Classification and epidemiology of MCI. *Clin Geriatr Med*. (2013) 29(4):753–72. doi: 10.1016/j.cger.2013.07.003

7. Doody RS, Ferris SH, Salloway S, Sun Y, Goldman R, Watkins WE, et al. Donepezil treatment of patients with MCI: a 48-week randomized, placebo-controlled trial. *Neurology*. (2009) 72(18):1555–61. doi: 10.1212/01.wnl.0000344650.95823.03

8. Baker LD, Frank LL, Foster-Schubert K, Green PS, Wilkinson CW, McTiernan A, et al. Effects of aerobic exercise on mild cognitive impairment: a controlled trial. *Arch Neurol*. (2010) 67(1):71–9. doi: 10.1001/archneurol.2009.307

9. Etgen T, Sander D, Bickel H, Forstl H. Mild cognitive impairment and dementia: the importance of modifiable risk factors. *Dtsch Arztebl Int*. (2011) 108(44):743–50. doi: 10.3238/arztebl.2011.0743

10. World Health Organization. Risk reduction of cognitive decline and dementia: WHO guidelines. Geneva: World Health Organization (2019). Available at: https://apps.who.int/iris/bitstream/handle/10665/312180/9789241550543-eng.pdf?ua=1 (Accessed June 8, 2023).

11. Li J, Wang YJ, Zhang M, Xu ZQ, Gao CY, Fang CQ, et al. Vascular risk factors promote conversion from mild cognitive impairment to Alzheimer disease. *Neurology*. (2011) 76(17):1485–91. doi: 10.1212/WNL.0b013e318217e7a4

12. Chaudhry BM, Dasgupta D, Chawla NV. Successful aging for community-dwelling older adults: an experimental study with a tablet app. *Int J Environ Res Public Health*. (2022) 19:13148. doi: 10.3390/ijerph192013148

13. Dasgupta D, Chaudhry B, Koh E, Chawla NV. A survey of tablet applications for promoting successful aging in older adults. *IEEE Access*. (2016) 4:9005–17. doi: 10.1109/ACCESS.2016.2632818

14. Bateman DR, Srinivas B, Emmett TW, Schleyer TK, Holden RJ, Hendrie HC, et al. Categorizing health outcomes and efficacy of mHealth apps for persons with cognitive impairment: a systematic review. *J Med Internet Res*. (2017) 19(8):e301. doi: 10.2196/jmir.7814

15. Chaudhry BM, Smith J. Refinemind: a mobile app for people with dementia and their caregivers. In: Chandra Kruse L, Seidel S, Hausvik GI, editors. *The next wave of sociotechnical design. DESRIST 2021. Lecture notes in computer science*. Vol. 12807. Cham: Springer (2021). p. 16–21.

16. Onoda K, Hamano T, Nabika Y, Aoyama A, Takayoshi H, Nakagawa T, et al. Validation of a new mass screening tool for cognitive impairment: cognitive assessment for dementia, iPad version. *Clin Interv Aging*. (2013) 8:353–60. doi: 10.2147/CIA.S42342

17. Edgar CJ, Siemers E, Maruff P, Petersen RC, Aisen PS, Weiner MW, et al. Pilot evaluation of the unsupervised, at-home cogstate brief battery in ADNI-2. *J Alzheimers Dis*. (2021) 83(2):915–25. doi: 10.3233/JAD-210201

18. Suzuki T, Shimada H, Makizako H, Doi T, Yoshida D, Ito K, et al. A randomized controlled trial of multicomponent exercise in older adults with mild cognitive impairment. *PLoS One*. (2013) 8(4):e61483. doi: 10.1371/journal.pone.0061483

19. Shellington EM, Felfeli T, Shigematsu R, Gill DP, Petrella RJ. Healthebrain: an innovative smartphone application to improve cognitive function in older adults. *Mhealth*. (2017) 3:17. doi: 10.21037/mhealth.2017.04.05

20. Gates NJ, Vernooij RW, Di Nisio M, Karim S, March E, Martínez G, et al. Computerised cognitive training for preventing dementia in people with mild cognitive impairment. *Cochrane Database Syst Rev*. (2019) 3(3):Cd012279. doi: 10.1002/14651858.CD012279.pub2

21. Baric V, Andreassen M, Ohman A, Hemmingsson H. Using an interactive digital calendar with mobile phone reminders by senior people—a focus group study. *BMC Geriatr*. (2019) 19(1):116. doi: 10.1186/s12877-019-1128-9

22. Kanda Y. Investigation of the freely available easy-to-use software "EZR" for medical statistics. *Bone Marrow Transplant*. (2013) 48(3):452–8. doi: 10.1038/bmt.2012.244

23. Gauthier S, Reisberg B, Zaudig M, Petersen RC, Ritchie K, Broich K, et al. Mild cognitive impairment. *Lancet*. (2006) 367(9518):1262–70. doi: 10.1016/S0140-6736(06)68542-5

24. Petersen RC. Clinical practice. Mild cognitive impairment. *N Engl J Med*. (2011) 364(23):2227–34. doi: 10.1056/NEJMcp0910237

25. Kueper JK, Speechley M, Montero-Odasso M. The Alzheimer's disease assessment scale-cognitive subscale (ADAS-cog): modifications and responsiveness in pre-dementia populations: a narrative review. *J Alzheimers Dis*. (2018) 63(2):423–44. doi: 10.3233/JAD-170991

26. Winblad B, Amouyel P, Andrieu S, Ballard C, Brayne C, Brodaty H, et al. Defeating Alzheimer's disease and other dementias: a priority for European science and society. *Lancet Neurol*. (2016) 15(5):455–532. doi: 10.1016/S1474-4422(16)00062-4

27. Reeder B, Meyer E, Lazar A, Chaudhuri S, Thompson HJ, Demiris G. Framing the evidence for health smart homes and home-based consumer health technologies as a public health intervention for independent aging: a systematic review. *Int J Med Inform*. (2013) 82(7):565–79. doi: 10.1016/j.ijmedinf.2013.03.007

28. Wildenbos GA, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: a literature based framework (MOLD-US). *Int J Med Inform*. (2018) 114:66–75. doi: 10.1016/j.ijmedinf.2018.03.012

29. Gordon NP, Hornbrook MC. Older adults' readiness to engage with eHealth patient education and self-care resources: a cross-sectional survey. *BMC Health Serv Res*. (2018) 18(1):220. doi: 10.1186/s12913-018-2986-0

30. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. *Am J Prev Med*. (2016) 51(5):833–42. doi: 10.1016/j.amepre.2016.06.015

31. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. *J Med Internet Res*. (2009) 11(2):e13. doi: 10.2196/jmir.1194

32. Eysenbach G. The law of attrition. *J Med Internet Res*. (2005) 7(1):e11. doi: 10.2196/jmir.7.1.e11

33. Hill NT, Mowszowski L, Naismith SL, Chadwick VL, Valenzuela M, Lampit A. Computerized cognitive training in older adults with mild cognitive impairment or dementia: a systematic review and meta-analysis. *Am J Psychiatry*. (2017) 174(4):329–40. doi: 10.1176/appi.ajp.2016.16030360

frontiers | Frontiers in Digital Health

TYPE Brief Research Report
PUBLISHED 12 September 2023
DOI 10.3389/fdgth.2023.1215187

OPEN ACCESS

EDITED BY
Edith Talina Luhanga,
Carnegie Mellon University Africa, Rwanda

REVIEWED BY
Karen Sowon,
Carnegie Mellon University, United States,
Carine Pierrette Mukamakuza,
Carnegie Mellon University Africa, Rwanda

*CORRESPONDENCE
Nele Albers
✉ n.albers@tudelft.nl

RECEIVED 01 May 2023
ACCEPTED 25 August 2023
PUBLISHED 12 September 2023

CITATION
Ghantasala RP, Albers N, Penfornis KM, van Vliet MHM and Brinkman W-P (2023) Feasibility of generating structured motivational messages for tailored physical activity coaching. Front. Digit. Health 5:1215187. doi: 10.3389/fdgth.2023.1215187

Feasibility of generating structured motivational messages for tailored physical activity coaching

Ramya P. Ghantasala[1], Nele Albers[1]*, Kristell M. Penfornis[2,3], Milon H. M. van Vliet[4,5] and Willem-Paul Brinkman[1]

[1]Department of Intelligent Systems, Delft University of Technology, Delft, Netherlands, [2]Unit Health Medical and Neuropsychology, Institute of Psychology, Leiden University, Leiden, Netherlands, [3]Department of Cardiology, Leiden University Medical Center, Leiden, Netherlands, [4]Public Health and Primary Care, Leiden University Medical Center, Leiden, Netherlands, [5]National eHealth Living Lab, Leiden University Medical Center, Leiden, Netherlands

Tailored motivational messages are helpful to motivate people in eHealth applications for increasing physical activity, but it is not sufficiently clear how such messages can be effectively generated in advance. We, therefore, put forward a theory-driven approach to generating tailored motivational messages for eHealth applications for behavior change, and we examine its feasibility by assessing how motivating the resulting messages are perceived. For this, we designed motivational messages with a specific structure that was based on an adaptation of an existing ontology for tailoring motivational messages in the context of physical activity. To obtain tailored messages, experts in health psychology and coaching successfully wrote messages with this structure for personas in scenarios that differed with regard to the persona's mood, self-efficacy, and progress. Based on an experiment in which 60 participants each rated the perceived motivational impact of six generic and six tailored messages based on scenarios, we found credible support for our hypothesis that messages tailored to mood, self-efficacy, and progress are perceived as more motivating. A thematic analysis of people's free-text responses about what they found motivating and demotivating about motivational messages further supports the use of tailored messages, as well as messages that are encouraging and empathetic, give feedback about people's progress, and mention the benefits of physical activity. To aid future work on motivational messages, we make our motivational messages and corresponding scenarios publicly available.

KEYWORDS

physical activity, behavior change, eHealth, tailoring, personalization, motivation, feedback, monitoring

1. Introduction

Tailored motivational messages have been shown to be useful in eHealth applications for behavior change such as ones for becoming more physically active (1–4). For example, providing users with actionable feedback (i.e., feedback that can be acted on) in combination with information about their progress has been found to increase intrinsic motivation as well as adherence (5). For somebody creating an eHealth application, this means that they need to find a way to generate motivational messages adjusted to each combination of values for the tailoring factors that might be relevant to consider, such as a person's progress or motivation. Since it is likely not feasible for experts to write and send these messages in real time after the application has been deployed, the messages should be generated in advance.

A common approach to generating motivational messages used in current eHealth applications for behavior change is writing custom and hand-crafted messages without a specific structure (e.g., (6–10)). While this approach may lead to messages that are effective for the specific application at hand, it is difficult for other researchers to follow the same recipe for generating messages for a different behavior change application. This hinders the reuse, reproducibility, and generalizability of findings, especially when the messages are also not publicly shared (e.g., (6–8, 11, 12)).

To obtain structured motivational messages, both data-driven and theory-driven approaches have been employed. For example, Tielman et al. (13) asked experts to write messages for scenarios differing based on user trust and therapy progress in the context of post-traumatic stress disorder therapy. The sentences in the resulting messages were used to identify categories such as "compliment" and "give perspective," and new messages were generated based on the probabilities of categories appearing for a certain combination of tailoring factors. The resulting messages thus do have a specific structure, but this structure is based on data from a specific domain and may hence not be applicable to other domains. An alternative to this data-driven approach is a theory-driven one. Thomas et al. (14), for example, proposed asking peers to write messages for healthy eating based on argumentation schemes (i.e., forms of argument). Because the resulting message structure is not tailored to a specific domain, it can be used in different behavior change domains, thus reducing the time needed for designing and validating messages for the new domain. However, limited work on such theory-driven approaches exists for generating motivational messages that are tailored.

Our goal was thus to develop and test the feasibility of a theory-driven approach to generating tailored motivational messages. We, therefore, asked experts in health psychology and coaching to write motivational messages with a specific structure derived from an ontology. To obtain tailored messages, these messages were written for personas in scenarios that differed with regard to the mood, self-efficacy, and progress of the persona in the context of becoming more physically active. A variety of tailoring factors for motivational as well as persuasive messages more generally has been examined in previous work, including both dynamic factors (e.g., progress (15), self-efficacy (16), states derived from the Capability-Opportunity-Motivation-Behavior (COM-B) model (17)) and more stable user characteristics (e.g., gender (18), need for cognition (19), physical activity identity (20)). In line with our goal, we chose three factors that have been shown to be relevant in previous work and not necessarily the most important factors. Progress was chosen to make progress feedback possible, which can help users keep a positive mindset about physical activity (15), increase users' motivation to reach their goal behavior (21), and make messages more interesting (22) and thus more likely to be processed in detail (23) and with a persistent impact (24); mood and self-efficacy were primarily chosen because they can affect how messages and feedback, in particular, are processed. Specifically, mood has been shown to affect how a user processes messages (25) and feedback regarding behavior change (26, 27), and self-efficacy influences how a user absorbs message content (28) and is persuaded by differently framed messages for behavior change (16). Mood and self-efficacy can thus affect how effective different motivational messages are. Furthermore, mood and self-efficacy both influence physical activity and are influenced by it (26, 29, 30). This can be taken into account in motivational messages for physical activity if a user's mood and self-efficacy are known.

To validate our resulting messages, we conducted an experiment in which 60 participants each rated the perceived motivational impact of six tailored and six generic messages based on scenarios. Our hypothesis was that the tailored messages are perceived as more motivating than the generic ones. To improve motivational messages in the future, we further qualitatively analyzed people's free-text responses about what they find motivating and demotivating about motivational messages. Our results support the feasibility of theory-based generation and the use of messages tailored to mood, self-efficacy, and progress. To aid future work on motivational messages, we make the scenarios and resulting 60 tailored and 12 generic motivational messages publicly available (31).

2. Method

Our experiment was run from December 2021 to January 2022. The Human Research Ethics Committee of Delft University of Technology granted ethical approval for the research (Letter of Approval number: 1814) and the experiment was pre-registered on the Open Science Framework (OSF) (32).

2.1. Experimental design

The study was set up as a double-blind within-subjects experiment. The within-subjects factor was the type of message that participants rated for a hypothetical scenario. In the control condition, participants rated generic messages; in the experimental condition, they rated messages tailored to progress, mood, and self-efficacy. To mitigate order and learning effects, ABBA counterbalancing (33) was performed when assigning tailored or generic messages to participants. Furthermore, the progress, mood, and self-efficacy of the persona used in the hypothetical scenarios were counterbalanced with a modified Latin square.

2.2. Materials

The online crowdsourcing platform Prolific and the survey platform Qualtrics were used to recruit participants and host the questionnaires.

Scenarios. 30 hypothetical scenarios were designed to elicit motivational messages from experts. The scenarios were created by combining three levels of mood, two levels of self-efficacy, and five levels of progress ($2 \times 3 \times 5 = 30$). Each scenario had a persona with a specific mood, self-efficacy, and progress level of their physical activity program. The list of scenarios can be found in our online repository. Two independent coders labeled each scenario with a mood, self-efficacy, and progress level to

evaluate whether each scenario had a distinguishable mood, self-efficacy, and progress level. We obtained substantial to almost perfect agreement (34) for mood ($\kappa = 0.73$), self-efficacy ($\kappa = 0.73$) and progress ($\kappa = 0.96$).

Motivational messages. Using the framework by Op den Akker et al. (27) as inspiration, we created an ontology to model tailored motivational messages for physical activity coaching based on their intention (i.e., why the message is sent) and content (i.e., what is in the message) (**Figure 1**). The message structure (**Figure 2**) consists of the components feedback, argument, reinforcement, and suggestion following the model by Op den Akker et al. (27), which was created as a model of tailoring for real-time physical activity coaching applications based on a survey of the literature and of messages used in previous studies. Two health psychologists with experience in behavioral change coaching and developing eHealth applications for increasing physical activity were recruited from the network of the authors to write tailored motivational messages for the 30 hypothetical scenarios based on this ontology for a total of 60 messages. The psychologists were asked to consider the people in the scenarios as their clients or patients and write the messages as if they were motivating them to improve their physical activity levels. Along with the tailored messages, the psychologists were asked to write six generic messages each to motivate people to increase their physical activity regardless of their mood, self-efficacy, or progress. The psychologists were instructed to write all messages using the message structure depicted in **Figure 2** if possible. All messages are provided in our online repository.

2.3. Measures

We tested our main hypothesis using the following measure:

Perceived motivational impact of messages. For each message, participants rated the perceived motivational impact on an 11-point scale from −5 (“Very demotivating”) to 5 (“Very motivating”), adapted from the one by de Vries et al. (35). A higher value thus indicates a higher perceived motivational impact.

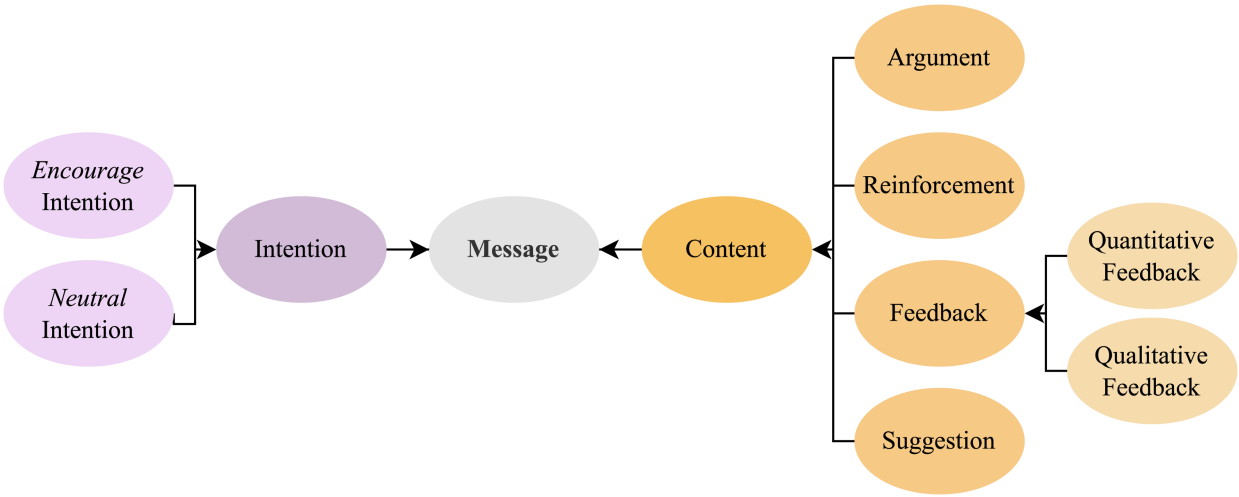


FIGURE 1
The ontology used for generating tailored motivational messages to improve physical activity. Based on the ontology developed by Op den Akker et al. (27), we have added the components of qualitative and quantitative feedback, while retaining the *encourage* and *neutral* intentions as well as the suggestion, argument, and reinforcement components.

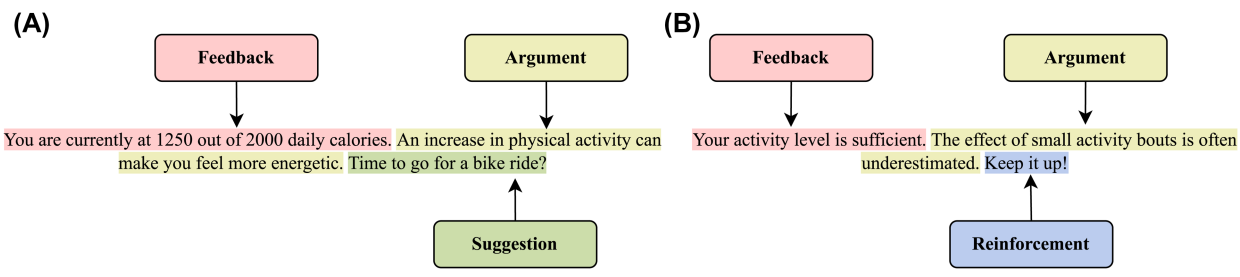


FIGURE 2
Message structure for messages with (A) *encourage* and (B) *neutral* intentions based on Op den Akker et al. (27). An *encourage* intention encourages users to do (more) physical activity, whereas a *neutral* intention acknowledges a user's progress and informs them to keep up the good work. If a user is consistently achieving their goals, they need not be reminded frequently about the arguments to stay active, and hence only the feedback and reinforcement components can then be used for messages with *neutral* intentions.

In addition, we collected the following measures for exploratory research:

Motivating and demotivating factors of messages. Participants were asked to provide free-text answers to the questions "How would a message motivate you to do physical activity?" and "How would a message demotivate you from doing physical activity?" The questions were adapted from the ones by Fukuoka et al. (36).

2.4. Participants

We conducted a power analysis for repeated measures, within-factors ANOVA with 6 repeated measurements per condition (i.e., 12 measurements in total), a small effect size ($f = 0.18$) (37), an alpha of 0.05, and a power of 0.90 to obtain a target sample size for our Bayesian analysis. The result was a sample size of 30. This estimate was conservative, as Bayesian analyses require no more samples than the corresponding frequentist ones (38). We also decided to add an additional 30 participants for a total of 60 if we had sufficient funds left after the first 30 participants had completed the study. This was done to increase the power, as well as to get more responses for the qualitative analysis. The decision to include more participants was taken without looking at experimental data.

To be eligible, participants had to meet the following criteria:

1. have an approval rate of ≥95% on the crowdsourcing platform Prolific to obtain participants with reliable submissions (39),
2. speak English fluently, and
3. have at least 10 previous submissions on Prolific, to avoid low-effort or low-quality submissions (as recommended by Prolific to recruit active and committed users (40)).

60 participants completed the study out of a total of 121 who started it. If participants passed at least half of the four and two attention check questions added in the two sections of the questionnaire respectively (i.e., the pre-questionnaire and scenario-rating sections), gave informed consent, and passed the pre-screener validation about English fluency, they were approved and paid according to the minimum payment rules on Prolific (i.e., five pounds sterling per hour). Participants were informed that their responses to the questionnaire would not affect their payment in any way unless it violated the conditions mentioned above. This served to mitigate biases such as the ones mentioned by Draws et al. (41), specifically loss aversion and self-interest bias. Participant characteristics such as age and gender are provided in the **Supplementary Material**. Participants on Prolific are nationals of or live in member countries of the Organisation for Economic Co-operation and Development (OECD) with the exception of Turkey, Lithuania, Colombia, and Costa Rica and the addition of South Africa (42).

2.5. Procedure

The study consisted of a questionnaire divided into two sections:

- *a pre-questionnaire section* to collect data on user characteristics, and
- *a scenario-rating section* in which participants were given 6 hypothetical scenarios with motivational messages intended to motivate the persona in the scenario per condition, for a total of 12 scenarios. To mitigate order effects and learning effects, ABBA counterbalancing was used when assigning tailored or generic messages. For each scenario, participants were asked to rate how motivating they would find the messages for themselves if they were the persona in the scenario. After the scenarios, participants were given two free-text questions about what they find motivating and demotivating in a motivational message.

2.6. Data preparation and analysis strategies

The data collected from the experiment was preprocessed by (1) removing data from participants whose submissions were not approved based on the earlier described criteria, and (2) anonymizing the data. The data and analysis script are available online (31) allowing to reproduce the analyses in a Docker container as recommended for Bayesian analyses by van de Schoot et al. (43).

We conducted a multi-level (i.e., hierarchical) Bayesian analysis using version 2.13 of the rethinking package (44). We fitted a model to a general mean, a random intercept for each participant, and a fixed effect for tailoring using diffuse priors based on the ones by McElreath (44).[1] We fit a t-distribution to our output variable, which is the perceived motivational impact of the messages. Using a prior sensitivity analysis to assess the impact of different settings for the priors, we found that the posterior probability of our main hypothesis holding remained unchanged for the tested priors. Based on the fitted model, we computed the posterior probability that our hypothesis was true. More precisely, we calculated the posterior probability that the fixed effect for tailoring was greater than zero. This posterior probability was evaluated based on the guidelines by Chechile (38) and their extension to posterior probabilities of less than 0.5 by Andraszewicz et al. (45). We also report the 95% Highest Posterior Density Interval (HPDI) for estimators, with an HPDI being "the narrowest interval containing the specified probability mass" (44).

In addition, we performed a thematic analysis (46) of participants' free-text responses about what they find motivating and demotivating about motivational messages for physical activity. After familiarizing herself with the data, RG created an initial coding scheme and coded all responses according to this scheme. This means that codes were largely created inductively. To assess the reliability of the coding, KP was trained on 20 responses for motivating factors and 10 responses for demotivating factors before independently coding all remaining responses based on the coding scheme. Afterward, the coding scheme was refined by RG and KP, and coding disagreements

[1]We used normal priors with $\mu = 0$ and $\sigma = 10$ for both the general mean and the fixed effect for tailoring. The full model specification can be found in the online repository (31).

were resolved by means of discussion. The updated coding scheme was thus developed by multiple researchers, which has been described as a way to increase the validity of qualitative research (47). To evaluate the reliability of this updated coding scheme, we conducted a second round of double coding with NA. After being trained on 20 responses each for motivating and demotivating factors, NA coded all remaining responses. We obtained fair to moderate agreement (34) based on a Cohen's κ of 0.40 for motivating and 0.46 for demotivating factors. In our analysis, we consider only those codes with at least moderate agreement (i.e., a Cohen's κ of at least 0.41). For these codes that we consider in our analysis, the agreement is substantial to almost perfect based on a Cohen's κ of 0.69 for motivating and 0.81 for demotivating factors. The final coding scheme together with the corresponding Cohen's κ values can be found in the **Supplementary Material**.

3. Results

3.1. Perceived motivational impact of tailored vs. generic messages

Figure 3A compares the sample motivational impact of the two message types. It shows that the sample mean perceived motivational impact is higher for tailored ($M = 2.33$, SD $= 2.11$) than for generic messages ($M = 1.32$, SD $= 2.29$). Quantifying this by means of our Bayesian analysis shows that the perceived motivational impact of tailored messages is 1.02 (SD $= 0.13$) scale points higher than the one of generic messages. The corresponding 95%-HDPI ranges from 0.76 to 1.28, with >99.999% of the credibility mass favoring the higher motivational impact of tailored messages (**Figure 3B**). According to the guidelines by Chechile (38), this can be qualified, or can be "bet on," as our hypothesis being virtually certainly true.

3.2. Exploratory analysis: motivating and demotivating factors of messages

We identified four themes describing what participants found motivating and demotivating in a message (**Figure 4**).

Tailored or generic. The most frequent theme was whether a message is tailored to a particular user's situation, or contains cliched, empty platitudes. Tailored messages were perceived as motivating, with 17% of the responses about motivating factors containing references to having personalized information in a message. On the other hand, messages with empty, cliched platitudes, or which conveyed a false sense of positivity, were reported to be demotivating. 40% of the responses about demotivating factors mentioned these attributes of generic messages. This is also demonstrated by the answer of P56 when asked about demotivating factors of a message: "If the message is generalized, I would rather not hear it at all and have an automated system, not a doctor, say it to me." And when asked about motivating factors, P56 said: "If the message is specific to my own situation, so it seems more personal and tailored to myself, I would hear it with greater care and would adopt it easier."

Intention. The intention of the message is the primary takeaway the message has. Encouraging and empathetic messages were the most commonly mentioned motivating factors here, appearing in 40% and 17% of the responses. On the other hand, messages that focus on failures or missed goals (22%) or lack empathy (10%) were seen as demotivating. P44 reported this when asked about what motivates them: "It must have some kind of encouragement for me, to make me understand that I can do it and I have what it takes." Notably, however, too much empathy was also seen as demotivating by 7% of participants.

Goal-related information. Goal-related information encompasses all information related to a user's physical activity goals. In motivating factors, the most commonly recurring sub-themes in goal-related information were receiving information on one's progress toward a goal (20%), learning about the benefits

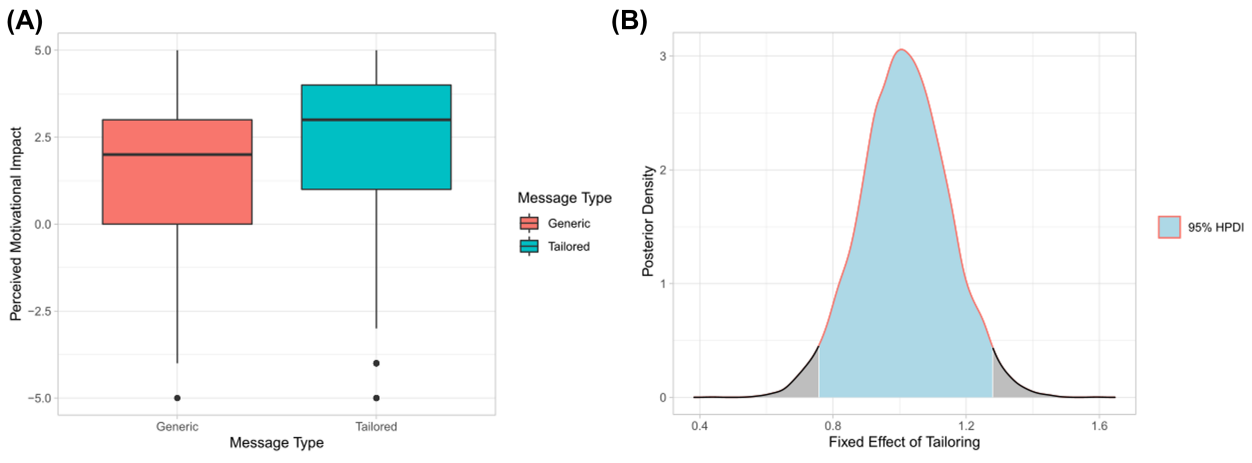


FIGURE 3
Perceived motivational impact of generic vs. tailored motivational messages. (A) Sample motivational impact ratings for the two types of messages. (B) Posterior density for the fixed effect of tailoring.

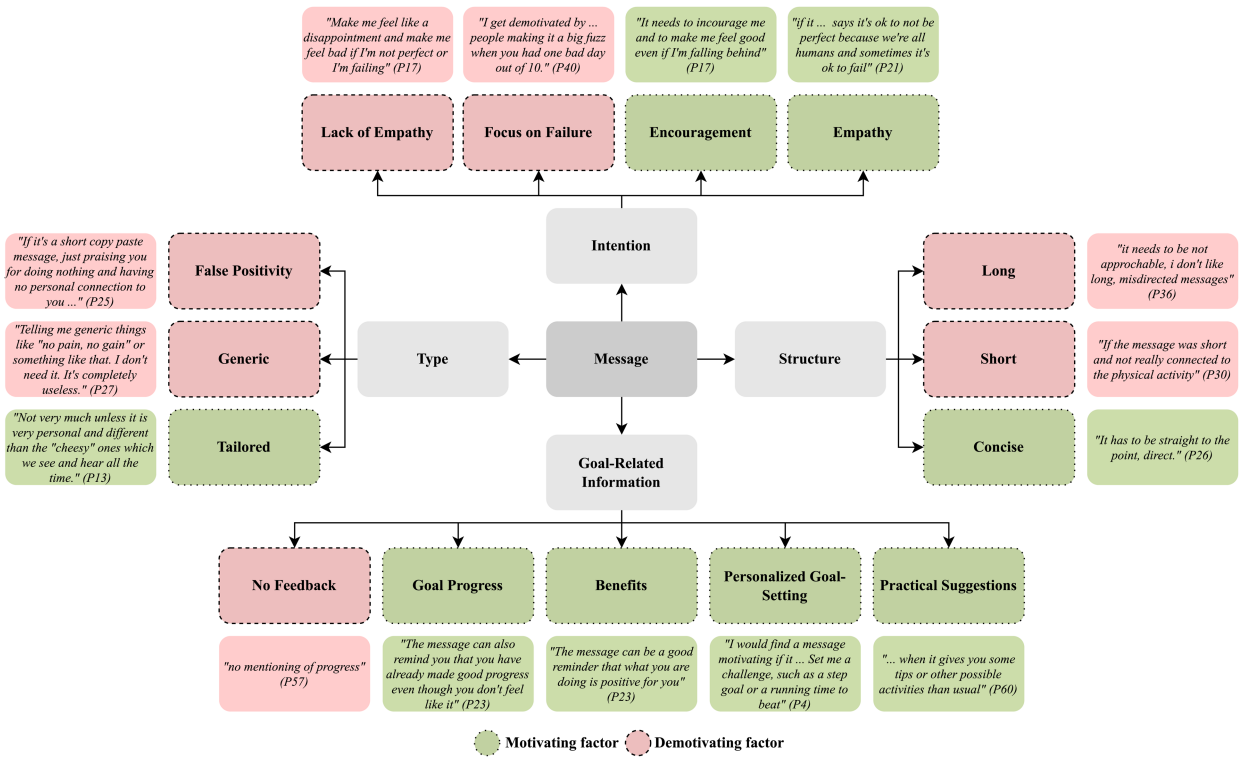


FIGURE 4
Themes describing factors participants found motivating and demotivating in a message, along with a few illustrative responses.

of physical activity (17%) as well as alternatives for goals and ways to be physically active (17%), and setting personalized goals (10%). Conversely, a demotivating factor was a lack of feedback regarding goals (8%). P18, for example, reported this when asked about demotivating factors: "A message would demotivate me by not validating any progress made, not being understanding of my needs and not providing alternative solutions to reach my goals."

Message structure. The structure of the message, specifically length and content style, was mentioned by a handful of participants. The responses about the message structure were conflicting, with short and long messages seen as demotivating in an equal percentage of responses (3%). For instance, P36 reported this when asked about demotivating factors: "…i don't like long, misdirected messages." P31, on the other hand, mentioned the following: "It is short and not personal. Completely random advice which does not help with my problem." Overall it thus seems that participants agreed on not wanting generic, misdirected messages, and associated either short or long messages with this. This matches the observation that concise messages were seen as motivating.

4. Discussion

Based on 60 participants each rating the perceived motivational impact of six generic and six tailored messages based on scenarios, we find that messages tailored to a user's mood, self-efficacy, and progress are perceived as more motivating than generic ones. This is in line with existing work which shows that tailored messages have modest success in motivating users in the context of health behavior change (1–4). Our results also complement existing research individually linking mood (48), self-efficacy (28), and progress (5) directly to motivation.

The thematic analysis of the motivating and demotivating factors of messages further revealed that motivation and de-motivation had common but complementary themes. For instance, goal-related information was a theme that was discovered in both motivating and demotivating factors, where participants found information about their progress and the benefits of physical activity motivating. A less-frequently mentioned factor was personalized goal-setting. On the other hand, participants found a lack of feedback demotivating. These results can be seen as validating the design of our messages, as the inclusion of progress and feedback was a key component of the design. Several of these themes have also been found in previous qualitative studies on motivation for physical activity. Kappen et al. (49), for instance, saw in the context of an eight-week physical activity intervention for older adults that people were motivated to do physical activity by "accomplishing a goal," which included being inspired by the in-app progress reports as well as by hitting pedometer targets. And participants of the physical and psychological intervention for breast cancer survivors by Sebri et al. (50) were primarily motivated to engage in physical activity by the benefits of physical activity such as improving their physical well-being. The fact that not all of the themes from previous studies on motivation for physical activity

(e.g., enjoying outdoors (49)) appeared in our study suggests, however, that what people find motivating in general and in motivational messages may not necessarily be the same.

Furthermore, participants highlighted that messages pertaining to their own situation (i.e., tailored messages) would be motivating, whereas cliched, generic messages would be demotivating. It stands to reason that motivating and demotivating factors are two sides of the same coin. The observation that tailored messages are perceived as motivating supports our quantitative findings as well as previous qualitative findings on motivation for physical activity. Albers et al. (51), for example, saw in the context of examples from other people shown in a goal-setting dialog for physical activity that examples from people that participants could relate to, and were thus perceived as tailored to the participant, were perceived as motivating.

4.1. Limitations

The main limitation of our study is that participants rated the perceived motivational impact of the messages for themselves if they were the personas in hypothetical scenarios instead of in real situations. To minimize the risk involved in the experiment such as the risk of injury, and due to the restrictions around Covid-19, we regarded the hypothetical scenarios as a good alternative to a real-world evaluation. Such scenarios have, for example, also successfully been used in studies by De Vries et al. (52) and Tielman et al. (13) to evaluate tailored motivational messages. Nevertheless, it is unclear how well our scenario-based findings generalize to other scenarios as well as users who are themselves in the situations described by the scenarios. To obtain a more accurate assessment of the messages' motivational impact, the messages would ideally be shown to users when they have made similar progress in their physical activity, or have a similar mood or self-efficacy. Conducting such a study also with other users besides crowd workers may further enhance the generalizability of the findings.

Another limitation is that only the perceived motivational impact of the messages was measured, and not the impact of the messages on users' physical activity. As stated in the COM-B model (30), motivation is only one of the three predictors of behavior, with opportunity and capability being the other two. Nevertheless, motivation has been shown to be a good predictor of behavior change (53–55), with the advantage of being easier to measure and being less noisy of a signal than actual behavior. Notably, even motivation can be talked about in terms of automatic vs. reflective according to the COM-B model (30). In our study, we considered solely reflective motivation, in which a user is actively and consciously involved in motivation, as opposed to automatic motivation, which is a result of impulsive, habitual, or drive-related behavior. However, the messages could also affect automatic motivation by, for example, influencing people's self-identity (30).

Furthermore, our evaluation of the tailored motivational messages was based on comparing them to expert-written generic messages. Since the tailored messages tended to be more detailed than the generic ones, it could be that the higher perceived motivational impact of the tailored messages is partially due to their higher level of detail or larger number of characters. To rule this out, one could compare tailored messages that match scenarios to tailored messages that do not match scenarios. An example of the latter is a message written for a confident persona being evaluated on a scenario with a persona with low confidence. However, given that congratulating a person with low confidence for their high confidence is unlikely to be very motivating, we regarded generic messages as a stronger baseline.

4.2. Directions for future work

This research could be extended by automatically detecting a user's state to tailor the messages and adapt content accordingly, as has been done based on affect by Grawemeyer et al. (56) in the context of student learning support. Similarly, Yang et al. (57) used sensors to automatically detect negative affect and send corresponding messages in the context of smoking cessation. To obtain more effective motivational messages, users' responses to the messages could as a next step be recorded to learn what kinds of messages users prefer. Current work on reinforcement learning for determining the best time to send a message (58) and adapting the framing of messages for inducing healthy nutritional habits (59) makes the idea of tailoring messages by automatically adapting to user state variables a feasible next step.

Second, our ontology and resulting message structure can be used by developers of eHealth applications to automatically generate messages, similar to the work by Tielman et al. (13), Ghosh et al. (60) and Thomas et al. (14). For example, given a set of motivational messages like ours, new messages can be generated by combining the components from different messages. The fact that messages written based on the ontology can be broken down into components also makes it easier for future researchers to understand and reproduce our work. With the ontology given, messages can be obtained from crowd-workers as well, as demonstrated by De Vries et al. (52), making message generation cheap as well as time- and resource-efficient. Directly comparing different approaches to generating tailored motivational messages in terms of time and resource efficiency as well as effectiveness would also be worthwhile.

Thirdly, we only considered messages represented in text form. However, other types of representation such as images (61), videos (62), or even emoticons (63) are possible. For example, auditory feedback was used in a study by Singh et al. (64) to motivate people with chronic pain to be more physically active, and audio, visual, and haptic representations of motivational messages were proposed by Op den Akker et al. (27). Comparing how these various types of representation affect motivation and behavior is an interesting direction for future work. For instance, Symons et al. (65) found based on an exploratory study that people preferred GIFs over text reminders and pictures for being motivated to take a brisk walk in a difficult moment. As with the message content, the message representation could also be automatically adapted based on user feedback.

A fourth direction for future work is to tailor the messages to further factors besides mood, self-efficacy, and progress. In the context of persuasive messages more generally, both other dynamic factors (e.g., states derived from the COM-B model

(17)) and more stable user characteristics (e.g., personality (18, 66), cultural background (67), regional factors (68), age (18)) have been shown to affect the effectiveness of different persuasive strategies. As the number of factors increases, so does the cost of collecting (expert-written or crowdsourced) messages tailored to these factors. However, not all of these factors are equally relevant. Albers et al. (20), for example, saw in the context of preparing for quitting smoking that dynamic factors could better predict people's behavior after persuasive attempts than more stable user characteristics. First gaining a thorough understanding of which factors matter is thus important (20).

4.3. Conclusion

In conclusion, we provided a systematic theory-driven way to generate structured motivational messages with the help of experts that is feasible and can thus be used by developers of other eHealth applications for behavior change. Based on a scenario-based evaluation, we found credible support for our hypothesis that messages tailored to mood, self-efficacy, and progress are perceived as more motivating than generic messages in the context of physical activity coaching. Testing the combination of mood, self-efficacy, and progress as tailoring factors is completely new, to the best of our knowledge. A thematic analysis of people's free-text responses about what they find motivating and demotivating about motivational messages provided further support for the use of tailored motivational messages, as well as messages that are encouraging and empathetic, give feedback about people's progress and mention the benefits of physical activity. Our findings thus support the use of motivational messages tailored to mood, self-efficacy, and progress in eHealth applications for physical activity coaching. We share our dataset of motivational messages that can be used during various stages of a user's physical activity intervention, along with a set of scenarios containing the aforementioned user state.

Data availability statement

All data and analysis code is publicly available on 4TU. Research Data (31).

Ethics statement

The studies involving humans were approved by the Human Research Ethics Committee of Delft University of Technology. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

RG, NA, and WB closely collaborated on the drafting and planning of the message generation, experiment, and analysis of the data. KP provided input during the initial drafting of the research. RG created the scenarios, conducted the experiment, analyzed the quantitative data, and created the initial figure drafts. RG, NA, and KP analyzed the qualitative data, and RG, NA, and MV contributed to the analysis code and documentation. NA and WB supervised the entire research process, providing critical feedback and guidance throughout. RG, NA, and WB jointly wrote the article, and KP and MV provided critical feedback on the first manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work is part of the multidisciplinary research project Perfect Fit, which is supported by several funders organized by the Netherlands Organization for Scientific Research (NWO), program Commit2Data - Big Data & Health (project number 628.011.211). Besides NWO, the funders include the Netherlands Organisation for Health Research and Development (ZonMw), Hartstichting, the Ministry of Health, Welfare and Sport (VWS), Health Holland, and the Netherlands eScience Center.

Acknowledgments

This work is based on the study conducted for the publicly available Master's thesis by Ghantasala (69). The authors would like to thank Eline Meijer and Bouke Scheltinga for their input on the study design as well as Debadeep Basu and the second coder for serving as the independent coders for the scenarios.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2023.1215187/full#supplementary-material

References

1. Bol N, Smit ES, Lustria MLA. Tailored health communication: opportunities, challenges in the digital era. *Digit Health*. (2020) 6:2055207620958913. doi: 10.1177/2055207620958913

2. Curry SJ. eHealth research, healthcare delivery: beyond intervention effectiveness. *Am J Prev Med*. (2007) 32:S127–30. doi: 10.1016/j.amepre.2007.01.026

3. Kreuter MW, Wray RJ. Tailored, targeted health communication: strategies for enhancing information relevance. *Am J Health Behav*. (2003) 27:S227–32. doi: 10.5993/AJHB.27.1.s3.6

4. Smeets T, Brug J, de Vries H. Effects of tailoring health messages on physical activity. *Health Educ Res*. (2008) 23:402–13. doi: 10.1093/her/cyl101

5. Rutherford WJ, Corbin CB, Chase LA. Factors influencing intrinsic motivation towards physical activity. *Health Values: The Journal of Health Behavior, Education & Promotion*. (1992) 16(5):19–24.

6. Figueroa CA, Deliu N, Chakraborty B, Modiri A, Xu J, Aggarwal J, et al. Daily motivational text messages to promote physical activity in university students: results from a microrandomized trial. *Ann Behav Med*. (2022) 56:212–8. doi: 10.1093/abm/kaab028

7. Hors-Fraile S, Schneider F, Fernandez-Luque L, Luna-Perejon F, Civit A, Spachos D, et al. Tailoring motivational health messages for smoking cessation using an mHealth recommender system integrated with an electronic health record: a study protocol. *BMC Public Health*. (2018) 18:1–10. doi: 10.1186/s12889-018-5612-5

8. Hors-Fraile S, Malwade S, Luna-Perejon F, Amaya C, Civit A, Schneider F, et al. Opening the black box: explaining the process of basing a health recommender system on the I-change behavioral change model. *IEEE Access*. (2019) 7:176525–40. doi: 10.1109/ACCESS.2019.2957696

9. Maurits K, Ruyter Boris D, Panos M, Emile A. Adaptive persuasive systems: a Study of tailored persuasive text messages to reduce snacking. *ACM Transactions on Interactive Intelligent Systems (TiiS)*. Vol. 2. New York, NY, USA: Association for Computing Machinery (2012). doi: 10.1145/2209310.2209313

10. Ybarra ML, Prescott TL, Holtrop JS. Steps in tailoring a text messaging–based smoking cessation program for young adults. *J Health Commun*. (2014) 19:1393–407. doi: 10.1080/10810730.2014.901441

11. Beristain Iraola A, Álvarez Sánchez R, Petsani D, Hors-Fraile S, Bamidis P, Konstantinidis E. Virtual coaching for older adults at home using smart goal supported behavior change. *Companion Publication of the 2020 International Conference on Multimodal Interaction*. New York, NY, USA: Association for Computing Machinery (2020). p. 251–5. doi: 10.1145/3395035.3425311

12. Kang Y, Tan AH, Miao C. An adaptive computational model for personalized persuasion. *Proceedings of the 24th International Conference on Artificial Intelligence*. AAAI Press / International Joint Conferences on Artificial Intelligence (2015). p. 61–7.

13. Tielman ML, Neerincx MA, Brinkman WP. Design, evaluation of personalized motivational messages by a virtual agent that assists in post-traumatic stress disorder therapy. *J Med Internet Res*. (2019) 21:e9240. doi: 10.2196/jmir.9240

14. Thomas R, Oren N, Masthoff J. Argumessage: a system for automation of message generation using argumentation schemes. *Is ArguMessage Effective? A Critical Evaluation of the Persuasive Message Generation System* (2018). p. 27–31. https://link.springer.com/chapter/10.1007/978-3-030-17287-9_8

15. Cauchard JR, Frey J, Zahrt O, Krister J, Alia C, Landay JA. The positive impact of push vs pull progress feedback: a 6-week activity tracking study in the wild. *Proceedings of the ACM on Interactive, Mobile, Wearable, Ubiquitous Technologies*. New York, NY, USA: Association for Computing Machinery (2019) doi: 10.1145/3351234.

16. Bertolotti M, Carfora V, Catellani P. Different frames to reduce red meat intake: the moderating role of self-efficacy. *Health Commun*. (2019) 35:475–82. doi: 10.1080/10410236.2019.1567444

17. Albers N, Neerincx MA, Brinkman WP. Addressing people's current and future states in a reinforcement learning algorithm for persuading to quit smoking and to be physically active. *PLoS ONE*. (2022) 17:e0277295. doi: 10.1371/journal.pone.0277295

18. Kaptein M, Eckles D. Heterogeneity in the effects of online persuasion. *J Interact Mark*. (2012) 26:176–88. doi: 10.1016/j.intmar.2012.02.002

19. Steward WT, Schneider TR, Pizarro J, Salovey P. Need for cognition moderates responses to framed smoking-cessation messages. *J Appl Soc Psychol*. (2003) 33:2439–64. doi: 10.1111/j.1559-1816.2003.tb02775.x

20. Albers N, Neerincx MA, Brinkman WP. Persuading to prepare for quitting smoking with a virtual coach: using states and user characteristics to predict behavior. *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*. Richland, SC, USA: International Foundation for Autonomous Agents and Multiagent Systems (2023). p. 717–26.

21. Bandura A, Cervone D. Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. *J Pers Soc Psychol*. (1983) 45:1017. doi: 10.1037/0022-3514.45.5.1017

22. Dijkstra A. Working mechanisms of computer-tailored health education: evidence from smoking cessation. *Health Educ Res*. (2005) 20:527–39. doi: 10.1093/her/cyh014

23. Maheswaran D, Meyers-Levy J. The influence of message framing and issue involvement. *J Mark Res*. (1990) 27:361–7. doi: 10.1177/002224379002700310

24. Petty RE, Cacioppo JT. The elaboration likelihood model of persuasion. *Communication and persuasion*. Springer (1986). p. 1–24

25. Bless H, Bohner G, Schwarz N, Strack F. Mood and persuasion. *Pers Soc Psychol Bull*. (1990) 16:331–45. doi: 10.1177/0146167290162013

26. Biddle SJ, Mutrie N, Gorely T, Faulkner G. *Psychology of physical activity: determinants, well-being and interventions*. New York: Routledge (2021).

27. Op den Akker H, Cabrita M, Jones VM, Hermens HJ. Tailored motivational message generation: a model, practical framework for real-time physical activity coaching. *J Biomed Inform*. (2015) 55:104–15. doi: 10.1016/j.jbi.2015.03.005

28. Riet J, Ruiter RA, Werrij MQ, De Vries H. The influence of self-efficacy on the effects of framed health messages. *Eur J Soc Psychol*. (2008) 38:800–9. doi: 10.1002/ejsp.496

29. Hooker SA, Masters KS. Daily meaning salience, physical activity in previously inactive exercise initiates. *Health Psychol*. (2018) 37:344. doi: 10.1037/hea0000599

30. Michie S, Van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci*. (2011) 6:1–12. doi: 10.1186/1748-5908-6-42

31. Ghantasala R, Albers N, Penfornis KM, van Vliet M, Brinkman W-P. Feasibility of generating structured motivational messages for tailored physical activity coaching: data and analysis code. Version 1. 4TU. Research Data. (2023). doi: 10.4121/33888406-2d4e-4365-bf6e-0a45616842ef.v1

32. Ghantasala RP, Albers N. Motivational messages adapted to mood, self-efficacy, and progress (2021). Available from: https://doi.org/10.17605/OSF.IO/MT3FJ

33. Stoet G, Hommel B. Action planning and the temporal binding of response codes. *J Exp Psychol Hum Percept Perform*. (1999) 25:1625. doi: 10.1037/0096-1523.25.6.1625

34. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. (1977) 6:159–74. doi: 10.2307/2529310

35. de Vries RA, Truong KP, Zaga C, Li J, Evers V. A word of advice: how to tailor motivational text messages based on behavior change theory to personality and gender. *Pers Ubiquitous Comput*. (2017) 21:675–87. doi: 10.1007/s00779-017-1025-1

36. Fukuoka Y, Lindgren TG, Mintz YD, Hooper J, Aswani A. Applying natural language processing to understand motivational profiles for maintaining physical activity after a mobile app and accelerometer-based intervention: the mPED randomized controlled trial. *JMIR Mhealth Uhealth*. (2018) 6:e10042. doi: 10.2196/10042

37. Cohen BH. *Explaining psychological statistics*. John Wiley & Sons (2008).

38. Chechile RA. *Bayesian statistics for experimental scientists*. MIT Press (2020). p. 238.

39. Schild C, Lilleholt L, Zettler I. Behavior in cheating paradigms is linked to overall approval rates of crowdworkers. *J Behav Decis Mak*. (2021) 34:157–66. doi: 10.1002/bdm.2195

40. Prolific Team. How do I set up a longitudinal/multi-part study? (2021). Available from: https://researcher-help.prolific.co/hc/en-gb/articles/360009222733-How-do-I-set-up-a-longitudinal-multi-part-study- (Accessed December 21, 2021).

41. Draws T, Rieger A, Inel O, Gadiraju U, Tintarev N. A checklist to combat cognitive biases in crowdsourcing. *Proceedings of the AAAI Conference on Human Computation, Crowdsourcing*. Vol. 9. Palo Alto, California, USA: The AAAI Press (2021). p. 48–59. doi: 10.1609/hcomp.v9i1.18939

42. Prolific Team. Who are the participants on prolific? (2022). Available from: https://researcher-help.prolific.co/hc/en-gb/articles/360009220833-Who-are-the-participants-on-Prolific (Accessed June 20, 2022).

43. van de Schoot R, Depaoli S, King R, Kramer B, Märtens K, Tadesse MG, et al. Bayesian statistics, modelling. *Nat Rev Methods Primers*. (2021) 1:1–26. doi: 10.1038/s43586-020-00001-2

44. McElreath R. *Statistical rethinking: a Bayesian course with examples in R and Stan*. New York: Chapman and Hall/CRC (2020).

45. Andraszewicz S, Scheibehenne B, Rieskamp J, Grasman R, Verhagen J, Wagenmakers EJ. An introduction to Bayesian hypothesis testing for management research. *J Manage*. (2015) 41:521–43. doi: 10.1177/0149206314560412

46. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. (2006) 3:77–101. doi: 10.1191/1478088706qp063oa

47. Nancy Carter R, Bryant-Lukosius D, Alba DiCenso R. The use of triangulation in qualitative research. *Oncology nursing forum*. Vol. 41. Oncology Nursing Society (2014). p. 545–7.

48. Bohner G, Crow K, Erb HP, Schwarz N. Affect and persuasion: mood effects on the processing of message content and context cues and on subsequent behaviour. *Eur J Soc Psychol*. (1992) 22:511–30. doi: 10.1002/ejsp.2420220602

49. Kappen DL, Mirza-Babaei P, Nacke LE. Technology facilitates physical activity through gamification: a thematic analysis of an 8-week study. *Front Computer Sci*. (2020) 2:530309. doi: 10.3389/fcomp.2020.530309

50. Sebri V, Durosini I, Mazzoni D, Pravettoni G. Breast cancer survivors' motivation to participate in a tailored physical and psychological intervention: a qualitative thematic analysis. *Behav Sci*. (2022) 12:271. doi: 10.3390/bs12080271

51. Albers N, Hizli B, Scheltinga BL, Meijer E, Brinkman WP. Setting physical activity goals with a virtual coach: vicarious experiences, personalization and acceptance. *J Med Syst*. (2023) 47:15. doi: 10.1007/s10916-022-01899-9

52. De Vries RA, Truong KP, Kwint S, Drossaert CH, Evers V. Crowd-designed motivation: motivational messages for exercise adherence based on behavior change theory. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery (2016). p. 297–308. doi: 10.1145/2858036.2858229

53. Edmunds J, Ntoumanis N, Duda JL. A test of self-determination theory in the exercise domain. *J Appl Soc Psychol*. (2006) 36:2240–65. doi: 10.1111/j.0021-9029.2006.00102.x

54. Sheeran P, Wright CE, Avishai A, Villegas ME, Rothman AJ, Klein WM. Does increasing autonomous motivation or perceived competence lead to health behavior change? A meta-analysis. *Health Psychol*. (2021) 40:706. doi: 10.1037/hea0001111

55. Sweet SN, Fortier MS, Strachan SM, Blanchard CM. Testing and integrating self-determination theory and self-efficacy theory in a physical activity context. *Can Psychol*. (2012) 53:319. doi: 10.1037/a0030280

56. Grawemeyer B, Mavrikis M, Holmes W, Gutierrez-Santos S, Wiedmann M, Rummel N. Affecting off-task behaviour: how affect-aware feedback can improve student learning. *Proceedings of the Sixth International Conference on Learning Analytics & knowledge*. New York, NY, USA: Association for Computing Machinery (2016). p. 104–13. doi: 10.1145/2883851.2883936

57. Yang MJ, Sutton SK, Hernandez LM, Jones SR, Wetter DW, Kumar S, et al. A just-in-time adaptive intervention (JITAI) for smoking cessation: feasibility and acceptability findings. *Addict Behav*. (2023) 136:107467. doi: 10.1016/j.addbeh.2022.107467

58. Liao P, Greenewald K, Klasnja P, Murphy S. Personalized heartsteps: a reinforcement learning algorithm for optimizing physical activity. *Proc ACM Interact Mob Wearable Ubiquitous Technol*. (2020) 4:1–22. doi: 10.1145/3381007

59. Carfora V, Di Massimo F, Rastelli R, Catellani P, Piastra M. Dialogue management in conversational agents through psychology of persuasion, and machine learning. *Multimed Tools Appl*. (2020) 79:35949–71. doi: 10.1007/s11042-020-09178-w

60. Ghosh S, Chollet M, Laksana E, Morency LP, Scherer S. Affect-LM: a neural language model for customizable affective text generation [Preprint] (2017). Available at: https://doi.org/10.48550/arXiv.1704.06851

61. Wang W, van Lint CL, Brinkman WP, Rövekamp TJ, van Dijk S, van der Boog P, et al. Guided or factual computer support for kidney patients with different experience levels and medical health situations: preferences and usage. *Health Technol*. (2019) 9:329–42. doi: 10.1007/s12553-019-00295-7

62. Whittaker R, Dorey E, Bramley D, Bullen C, Denny S, Elley CR, et al. A theory-based video messaging mobile phone intervention for smoking cessation: randomized controlled trial. *J Med Internet Res*. (2011) 13:e1553. doi: 10.2196/jmir.1553

63. Gill A, Plasqui G, Kok G, Schols AM, Ruiter RA, Kremers SP, et al. Weight-status related differences in reflective and impulsive determinants of physical activity in youngsters (8–18 years old). *Health Psychol Bull*. (2020) 4:29–38. doi: 10.5334/hpb.14

64. Singh A, Klapper A, Jia J, Fidalgo A, Tajadura-Jiménez A, Kanakam N, et al. Motivating people with chronic pain to do physical activity: opportunities for technology design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery (2014). p. 2803–12. doi: 10.1145/2556288.2557268

65. Symons M, Vandebosch H, Cutello CA, Poels K. Message reminders encouraging brisk walking by considering the dynamic factor of cognitive fatigue. *Eur J Health Commun*. (2022) 3:41–68. doi: 10.47368/ejhc.2022.303

66. Zalake M, de Siqueira AG, Vaddiparti K, Lok B. The effects of virtual human's verbal persuasion strategies on user intention and behavior. *Int J Hum Comput Stud*. (2021) 156:102708. doi: 10.1016/j.ijhcs.2021.102708

67. Oyibo K, Adaji I, Orji R, Olabenjo B, Vassileva J. Susceptibility to persuasive strategies: a comparative analysis of Nigerians vs. Canadians. *Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization*. New York, NY, USA: Association for Computing Machinery (2018). p. 229–38. doi: 10.1145/3209219.3209239

68. Xu W, Legaspi R, Ishikawa Y. Does the association between persuasive strategies and Personality types vary across regions. *Persuasive Technology: 18th International Conference, PERSUASIVE 2023, Proceedings*; 2023 Apr 19–21; Eindhoven, The Netherlands. Springer (2023). p. 359–68.

69. Ghantasala R. *Motivating, your way: tailoring your fitness journey* [Master's thesis] (2022). http://resolver.tudelft.nl/uuid:d2a3e006-5d75-404f-a0b9-89423ea97c63

frontiers | Frontiers in Digital Health

TYPE Systematic Review
PUBLISHED 20 October 2023
DOI 10.3389/fdgth.2023.1253390

OPEN ACCESS

EDITED BY
Kim Cornelia Marie-Louise Bul,
Coventry University, United Kingdom

REVIEWED BY
Erik Hoencamp,
University Department Psychology, Netherlands
Karin Waldherr,
Ferdinand Porsche FernFH—Distance Learning University of Applied Sciences, Austria

*CORRESPONDENCE
Karolina Morello
karolina.morello@lir-mainz.de

†These authors share last authorship

RECEIVED 05 July 2023
ACCEPTED 09 October 2023
PUBLISHED 20 October 2023

CITATION
Morello K, Schäfer SK, Kunzler AM, Priesterroth L-S, Tüscher O and Kubiak T (2023) Cognitive reappraisal in mHealth interventions to foster mental health in adults: a systematic review and meta-analysis.
Front. Digit. Health 5:1253390.
doi: 10.3389/fdgth.2023.1253390

Cognitive reappraisal in mHealth interventions to foster mental health in adults: a systematic review and meta-analysis

Karolina Morello[1]*, Sarah K. Schäfer[1,2], Angela M. Kunzler[1,3], Lilli-Sophie Priesterroth[4], Oliver Tüscher[1,5,6†] and Thomas Kubiak[4†]

[1]Leibniz Institute for Resilience Research (LIR) Mainz, Mainz, Germany, [2]Clinical Psychology, Psychotherapy and Diagnostics, Technische Universität Braunschweig, Brunswick, Germany, [3]Institute for Evidence in Medicine, Medical Center - University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany, [4]Health Psychology, Institute of Psychology, Johannes Gutenberg University Mainz, Mainz, Germany, [5]Department of Psychiatry and Psychotherapy, University Medical Center of the Johannes Gutenberg University Mainz, Johannes Gutenberg University Mainz, Mainz, Germany, [6]Institute of Molecular Biology (IMB), Mainz, Germany

Background: An increasing number of mHealth interventions aim to contribute to mental healthcare of which interventions that foster cognitive reappraisal may be particularly effective.
Objectives: To evaluate the efficacy of mHealth interventions enhancing cognitive reappraisal to improve mental health in adult populations.
Methods: The literature search (four databases) yielded 30 eligible randomized controlled trials (comprising 3,904 participants). We performed a multi-level meta-analysis to examine differences between intervention and comparator conditions at post-intervention assessment. Moderator analyses were conducted for potential moderator variables (e.g., type of comparators).
Results: Most interventions were CBT-based with other training components in addition to cognitive reappraisal. We found preliminary evidence for a small to medium effect favouring mHealth interventions to enhance cognitive reappraisal over comparators, M(SMD) = 0.34, p = .002. When analysing single symptoms, there was evidence for a small to medium effect of mHealth interventions on anxiety and depressive symptoms, but not for psychological distress and well-being. All analyses showed substantial heterogeneity. Moderator analyses revealed evidence for more favourable effects in studies with passive comparators. There was an overall high risk of bias in most of the studies.
Conclusions: We found preliminary evidence for a small to medium effect of mHealth interventions including a cognitive reappraisal component to improve mental health. However, most of the interventions were complex (i.e., reappraisal was provided alongside other components), which prevents us from examining reappraisal-specific effects beyond general mental health promotion in mHealth. Dismantling studies examining the effects of single intervention components are warranted to corroborate these promising results.

Systematic Review Registration: https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=142149, identifier [CRD42019142149].

KEYWORDS

mHealth, mental health apps, ecological momentary intervention, mental health, reappraisal, cognitive restructuring, systematic review, meta-analysis

Introduction

The high prevalence of mental disorders is a key challenge for healthcare (1). Digital technologies such as mobile phones can potentially improve the dissemination of evidence-based mental health interventions (2, 3). Thus, numerous mobile-based programs, also referred to as mobile health (mHealth) interventions, were developed and distributed in the last years. These aim at improving mental health and well-being or reducing symptoms of mental disorders of their users. Using mobile apps, trainings can easily be delivered to users in their everyday lives and natural settings in a timely fashion (4).

mHealth apps can be implemented as stand-alone interventions, that might also be used as easy-to-access self-help programs. Alternatively, mHealth can be used as an add-on to face-to-face treatments, e.g., for therapeutic guidance (2). Thus, mHealth is an easily accessible, low-cost method that has the potential to help overcoming supply gaps in mental healthcare (5, 2, 6).

A number of clinical trials as well as recent systematic reviews and meta-analyses pointed to an increasing importance of mHealth interventions for mental health care and promotion (3, 7–12). Those studies provided evidence for mHealth interventions being feasible, acceptable, and effective in different populations and for different purposes. In line with common classifications (13, 14), mHealth interventions can be used in the following domains: First, many mHealth interventions aim at reducing symptoms of mental disorders (i.e., treatment) such as anxiety disorders and depression (15–18, 11). However, not only patients with manifest mental disorders can be targeted by mHealth interventions. For at-risk groups, subpopulations, or the general community mHealth interventions may also have the potential to prevent mental disorders [i.e., prevention (19)] or to promote positive aspects of functioning and well-being [i.e., mental health promotion (20)]. Research focuses primarily on mHealth interventions, which aim at reducing symptoms of mental disorders (21, 9), but there is some evidence on the effects of mHealth interventions for the general community (22, 20) which highlights the benefits for a broader target group.

The increasing popularity of mHealth interventions to reduce symptoms of mental illness as well as to promote mental health claims for their critical evaluation to provide users with effective and safe technologies. However, the scientific evaluation of mHealth trainings is still in its beginnings and the integration of scientifically validated theories and strategies into mHealth interventions in order to maximize a favorable mental health impact often remains unclear (23–25). For example, a recent systematic review examined the extent to which evidence-based contents are included in popular smartphone apps for the treatment of depression and anxiety (26). The study found that most of these apps included at least one evidence-based treatment element. However, specific evidence-based components were found rarely in those apps. Thus, there is evidence for a gap between empirically supported treatment components and the components used in mHealth interventions (26). Moreover, it becomes apparent that many mHealth interventions are multi-component interventions (27). They offer a range of components in combination with each other, of which some are evidence-based whereas others are not, which further complicates a statement on the effectiveness of single components.

One of the above-mentioned evidence-based intervention components is cognitive reappraisal, which pertains to modify dysfunctional thoughts. Cognitive reappraisal is an emotion regulation strategy that has been defined as creating alternative appraisals or interpretations of a potentially emotion-eliciting situation to change its emotional impact (28–30). As a strategy of cognitive change, reappraisal can be achieved through techniques of cognitive restructuring, reframing or reinterpretation, which represent key tactics of cognitive behavioral therapy [CBT (31, 32)].

Evidence comes from meta-analyses on the impact of reappraisal on mental health: Cognitive reappraisal has been found to be negatively correlated with symptoms of mental distress (33–36) and positively with indicators of mental health such as well-being (37, 35). Moreover, reappraisal appears to be associated with successful coping and maintaining of well-being when facing stressful events [i.e., resilience (38)]. Therefore, cognitive reappraisal might be a valuable ingredient for all types of mHealth interventions (treatment, prevention, or mental health promotion).

It is well established that reappraisal processes play a central role in the development and maintenance of mental illness, as changing dysfunctional cognitions is essential to reduce symptoms of mental disorders (31). Following a mechanistic approach, reappraisal might be seen as an active therapeutic mechanism. This means, that an effective psychological treatment might enhance reappraisal skills of an individual, which in turn contributes to symptom change (39). In line with this notion, there is evidence from CBT-based interventions that changes in symptom-related cognitions (i.e., reappraisal) mediate changes in symptoms, for example of panic disorder (40) and post-traumatic stress disorder [PTSD (41)].

On the other hand, research on the specific effects of reappraisal components in CBT interventions is still ongoing. Some mixed evidence is coming from research on the effectiveness of specific intervention components in complex CBT interventions. While some studies found no evidence for specific effects of any components (42) or no additive effect of reappraisal-related components (43, 44), Pompoli et al. (45) found that cognitive restructuring components are associated with the largest remission rates in panic disorder.

Notably, this research on reappraisal components does not focus the setting of mHealth interventions. Therefore, it remains unclear whether the advantages of mHealth, such as facilitating a continuous training of cognitive reappraisal in everyday life at a time, when a person needs support, can be put into practice.

Therefore, this systematic review and meta-analysis aims to examine the efficacy of psychological mHealth interventions comprising a component of cognitive reappraisal to enhance mental health in clinical and non-clinical adult populations.

Methods

A review protocol was registered with PROSPERO (ID: CRD42019142149). The review is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analysis [PRISMA (46)]. We used the web-based platforms Covidence (47) and Rayyan (48) for the screening process.

Search strategy

A comprehensive literature search in MEDLINE (via PubMed), PsycArticles (via Ebscohost), Embase (via Ovid) and CENTRAL was conducted (search starting from the incept of each database; last update: March 7, 2022). The search process for this review integrated search terms for several higher-level resilience mechanisms identified in a landmark theory paper [(49); i.e., cognitive reappraisal, cognitive flexibility, extinction learning, interference inhibition, and stress immunization]. Of these resilience mechanisms, only cognitive reappraisal is relevant for the present review. Records addressing other resilience mechanisms than cognitive reappraisal were excluded throughout the process of study selection to ensure sufficient comparability of the included studies. The search strategy comprised four clusters of search terms that were searched in title, abstracts and (partly) in keywords: (1) cognitive reappraisal (e.g., cognitive restructuring, emotion regulation); (2) intervention (e.g., training, program); (3) mobile delivery [e.g., ecological momentary intervention (EMI), mobile application] and (4) study design [e.g., randomized controlled trial (RCT)]. Within clusters, search terms were combined using the Boolean operator OR, while clusters were linked using the operator AND. The search strategies for all databases are presented in **Supplementary Material A**.

Study selection

All identified records were screened by two independent reviewers. First, clearly irrelevant records were excluded at title/abstract level. Second, eligibility was assessed at full-text level (kappa = .65), where we examined whether the mHealth interventions contained at least one reappraisal-enhancing component. For this purpose, we retrieved additional materials (i.e., study protocols, websites, app store or play store entries) related to the respective mHealth interventions, to gather information on intervention components. Discrepancies were resolved through discussion or by consulting a third author at each stage, and consensus was achieved in all cases.

Eligibility criteria

Studies were eligible if they met the following criteria: (a) An adult study sample (≥18 years) from a clinical or non-clinical population was assessed. (b) The intervention used mobile devices (e.g., smartphones, tablets) to deliver intervention components to participants in their everyday lives. We considered stand-alone mHealth interventions as well as interventions that included a mobile component as an adjunct to other delivery formats (e.g., face-to-face). The examined interventions did have to address cognitive reappraisal and provide related techniques (e.g., cognitive restructuring; see **Supplementary Material B** for more information). (c) Included studies were RCTs, controlled trials (CTs), or cluster-randomized trials with a comparison of the mHealth intervention against any control group (e.g., wait-list control, no treatment control) or any other treatment not comprising a cognitive reappraisal component (e.g., face-to-face intervention, treatment as usual, other mHealth intervention). (d) The studies assessed at least one mental health outcome (primary outcome), that is, depressive symptoms and anxiety symptoms, measures of general psychological distress, and well-being outcomes (e.g., quality of life). Secondary outcome was cognitive reappraisal. We considered results of self-reported questionnaires, for example the Patient Health Questionnaire [PHQ-9 (50)] for depression, or the Cognitive Emotion Regulation Questionnaire [CERQ (51)] for cognitive reappraisal, and clinician-administered interviews. (e) Eligible studies were peer-reviewed publications published in English.

Data extraction and coding

Two independent reviewers extracted relevant information of the included studies using a predefined data extraction sheet. The following information were extracted: (a) study details (e.g., first author, publication year, country); (b) sample characteristics (e.g., sample size, mean age, gender balance); (c) intervention characteristics (e.g., name of the mHealth intervention, intervention length, availability of human support, proportion of cognitive reappraisal); (d) study methods (e.g., design, type of comparator); (e) mental health outcomes and measures.

For all included studies, we defined relevant intervention arms and outcome measures. Arms comprising a mHealth intervention with a cognitive reappraisal component were defined as intervention group; other arms were defined as comparators. For each study, all reported outcomes of interest for these meta-analyses were considered. More details on data extraction and coding (e.g., definition of availability of human support, intervention type, intervention triggering as well as proportion of cognitive reappraisal) are presented in **Supplementary Material B**.

Quality and bias assessment

Risk of bias of primary studies

As only RCTs were identified, risk of bias was assessed independently and in duplicate in the following domains using the Cochrane Risk of Bias Assessment Tool 2 [RoB2 (52)]: (a)

bias arising from the randomization process; (b) bias due to deviations from intended interventions; (c) bias due to missing outcome data; (d) bias in measurement of the outcome; and (e) bias in selection of the reported result. With one identified crossover RCT, we proceeded according to the additional considerations for crossover trials of the RoB2 tool. In addition to bias ratings per domain, the overall risk of bias was assessed at study level. Ratings could be "low" or "high" risk of bias or express "some concern" (52). The interrater agreement was high (96.2%), and all disagreements were resolved through discussion.

Publication bias

The potential impact of publication bias was examined by using visual inspections of contour-enhanced funnel plots (53) as well as by means of statistical analyses by approximating the Begg Mazumdar rank correlation test (54). This test is not available for multilevel meta-analysis but can be approximated by including the sampling error as moderator variable to the main analysis. In case the sampling error would significantly predict study effect sizes, this can be interpreted as evidence of a non-normal distribution of effect sizes and thus, suggest the presence of a publication bias.

Data synthesis

Included studies were summarized narratively and in tabular form. Pairwise meta-analyses (intervention vs. comparator) were performed for primary outcomes if at least two studies were available, and these were sufficiently homogeneous in terms of intervention types and outcomes. For studies with multiple intervention arms, it was determined which arm was most relevant for this review. In case more than one arm was relevant, these were averaged according to the recommendations of the Cochrane Collaboration (55). For our meta-analyses, we included multiple types of comparators (e.g., wait-list and active controls, face-to-face interventions). However, we examined the impact of this analytical decision by means of subgroup moderator analyses. In case data needed for effect size calculation was missing or ambiguous, study authors were contacted by the review team.

All statistical analyses were performed in *R* version 4.2.3 (56) using the package *metafor* (57). As we expected relevant between-study heterogeneity, all analyses used random-effects models and maximum likelihood estimations.

As effect size measure, we used standardized mean differences (SMDs, Hedges' *g*) at post-intervention assessment and 95% confidence intervals (CIs) as indicators of their significance. SMDs were calculated based on the means, standard deviations (SDs), or alternative statistical information (e.g., standard error), with positive SMDs indicating a favorable effect of the intervention over the comparator condition (i.e., better mental health at post-intervention assessment). Moreover, as we expected to find heterogeneous results, 95% prediction intervals (PIs) were used to account for uncertainty of meta-analytical findings (55). PIs provide an estimate of the interval in which 95% of future observations will fall. In line with previous recommendations, PIs were calculated when 10 or more effect sizes were available per analysis (55).

Our main analysis aimed at answering the question of whether there is an effect of mHealth interventions aiming to enhance cognitive reappraisal on overall mental health. For this purpose, we used a multilevel approach nesting effect sizes of mental health indicators defined as primary outcomes within studies and outcome types [i.e., depressive symptoms and anxiety symptoms, measures of general psychological distress, and well-being (58)]. Thereby, our model allowed for correlations of effect sizes coming from the same study as well as for correlations of effect sizes coming from different studies but assessing the same outcome type (e.g., depressive symptoms).

If a study used more than one questionnaire to assess an outcome domain, for example the PHQ-9 and BDI-II (59) for depressive symptoms, we used the questionnaire that was administered most frequently across all studies in our analyses. In the few cases, where outcome data were available for more than one time point after the end of intervention, data of the post-intervention assessment was used to ensure between-study comparability.

Additional analyses aimed at answering the question of whether there is an effect of mHealth interventions aiming to enhance cognitive reappraisal on specific symptom types (e.g., depressive symptoms) or mental well-being. For these analyses, we used traditional meta-analyses for all primary outcomes.

Statistical heterogeneity was assessed using the Cochran's *Q* statistic (60), with a significant *Q* statistic indicating the presence of heterogeneity. To further describe the amount of heterogeneity in our analyses, we used the I^2 statistics (range: 0%–100%), with values of 50% and above indicating substantial heterogeneity (55). As part of our multilevel approach, we estimated heterogeneity related to between-study and between-outcome differences separately (58).

We had planned several moderator analyses in advance (see preregistration: CRD42019142149), which were also conducted due to the substantial heterogeneity as indicated by the above-mentioned indicators. All moderator analyses were performed using the multilevel approach of our primary analysis. We used meta-regressions for moderator analyses, that is, omnibus moderation tests (Q_M) for categorical moderators (e.g., type of comparator) and to examine the impact of continuous moderators (e.g., sample mean age), with a significant Q_M test indicating the presence of a moderator effect. In case of categorical moderators, we obtained subgroup estimates from the multilevel model. Moderator analyses were performed for sociodemographic sample characteristics (i.e., mean sample age, and proportion of female participants), year of publication, population type (clinical vs. non-clinical population), proportion of cognitive reappraisal enhancing components, intervention duration, intervention type (treatment vs. prevention vs. mental health promotion), availability of human support, stand-alone mHealth interventions vs. combined interventions (e.g., mHealth intervention with face-to-face sessions) and comparator type (passive vs. active control condition).

Sensitivity analyses were performed based on the multilevel model used for primary analysis and examined the impact of risk of bias (only for risk of bias domains with sufficient

between-study variation), the inclusion of a crossover RCT for which no ideal outcome data was available, and the impact of considering studies in which the cognitive reappraisal condition was defined as control condition by primary study authors.

Results

The literature search yielded 2,502 records after duplicates had been removed for title/abstract screening (see **Figure 1**). Of these records, 266 were assessed on the full-text level. Full-text level screenings resulted in 30 studies (from 34 reports) that were eligible for the review and included in the qualitative and quantitative summary.

Study characteristics

Supplementary Material C provides an overview of the characteristics of the 30 studies [corresponding to 34 reports (61–95)] comprising a total of 5,377 participants at baseline ($n_{IG} = 2{,}697$, $n_{CG} = 2{,}680$). All studies were RCTs, including one crossover RCT (72). They were all published recently (between 2014 and 2022). Ten studies were conducted in the United States, five in Germany, three were from the Republic of Korea, two from Iran and two from China, with the remaining studies being from Canada, Australia, Japan, South Africa, Sweden, Switzerland, Vietnam, and the United Kingdom.

The mean age of the participants ranged from 18.68 to 74.93 years. The mean percentage of female participants across all studies was 68.51% ($SD = 18.64$) indicating an overrepresentation of female participants. Most studies assessed non-clinical and/or sub-clinical samples ($n = 20$), whereas samples of 10 studies were clinical (i.e., referred to as patients).

Nevertheless, also in non-clinical samples, most of the participants showed elevated symptoms of mental disorders (15 out of 20 non-clinical samples) but were not referred to as patients in a clinical setting and therefore classified as non-clinical. In total, participants of 23 out of 30 studies showed at least some elevated mental health symptoms. Sixteen studies

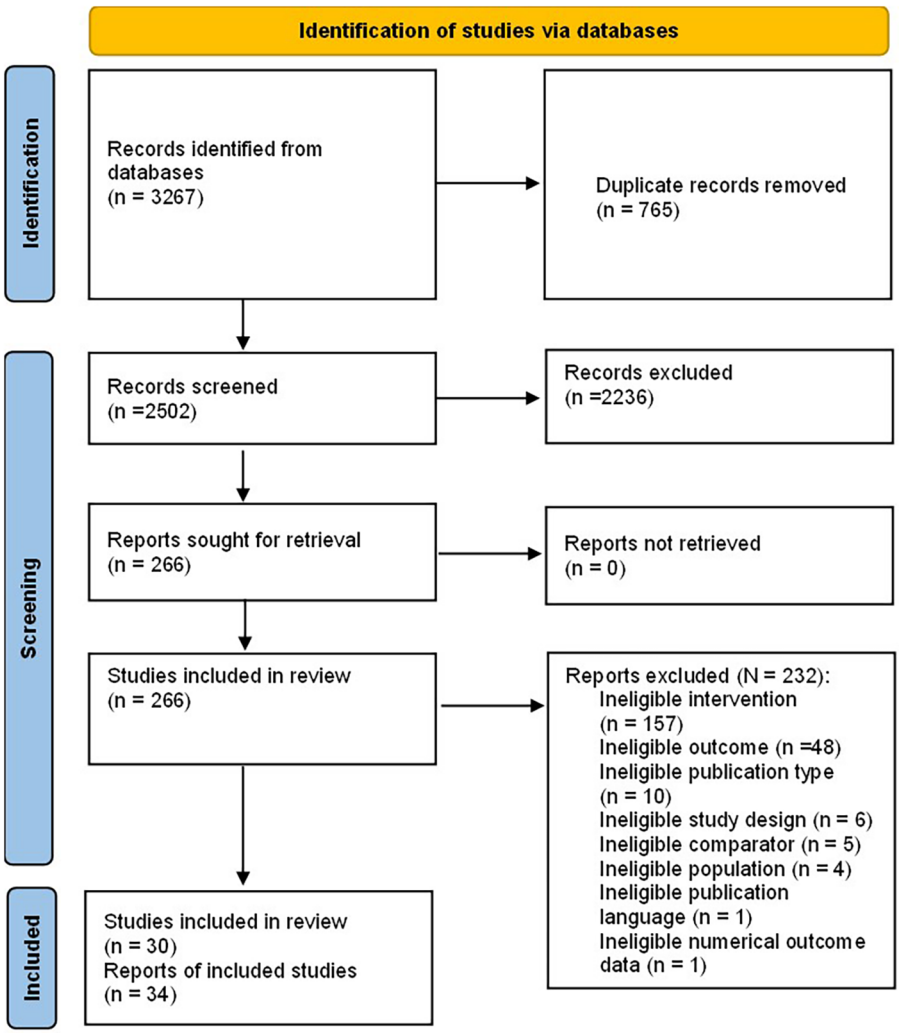


FIGURE 1
PRISMA flow diagram.

assessed participants with depressive symptoms, nine studies reported on samples with anxiety or trauma-related symptoms, and one study examined patients with schizophrenia, schizoaffective disorder or bipolar I disorder (69). Six studies reported on samples with somatic diseases, which were the human immunodeficiency virus [HIV (94) epilepsy (61, 86), irritable bowel syndrome (72), and cancer (70, 71)]. Two studies assessed samples without any somatic or mental disease (63, 77). In line with the notion that most samples showed symptoms of mental disorders, the majority of studies ($n = 23$) mostly aimed at reducing mental symptoms. Five studies were prevention interventions and two studies focused on the promotion of mental health. Most studies primarily aimed at reducing depressive symptoms. Intervention duration ranged from 14 to 112 days ($M = 54.57$, $SD = 26.63$).

All mHealth interventions used smartphones and/or apps. For some interventions, materials could alternatively be accessed via computer or tablet, if preferred (66, 74, 86, 90). Within the mHealth interventions, the full range of multimedia materials was presented: Interventions used text-based elements combined with other materials like voice mails (77), cartoons (81, 96), interactive games (97, 90), and video material (66, 70). In addition to educational intervention elements, all studies required active engagement from the participants. Active engagement was facilitated by various types of interaction with the app or the study personnel, for example, diary entries of negative beliefs (93), multi-step questions (73, 81), rating scales (69), or games (90). Some studies were based on messenger apps (94), or integrated a chatbot (89). The majority of studies ($n = 22$) did not provide human support in addition to the mHealth intervention, while in eight studies human guidance or feedback was available. Moreover, most mHealth interventions ($n = 23$) were established in a stand-alone manner without non-mHealth components. Most studies ($n = 14$) delivered the intervention components on-demand, that is, the training was available at any time throughout the study period. Another five mHealth interventions prescribed a fixed schedule for usage (e.g., three exercises per day) and one intervention triggered participants (i.e., prompts to engage in exercises). Ten studies used a combination of on-demand, fixed and triggered delivery.

Most studies ($n = 23$) evaluated the mHealth interventions against a passive comparator (i.e., wait-list, treatment as usual, no intervention), whereas 14 studies used active comparators, with most studies using a comparable mobile app with educational material as active comparator. Seven studies assessed both, a passive as well as an active control group.

Three studies (81, 77, 69) used clinician-administered interviews for outcome assessment whereas the outcomes of the remaining studies were self-reported. Depressive symptoms were assessed in the majority of studies ($n = 27$). Nineteen studies reported on anxiety symptoms, 14 studies on well-being, and two on general psychological distress.

Implementation of cognitive reappraisal

All interventions were theoretically based on cognitive behavioral therapy (CBT), cognitive theory (92), or cognitive behavioral stress management [CBSM, (94)]. Most of the identified mHealth interventions ($n = 27$) were multi-component interventions, meaning that cognitive reappraisal was promoted alongside other training components (e.g., behavioral activation). There were three studies (67, 73, 92) solely focusing on cognitive reappraisal. **Supplementary Material D** gives an overview of the components of each study as well as the relative importance of the cognitive reappraisal component in relation to the full intervention. The proportion of cognitive reappraisal in the interventions was relatively low ($M = 28.27\%$, $SD = 25.21$).

Study quality

Risk of bias

There was an overall high risk of bias in 27 of 30 studies (see **Figure 2** and **Supplementary Material E**). Main flaws (≥20% some concerns or high-risk ratings) across the included studies and outcomes came from outcome measurement (89.7%, i.e., most studies used self-report measures for outcome assessment and participants were not blinded), deviations from

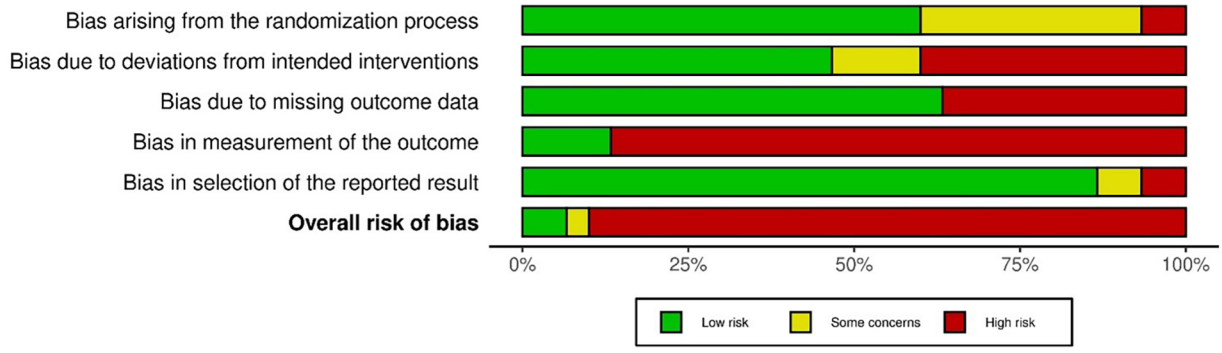


FIGURE 2
Distribution of risk-of-bias ratings within each of five bias domains and overall risk of bias according to the Cochrane risk-of-bias tool for randomized trials (RoB2) (52). All percentages are related to 30 studies (62 effect sizes) included in the review.

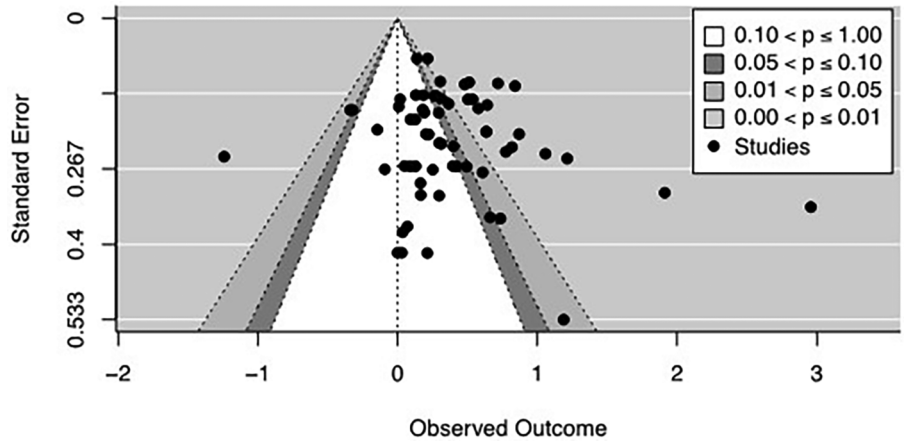


FIGURE 3
Contour-enhanced funnel plot of the main analysis. Contour-enhanced funnel plots are used to highlight significance levels of each effect size in the plot and allow to examine whether studies are missing in specific areas of the plot, that is, the non-significant areas or findings pointing to the opposite direction. For the present analysis, the plot points to a mild publication bias with a relevant number of studies being at the significant side of the right border region between the non-significant and significant area.

intended interventions (53.3%), and missing outcome data (36.7%). For the study with a crossover design (72), risk of bias arising from period and carryover effects was low.

Publication bias

The meta-regression model provided evidence for a trend towards an association of sample errors and effect size estimates for the main analysis, $Q_M(1) = 3.20$, $p = .074$, that is, the effect sizes are likely to violate the assumption of a normal distribution. The visual inspection of the contour-enhanced funnel plot suggested the presence of a mild publication bias (see **Figure 3**), with effect sizes being more likely to fall into the right significant border area of the funnel plot, while the number of studies in non-significant areas was smaller.

Pairwise meta-analysis—intervention effect

Thirty studies [comprising 62 effect sizes from 9,458 observations of 3,904 participants (intervention: 4,692 observations from 1,881 participants; comparators: 4,766 observations from 2,023 participants)] were included in our quantitative synthesis (see **Figure 4**). Using a multilevel approach, we compared post-intervention means across all mental health outcomes between mHealth interventions to enhance cognitive reappraisal and any comparator condition. Moreover, we examined group differences of post-intervention means for single mental health indicators (see **Table 1**). Across all mental health outcomes, we found evidence for a small to medium effect favouring mHealth interventions to enhance cognitive reappraisal over comparators, $M(SMD) = 0.34$, 95% CI [0.12, 0.56], $p = .002$, with considerable heterogeneity between studies and mental health outcomes, $Q(61) = 295.42$, $p < .001$; $I^2 = 86.5\%$. However, the majority of the total heterogeneity was accounted for by between-study differences ($I^2 = 78.5\%$), while only 8.0% originated from between-outcome differences. The presence of considerable heterogeneity was further supported by the wide 95% PI covering small to medium adverse and large favourable intervention effects.

When analysing single symptoms, there was evidence for a small to medium effect of mHealth interventions to enhance cognitive reappraisal on anxiety symptoms, $M(SMD) = 0.33$, 95% CI [0.16, 0.50], $p = .001$, and depressive symptoms, $M(SMD) = 0.51$, 95% CI [0.30, 0.73], $p = <.001$. However, there was no evidence in favour of mHealth interventions to enhance cognitive reappraisal for psychological distress, $M(SMD) = 0.20$, 95% CI [−0.25, 0.66], $p = .380$, and well-being, $M(SMD) = 0.18$, 95% CI [−0.05, 0.42], $p = .126$. Overall, between-study heterogeneity was considerable for all analyses ($I^2 \geq 76.7\%$) except for the meta-analysis on psychological distress ($I^2 = 60.6\%$), however, this was based on only two studies.

Moderator analyses

Moderator analyses were performed for the main analysis comprising all mental health outcomes.

Sociodemographic variables

We found no evidence for a moderating effect of mean sample age, $Q_M(1) = 1.34$, $p = .247$, and proportion of female participants per sample, $Q_M(1) = 0.15$, $p = .698$.

Year of publication

We found no evidence for an association of publication year and effect size estimates, $Q_M(1) = 0.01$, $p = .907$, that is, more recent mHealth interventions were not associated with larger effects.

Population type

We found no evidence for differences between intervention effects in clinical (i.e., patients) and non-clinical populations,

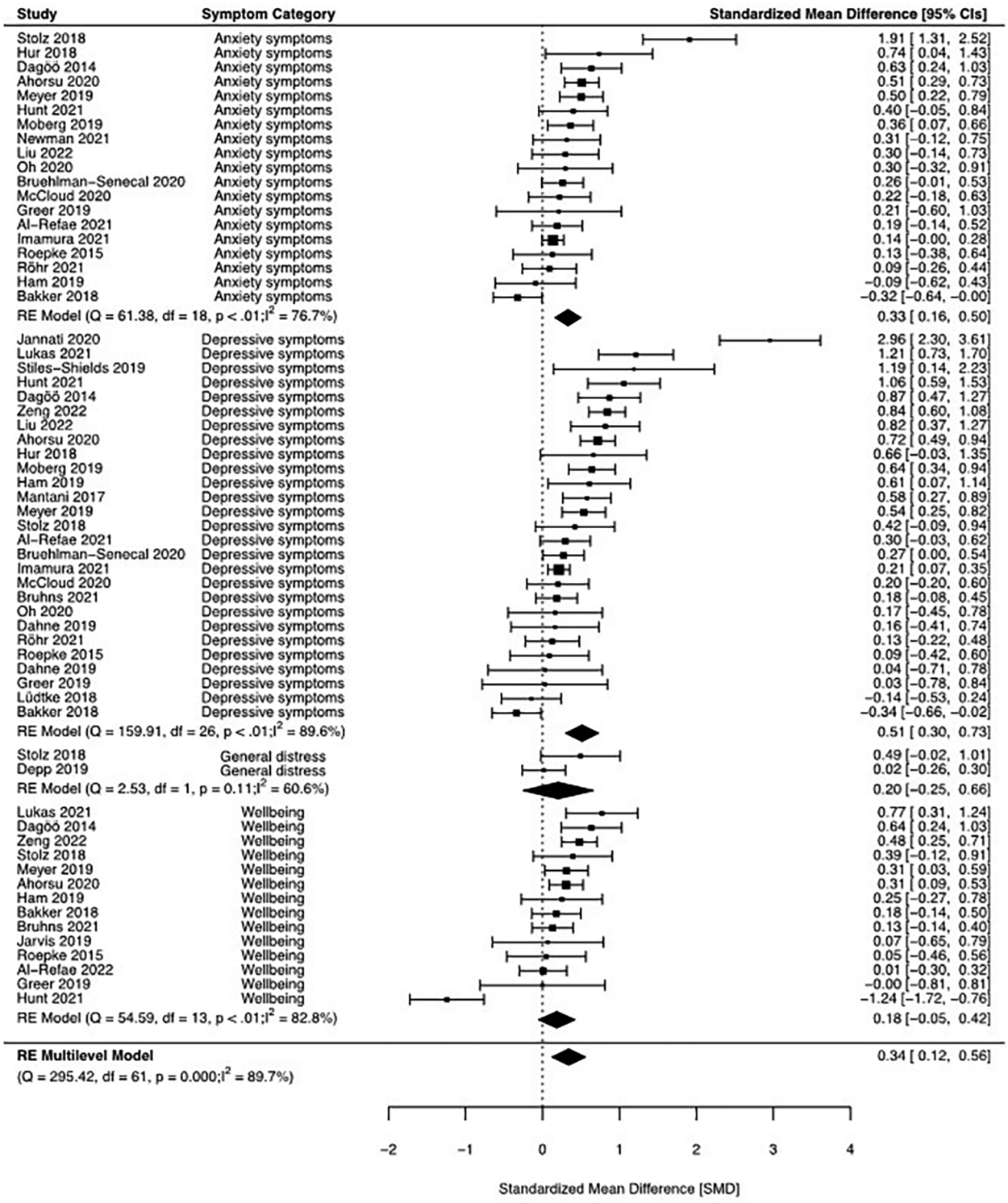


FIGURE 4
Forest plot of the multilevel meta-analysis (main analysis) comprising all mental health outcomes and traditional meta-analyses on single outcome categories. Positive effect size estimates indicate favorable effects of an intervention (i.e., reductions of mental symptoms or increases in well-being). CI, confidence interval; DF, degree of freedom; I^2, heterogeneity index in percentage (range: 0%–100%); Q, Cochran's Q statistic with p value.

$Q_M(1) = 0.27$, $p = .602$ (see Table 2). Further, there was no evidence for differences in effects between populations with elevated symptoms of mental disorders compared to populations without symptoms of mental disorders (See Supplementary Material F).

Intervention type

As not enough studies in the category mental health promotion ($n = 2$) were available, we performed the moderator analysis only for treatment vs. prevention. We found no

10.3389/fdgth.2023.1253390

TABLE 1 Results of main analyses (multilevel and traditional meta-analyses) for primary outcomes comparing mHealth interventions to enhance cognitive reappraisal with comparator conditions.

				Confidence interval		Prediction interval						
Analysis	*n*	*k*	*M*(SMD)	95%$_l$	95%$_u$	95%$_l$	95%$_u$	*p*	Q	df	*p*(Q)	I^2 (I^2_S, I^2_O)
Mental health (ML)	30	62	0.34	0.12	0.56	−0.54	1.23	.002	295.42	61	<.001	86.5 (78.5, 8.0)
Anxiety symptoms (T)	19	19	0.33	0.16	0.50	−0.31	0.96	<.001	61.38	18	<.001	76.7
Depressive symptoms (T)	27	27	0.51	0.30	0.73	−0.53	1.55	<.001	159.91	26	<.001	89.6
Psychological distress (T)	2	2	0.20	−0.25	0.66	–	–	.380	2.54	1	.111	60.6
Well-being (T)	14	14	0.18	−0.05	0.42	−0.62	0.99	.126	54.59	13	<.001	82.8

The multilevel meta-analysis on mental health indicators included anxiety symptoms, depressive symptoms, psychological distress, and well-being. For all outcomes, SMDs were calculated in a way that positive values indicate favorable effects of an intervention (i.e., reductions of mental symptoms or increases in well-being).
Df, degrees of freedom; I^2, heterogeneity index in percentage (range: 0%–100%); I^2_S, heterogeneity accounted for by between-study differences; I^2_O, heterogeneity accounted for by between-outcome differences; *k*, number of effect sizes; ML, multilevel meta-analysis, *n*, number of studies; Q, Cochran's *Q* statistic with *p* value; SMD, standardized mean difference; T, traditional meta-analysis; 95%$_l$, lower boundary of the 95% confidence/prediction interval; 95%$_u$, upper boundary of the 95% confidence/prediction interval.

TABLE 2 Results of the moderator analyses for categorical moderators.

Analysis	*n*	*k*	*M*(SMD)	95%CI$_l$	95%CI$_u$	*p*
Population type						
Clinical populations	10	22	0.28	−0.04	0.60	.083
Non-clinical populations	20	40	0.37	0.12	0.63	.004
$Q_M(1) = 0.27$, $p = .602$						
Intervention Type						
Treatment	23	45	0.39	0.12	0.65	.004
Prevention	5	11	0.19	−0.25	0.64	.397
$Q_M(1) = 0.73$, $p = .394$						
Availability of human support						
Human support available	8	16	0.50	0.15	0.85	.005
No human support available	8	22	0.28	0.03	0.52	.026
$Q_M(1) = 1.36$, $p = .243$						
Stand alone vs. combined						
Stand alone	7	9	0.37	0.14	0.61	.002
Add-on	23	53	0.23	−0.16	0.61	.254
$Q_M(1) = 0.49$, $p = .484$						
Control Condition						
Active comparator	14	30	0.13	−0.12	0.39	.316
Passive comparator	23	45	0.44	0.21	0.68	<.001
$Q_M(1) = 11.90$, $p = .001$						

Results for the categorical moderator variables in the multilevel model including all outcomes, i.e., anxiety symptoms, depressive symptoms, psychological distress, and well-being. Again, positive SMDs indicate positive intervention effects. We present results of omnibus moderator tests (i.e., Q_M test) along with estimates per subgroup (obtained from the multilevel model). To note, for the moderator analysis on control condition, we included all effect sizes reporting on active and passive comparators separately (nested within studies), thus, the number of effect sizes included in this analysis exceeds the number of effect sizes included in the main analysis. Further, the subgroup mental health promotion contains only two studies and was left out from the analysis on intervention type. Therefore, the number of studies in the moderator analysis is lower than in the main analysis.
Q_M(df), omnibus moderator test with degrees of freedom; *k*, number of effect sizes; *n*, number of studies; SMD, standardized mean difference, 95% CI$_l$, lower boundary of the 95% confidence interval; 95% CI$_u$, upper boundary of the 95% confidence interval.

evidence for differences between intervention effects in treatment and prevention interventions, $Q_M(1) = 0.73$, $p = .394$ (see Table 2).

Proportion of cognitive reappraisal component

We found no evidence for interventions comprising a higher proportion of cognitive reappraisal to be associated with larger intervention effects, $Q_M(1) = 0.02$, $p = .896$.

Intervention duration

We found no evidence for more intense interventions, that is, interventions with more days, being associated with larger intervention effects, $Q_M(1) = 0.39$, $p = .533$.

Availability of human support

There was no evidence for a moderator effect of the availability of human support, $Q_M(1) = 1.36$, $p = .243$ (see Table 2).

Stand-alone versus combined interventions

Also when comparing stand-alone mHealth interventions with combined interventions (e.g., mHealth intervention with face-to-face sessions), we found no evidence for a moderator effect, $Q_M(1) = 0.49$, $p = .484$ (see Table 2).

Control condition

There was evidence for a moderator effect of the type of comparator condition (i.e., passive vs. active comparators) $Q_M(1) = 11.90$, $p = .001$. While studies using passive comparators provided evidence for a small to medium favourable effect of interventions with a cognitive reappraisal component, there was no evidence for a favourable effect when active comparators were used (see Table 2).

Sensitivity analyses

Risk of bias

We re-estimated our main analysis on mental health outcomes excluding effect sizes rated to be at least "at some concern" for risk of bias in different domains (see Supplementary Material E for detailed results). Excluding all effect sizes being at least at "some concern" regarding the randomization process, the analysis

TABLE 3 Results of main analysis excluding effect sizes rated to be at least "at some concern" for specific risk of bias domains.

				Confidence interval		Prediction interval						
Risk of bias domain	*n*	*k*	*M*(SMD)	95%$_l$	95%$_u$	95%$_l$	95%$_u$	*p*	Q	df	*p*(Q)	I^2 (I^2_S, I^2_O)
Randomization	18	38	0.42	0.13	0.70	−0.70	1.53	.004	195.24	37	<.001	92.1 (89.8, 2.4)
Deviations from intended intervention	14	31	0.47	0.10	0.84	−0.74	1.68	.014	186.22	30	<.001	94.2 (84.4, 9.8)
Missing outcome data	19	39	0.42	0.11	0.74	−0.75	1.60	.009	245.10	38	<.001	93.1 (86.7, 6.5)
Outcome measurement	4	8	0.23	−0.12	0.59	–	–	.202	53.95	7	<.001	86.4 (86.4, 0)
Selective reporting	26	55	0.35	0.10	0.59	−0.53	1.00	.005	285.21	54	<.001	61.5 (55.9, 5.6)

The multilevel meta-analysis on mental health included anxiety symptoms, depressive symptoms, psychological distress, and well-being. Effect-sizes that were at least 'at some concern' for the respective bias domain were excluded from these analyses. For all outcomes, SMDs were calculated in a way that positive values indicate favorable effects of an intervention (i.e., reductions of mental symptoms or increases in well-being).
Df, degrees of freedom; I^2, heterogeneity index in percentage (range: 0%–100%); I^2_S, heterogeneity accounted for by between-study differences; I^2_O, heterogeneity accounted for by between-outcome differences; *k*, number of effect sizes; ML, multilevel meta-analysis, *n*, number of studies; Q, Cochran's *Q* statistic with *p* value; SMD, standardized mean difference; T, traditional meta-analysis; 95%$_l$, lower boundary of the 95% confidence/prediction interval; 95%$_u$, upper boundary of the 95% confidence/prediction interval.

yielded similar results, *M*(SMD) = 0.42, 95% CI [0.13, 0.70], $p < .001$. The same applied to analyses excluding effect sizes with at least some concerns in the domains of deviation from intended intervention, missing outcome data, and selective reporting. An analysis based on eight effect sizes at low risk for bias for outcome assessment provided no evidence for a favourable effect of mHealth interventions to enhance cognitive reappraisal, *M*(SMD) = 0.23, 95% CI [−0.12, 0.59], $p = .202$ (see **Table 3**).

Crossover RCT

We included a crossover RCT in our main analysis for which no ideal outcome data was available. However, in- and excluding this trial left our results largely unchanged, *M*(SMD) = 0.37, 95% CI [0.17, 0.56], $p < .001$, as evidenced by overlapping confidence intervals.

Condition labelling in primary studies

In a few cases, we included conditions as intervention group that were labelled as control conditions by primary study authors. Excluding these studies from our main analysis left our results unchanged, *M*(SMD) = 0.36, 95% CI [0.14, 0.58], $p = .002$.

Discussion

Main findings

With this systematic review and meta-analysis, we aimed to evaluate the efficacy of mHealth interventions comprising a component of cognitive reappraisal to enhance mental health. mHealth interventions are popular and an increasing number of programs aiming at reducing mental symptoms, preventing mental disorders, or promoting mental health become available for users. Nevertheless, the scientific evaluation of effective components is still in its beginnings. Therefore, we aimed to contribute to close this gap. We followed a mechanistic approach assuming that cognitive reappraisal, as an evidence-based therapeutic mechanism, might be a particular promising target of mHealth interventions.

During the process of study selection, it became evident that cognitive reappraisal is usually included in multi-component mHealth interventions. This means that reappraisal is one of many mHealth intervention components (e.g., behavioral exercises, psychoeducation). Only three of 30 identified studies exclusively focused on the promotion of reappraisal. Thus, this review and meta-analysis cannot provide evidence on the effects of single-component reappraisal mHealth interventions as initially intended. However, the present review provided an evidence synthesis on complex mHealth interventions including a reappraisal component.

In our main analysis, we included 62 effect sizes of 30 studies. We found evidence for a small to medium overall effect favoring mHealth interventions to enhance cognitive reappraisal over comparators across all mental health outcomes. Analyses considering single symptoms (i.e., anxiety, depression, psychological distress, and well-being) revealed a more differentiated picture. Small to medium effects in favor of mHealth interventions were found for depressive and anxiety symptoms, but there was no evidence for an effect on well-being and general psychological distress. Notably, we identified only two studies with outcomes of general psychological distress. This means that mHealth interventions might be beneficial for the reduction of depressive and anxiety symptoms only and do not succeed in the improvement of positive indicators of mental health, such as well-being.

The detected effect sizes are comparable to recent meta-analyses that found evidence for small to medium effects of mHealth interventions on depression (15, 17) and anxiety (16).

Even though the meta-analysis revealed a significant overall effect of complex mHealth interventions including a cognitive

reappraisal component on mental health, moderator analyses provided a more nuanced picture with favorable effects only being found when interventions were compared with passive but not active controls. In line with other meta-analyses of mHealth interventions (15, 16), this result suggests that the use of a digital device itself may provide psychological benefits. In this regard, the concept of a digital placebo effect has been proposed (98). Thus, the favorable overall intervention effect may reflect the rather unspecific effect of mHealth interventions.

We also investigated various other moderator variables (e.g., sociodemographic variables, availability of human support), but none of them reached significance. This indicates that complex mHealth interventions including a cognitive reappraisal component may be applicable to a broad range of users and a variety of intervention features.

Notably, we did not find evidence for mHealth interventions including cognitive reappraisal to be more effective when used to reduce mental symptoms compared to mHealth programs with a preventive focus. However, only a small amount of mHealth interventions with a focus on the prevention of mental disorders or mental health promotion were eligible for our systematic review. This is in line with findings from similar fields showing a strong focus on symptom reduction with limited evidence available for mental health promotion and prevention programs (99). Mostly, mHealth interventions targeted individuals with (at least) elevated symptoms of mental disorders, meaning that cognitive reappraisal was used for symptom reduction. The effectiveness of cognitive reappraisal components in mHealth interventions for more general population samples (i.e., prevention of mental disorder or mental health promotion) need further research. This ties in with an overall need for intensified research in the areas of mental health promotion and prevention of mental disorders (100).

In contrast to our expectations, we found no evidence for mHealth interventions with larger proportions of reappraisal promotion being associated with larger intervention effects. Consequently, we did not find evidence for a dose-response relationship between reappraisal proportion and intervention effects. First, this finding might be explained by the small proportion of cognitive reappraisal in most of the interventions (70% of the mHealth interventions included 20% or less reappraisal training). Second, we calculated the relation of the reappraisal-component relative to the total number of intervention components as the best available proxy for the intensity of reappraisal training. However, this might only insufficiently capture the intensity of reappraisal training. For example, a mHealth intervention containing one module on cognitive reappraisal next to a second module has a reappraisal-proportion of 50%. However, having less training modules might not automatically result in a more intense, and in turn, effective reappraisal training. Insufficient reporting of intervention details in the included primary studies prevented a more elaborated analysis.

The absence of a specific effect of cognitive reappraisal components might be interpreted with regard to the usefulness of cognitive reappraisal itself. Some authors have started to question the usefulness of cognitive reappraisal for regulating negative emotions in comparison to other emotion regulatory and cognitive mechanisms in the past (101, 102). For example, an individual participant data component network meta-analysis (cNMA) of internet-based CBT trials for depression resulted in no additive effect of cognitive restructuring components (43).

Limitations

Our review should be interpreted in the light of its limitations. First, the included studies are limited to interventions that are provided via mobile devices such as smartphones. Therefore, we can only draw conclusions on the efficacy of cognitive reappraisal in the specific setting of mHealth and one should not generalize the findings to, for example, any digital intervention or face-to-face settings. This implies that the small effects of complex mHealth interventions including a cognitive reappraisal component might be not because of the impact of reappraisal itself but arise from problems with the mobile delivery. A training of cognitive reappraisal might be better addressed in face-to-face settings than in mHealth interventions, for example because the commitment to the intervention and participants' willingness to engage in cognitive, sometimes challenging exercises might be more successful when supported and encouraged by a human being. Even if the use of mobile devices offers novel possibilities for encouragement and individualization, smartphones may sometimes fail to reach the individual and to consolidate reappraisal skills as a long-term engagement is often lacking. However, we cannot make a statement on the participants' compliance in the included mHealth interventions as data on compliance were provided rarely in the studies and were not eligible for analyses.

Moreover, during the selection process, it became apparent that many studies failed to report sufficient intervention details (i.e., training content). We can therefore not rule out that we missed mHealth interventions that have used cognitive reappraisal training but failed to report it explicitly. Vice versa, we might have overrated reappraisal in some studies because we classified studies as reappraisal intervention even if the reappraisal training was part of a complex intervention. Studies might also be heterogeneous regarding their reappraisal training components in terms of schedule and content. We might have also missed eligible studies that are ongoing or unpublished as we focused the literature search on electronic databases and did not check for grey literature or perform citation searching. Moreover, our literature search ended in March 2022 and was not updated, which might have resulted in some publications being not included in our analyses. Thus, future systematic reviews will have to update our findings based on a larger number of primary studies.

Further, there are limitations that originate from the included studies. The reviewed studies were mainly multi-component interventions, which might include active intervention components next to cognitive reappraisal. Thus, the reviewed studies did not allow us to draw conclusions about a reappraisal-

specific intervention effect. In addition, information whether participants acquired reappraisal skills are lacking. We could not analyze intervention effects on reappraisal because the included studies did not evaluate reappraisal-specific outcomes. Thus, we cannot answer the question of whether the included interventions in fact enhanced cognitive reappraisal.

The risk of bias assessment points towards a high risk of bias in most of the included studies. This is due to the fact that outcomes in the studies were assessed using self-report measures and participants were aware of group assignment.

Implications and future research

Even in the light of the limitations outlined above, our findings have important implications for future research on mHealth interventions. First, more research evaluating the effects of reappraisal training via mHealth interventions are required. These are mHealth interventions consisting of an isolated training of cognitive reappraisal as well as trials with cognitive reappraisal specific outcomes. Moreover, it is important to examine whether cognitive reappraisal or its combination with other intervention components are superior in the promotion of mental health. A promising approach are dismantling studies, in which full mHealth interventions are experimentally compared with a disentangled variation of the same intervention that omits cognitive reappraisal. These studies can lead to evidence on the specific effects of single intervention components, i.e., cognitive reappraisal, and advance the research on the mechanisms of therapeutic change in mHealth interventions. Second, research on the effects and usability of mHealth interventions enhancing cognitive reappraisal skills in the real-world (i.e., effectiveness studies) should be expanded. Third, meta-analyses on individual participant data could consider Level 1-moderators (e.g., age, symptom severity) with greater statistical power, which may allow for future tailoring of interventions.

Conclusion

Our findings provide insights into the implementation and use of cognitive reappraisal training modules in mHealth interventions. Training components related to the enhancement of cognitive reappraisal have gained scientific interest in the last years and are implemented in apps and other mHealth formats for various population groups. As expected, it turned out that most of the identified mHealth interventions were multi-component interventions, in which cognitive reappraisal was promoted alongside multiple other components. We found first evidence for a small to medium favorable effect of these complex mHealth interventions that included a component of cognitive reappraisal on mental health. The favorable effects were found only when mHealth interventions were compared with passive but not active controls. There was also no evidence for a dose–response effect of reappraisal. Thus, our findings suggest that favorable effects may arise mainly from unspecific beneficial effect of mHealth interventions. Consequently, high-quality dismantling studies examining the effects of single intervention components are warranted to corroborate and further evaluate active therapeutic mechanisms in mHealth interventions such as cognitive reappraisal.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

KM: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review and editing. SS: Conceptualization, Data curation, Methodology, Supervision, Writing – original draft, Writing – review and editing. AK: Conceptualization, Formal analysis, Methodology, Writing – review and editing. L-SP: Data curation, Writing – review and editing. OT: Conceptualization, Project administration, Supervision, Writing – review and editing. TK: Conceptualization, Project administration, Supervision, Writing – review and editing.

Funding

This project has received funding from: the State of Rhineland-Palatinate, Germany (DRZ EMA/EMI program, Leibniz Institute for Resilience Research), the European Union's Horizon 2020 research and innovation program under Grant Agreement No. 777084 (DynaMORE project) and the Deutsche Forschungsgemeinschaft (DFG Grant CRC 1193, subprojects C04 & Z03). The publication of this article was funded by the Open Access Fund of the Leibniz Association.

Acknowledgments

We wish to thank all research assistants for their contribution in the literature screening, data extraction, and quality assessment.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of

their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2023.1253390/full#supplementary-material

SUPPLEMENTARY MATERIAL A
Full search strategy.

SUPPLEMENTARY MATERIAL B
Coding.

SUPPLEMENTARY MATERIAL C
Study characteristics.

SUPPLEMENTARY MATERIAL D
Proportion of cognitive reappraisal.

SUPPLEMENTARY MATERIAL E
Study quality assessment.

SUPPLEMENTARY MATERIAL F
Moderator analysis.

References

1. World Health Organization. *The WHO special initiative for mental health (2019–2023): universal health coverage for mental health; technical documents*. (2019).

2. Fairburn CG, Patel V. The impact of digital technology on psychological treatments and their dissemination. *Behav Res Ther*. (2017) 88:19–25. doi: 10.1016/j.brat.2016.08.012

3. Petrie K, Proudfoot J, Clarke J, Birch M-R, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. *J Med Internet Res*. (2013) 15(11):e247. doi: 10.2196/jmir.2791

4. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. *Br J Health Psychol*. (2010) 15(Pt 1):1–39. doi: 10.1348/135910709X466063

5. Donker T, Blankers M, Hedman E, Ljótsson B, Petrie K, Christensen H. Economic evaluations of internet interventions for mental health: a systematic review. *Psychol Med*. (2015) 45(16):3357–76. doi: 10.1017/S0033291715001427

6. Paganini S, Teigelkötter W, Buntrock C, Baumeister H. Economic evaluations of internet- and mobile-based interventions for the treatment and prevention of depression: a systematic review. *J Affect Disord*. (2018) 225:733–55. doi: 10.1016/j.jad.2017.07.018

7. Rathbone AL, Clarry L, Prescott J. Assessing the efficacy of mobile health apps using the basic principles of cognitive behavioral therapy: systematic review. *J Med Internet Res*. (2017) 19(11):e399. doi: 10.2196/jmir.8598

8. Rathbone AL, Prescott J. The use of mobile apps and SMS messaging as physical and mental health interventions: systematic review. *J Med Internet Res*. (2017) 19(8):e295. doi: 10.2196/jmir.7740

9. Wang K, Varma DS, Prosperi M. A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders. *J Psychiatr Res*. (2018) 107:73–8. doi: 10.1016/j.jpsychires.2018.10.006

10. Wickersham A, Petrides PM, Williamson V, Leightley D. Efficacy of mobile application interventions for the treatment of post-traumatic stress disorder: a systematic review. *Digit Health*. (2019) 5:1–11. doi: 10.1177/2055207619842986

11. Serrano-Ripoll MJ, Zamanillo-Campos R, Fiol-DeRoque MA, Castro A, Ricci-Cabello I. Impact of smartphone app-based psychological interventions for reducing depressive symptoms in people with depression: systematic literature review and meta-analysis of randomized controlled trials. *JMIR Mhealth Uhealth*. (2022) 10(1):e29621. doi: 10.2196/29621

12. 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. doi: 10.2196/17458

13. Purgato M, Uphoff E, Singh R, Thapa Pachya A, Abdulmalik J, van Ginneken N. Promotion, prevention and treatment interventions for mental health in low- and middle-income countries through a task-shifting approach. *Epidemiol Psychiatr Sci*. (2020) 29:e150. doi: 10.1017/S204579602000061X

14. Uphoff E, Robertson L, Cabieses B, Villalón FJ, Purgato M, Churchill R, et al. An overview of systematic reviews on mental health promotion, prevention, and treatment of common mental disorders for refugees, asylum seekers, and internally displaced persons. *Cochrane Database Syst Rev*. (2020) 9(9):CD013458. doi: 10.1002/14651858.CD013458.pub2

15. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry*. (2017) 16(3):287–98. doi: 10.1002/wps.20472

16. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. *J Affect Disord*. (2017) 218:15–22. doi: 10.1016/j.jad.2017.04.046

17. Senanayake B, Wickramasinghe SI, Chatfield MD, Hansen J, Edirippulige S, Smith AC. Effectiveness of text messaging interventions for the management of depression: a systematic review and meta-analysis. *J Telemed Telecare*. (2019) 25(9):513–23. doi: 10.1177/1357633X19875852

18. Versluis A, Verkuil B, Spinhoven P, van der Ploeg MM, Brosschot JF. Changing mental health and positive psychological well-being using ecological momentary interventions: a systematic review and meta-analysis. *J Med Internet Res*. (2016) 18(6):e152. doi: 10.2196/jmir.5642

19. Deady M, Choi I, Calvo RA, Glozier N, Christensen H, Harvey SB. Ehealth interventions for the prevention of depression and anxiety in the general population: a systematic review and meta-analysis. *BMC Psychiatry*. (2017) 17(1):310. doi: 10.1186/s12888-017-1473-1

20. Eisenstadt M, Liverpool S, Infanti E, Ciuvat RM, Carlsson C. Mobile apps that promote emotion regulation, positive mental health, and well-being in the general population: systematic review and meta-analysis. *JMIR Ment Health*. (2021) 8(11):e31170. doi: 10.2196/31170

21. Seegan PL, Miller MJ, Heliste JL, Fathi L, McGuire JF. Efficacy of stand-alone digital mental health applications for anxiety and depression: a meta-analysis of randomized controlled trials. *J Psychiatr Res*. (2023) 164:171–83. doi: 10.1016/j.jpsychires.2023.06.019

22. Coulon SM, Monroe CM, West DS. A systematic, multi-domain review of mobile smartphone apps for evidence-based stress management. *Am J Prev Med*. (2016) 51(1):95–105. doi: 10.1016/j.amepre.2016.01.026

23. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. *JMIR Ment Health*. (2016) 3(1):e7. doi: 10.2196/mental.4984

24. Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: where is the evidence? *PLoS Med*. (2013) 10(2):e1001382. doi: 10.1371/journal.pmed.1001382

25. Larsen ME, Huckvale K, Nicholas J, Torous J, Birrell L, Li E, et al. Using science to sell apps: evaluation of mental health app store quality claims. *NPJ Digit Med*. (2019) 2(1):18. doi: 10.1038/s41746-019-0093-1

26. Wasil AR, Venturo-Conerly KE, Shingleton RM, Weisz JR. A review of popular smartphone apps for depression and anxiety: assessing the inclusion of evidence-based content. *Behav Res Ther*. (2019) 123:103498. doi: 10.1016/j.brat.2019.103498

27. Martinengo L, Stona A-C, Griva K, Dazzan P, Pariante CM, Wangenheim Fv, et al. Self-guided cognitive behavioral therapy apps for depression: systematic assessment of features, functionality, and congruence with evidence. *J Med Internet Res*. (2021) 23(7):e27619. doi: 10.2196/27619

28. Gross JJ. Emotion regulation: affective, cognitive, and social consequences. *Psychophysiology*. (2002) 39(3):281–91. doi: 10.1017/S0048577201393198

29. McRae K, Ciesielski B, Gross JJ. Unpacking cognitive reappraisal: goals, tactics, and outcomes. *Emotion*. (2012) 12(2):250–5. doi: 10.1037/a0026351

30. McRae K, Gross JJ. Emotion regulation. *Emotion*. (2020) 20(1):1–9. doi: 10.1037/emo0000703

31. Beck AT, Dozois DJA. Cognitive therapy: current status and future directions. *Annu Rev Med*. (2011) 62:397–409. doi: 10.1146/annurev-med-052209-100032

32. Clark DA. Cognitive restructuring. In: Hofmann SG, editors. *The wiley handbook of cognitive behavioral therapy*. Chichester, UK: John Wiley & Sons, Ltd (2013). p. 1–22.

33. Aldao A, Nolen-Hoeksema S. Specificity of cognitive emotion regulation strategies: a transdiagnostic examination. *Behav Res Ther*. (2010) 48(10):974–83. doi: 10.1016/j.brat.2010.06.002

34. Garnefski N, Kraaij V. Relationships between cognitive emotion regulation strategies and depressive symptoms: a comparative study of five specific samples. *Pers Individ Dif*. (2006) 40(8):1659–69. doi: 10.1016/j.paid.2005.12.009

35. Hu T, Zhang D, Wang J, Mistry R, Ran G, Wang X. Relation between emotion regulation and mental health: a meta-analysis review. *Psychol Rep*. (2014) 114 (2):341–62. doi: 10.2466/03.20.PR0.114k22w4

36. LeMoult J, Gotlib IH. Depression: a cognitive perspective. *Clin Psychol Rev*. (2019) 69:51–66. doi: 10.1016/j.cpr.2018.06.008

37. Haga SM, Kraft P, Corby E-K. Emotion regulation: antecedents and well-being outcomes of cognitive reappraisal and expressive suppression in cross-cultural samples. *Journal of Happiness Studies: An Interdisciplinary Forum on Subjective Well-Being*. (2009) 10(3):271–91. doi: 10.1007/s10902-007-9080-3

38. Riepenhausen A, Wackerhagen C, Reppmann ZC, Deter H-C, Kalisch R, Veer IM, et al. Positive cognitive reappraisal in stress resilience, mental health, and well-being: a comprehensive systematic review. *Emot Rev*. (2022) 14(4):310–31. doi: 10.1177/17540739221114642

39. Lorenzo-Luaces L, German RE, DeRubeis RJ. It's complicated: the relation between cognitive change procedures, cognitive change, and symptom change in cognitive therapy for depression. *Clin Psychol Rev*. (2015) 41:3–15. doi: 10.1016/j.cpr.2014.12.003

40. Hofmann SG, Meuret AE, Rosenfield D, Suvak MK, Barlow DH, Gorman JM, et al. Preliminary evidence for cognitive mediation during cognitive-behavioral therapy of panic disorder. *J Consult Clin Psychol*. (2007) 75(3):374–9. doi: 10.1037/0022-006X.75.3.374

41. Kleim B, Grey N, Wild J, Nussbeck FW, Stott R, Hackmann A, et al. Cognitive change predicts symptom reduction with cognitive therapy for posttraumatic stress disorder. *J Consult Clin Psychol*. (2013) 81(3):383–93. doi: 10.1037/a0031290

42. López-López JA, Davies SR, Caldwell DM, Churchill R, Peters TJ, Tallon D, et al. The process and delivery of CBT for depression in adults: a systematic review and network meta-analysis. *Psychol Med*. (2019) 49(12):1937–47. doi: 10.1017/S003329171900120X

43. Furukawa TA, Suganuma A, Ostinelli EG, Andersson G, Beevers CG, Shumake J, et al. Dismantling, optimising, and personalising internet cognitive behavioural therapy for depression: a systematic review and component network meta-analysis using individual participant data. *Lancet Psychiatry*. (2021) 8(6):500–11. doi: 10.1016/S2215-0366(21)00077-8

44. Cuijpers P, Cristea IA, Karyotaki E, Reijnders M, Hollon SD. Component studies of psychological treatments of adult depression: a systematic review and meta-analysis. *Psychother Res*. (2019) 29(1):15–29. doi: 10.1080/10503307.2017.1395922

45. Pompoli A, Furukawa TA, Efthimiou O, Imai H, Tajika A, Salanti G. Dismantling cognitive-behaviour therapy for panic disorder: a systematic review and component network meta-analysis. *Psychol Med*. (2018) 48(12):1945–53. doi: 10.1017/S0033291717003919

46. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Br Med J*. (2021) 372:n71. doi: 10.1136/bmj.n71

47. Veritas Health Innovation. *Covidence—better systematic review management*. (2019).

48. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev*. (2016) 5(1):210. doi: 10.1186/s13643-016-0384-4

49. Kalisch R, Müller MB, Tüscher O. A conceptual framework for the neurobiological study of resilience. *Behav Brain Sci*. (2015) 38:e92. doi: 10.1017/S0140525X1400082X

50. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. (2001) 16(9):606–13. doi: 10.1046/j.1525-1497.2001.016009606.x

51. Garnefski N, Kraaij V. The cognitive emotion regulation questionnaire. *Eur J Psychol Assess*. (2007) 23(3):141–9. doi: 10.1027/1015-5759.23.3.141

52. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. Rob 2: a revised tool for assessing risk of bias in randomised trials. *Br Med J*. (2019) 366:l4898. doi: 10.1136/bmj.l4898

53. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *J Clin Epidemiol*. (2008) 61(10):991–6. doi: 10.1016/j.jclinepi.2007.11.010

54. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics*. (1994) 50(4):1088–101. doi: 10.2307/2533446

55. Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. *Cochrane handbook for systematic reviews of interventions version 6.3 (updated February 2022)*. Chichester, UK: Cochrane (2022). Available at: www.training.cochrane.org/handbook

56. R Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing (2022). Available from: https://www.R-project.org/

57. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw*. (2010) 36(3):1–48. doi: 10.18637/jss.v036.i03

58. Assink M, Wibbelink CJM. Fitting three-level meta-analytic models in R: a step-by-step tutorial. *TQMP*. (2016) 12(3):154–74. doi: 10.20982/tqmp.12.3.p154

59. Beck AT, Steer RA, Brown G. Beck depression inventory–II (BDI-II) [Database record]. *APA PsycTests*. (1996). https://doi.org/10.1037/t00742-000

60. Sedgwick P. Meta-analyses: what is heterogeneity? *Br Med J*. (2015) 350:h1435. doi: 10.1136/bmj.h1435

61. Ahorsu DK, Lin C-Y, Imani V, Carlbring P, Nygårdh A, Broström A, et al. Testing an app-based intervention to improve insomnia in patients with epilepsy: a randomized controlled trial. *Epilepsy Behav*. (2020) 112:107371. doi: 10.1016/j.yebeh.2020.107371

62. Al-Refae M, Al-Refae A, Munroe M, Sardella NA, Ferrari M. A self-compassion and mindfulness-based cognitive Mobile intervention (serene) for depression, anxiety, and stress: promoting adaptive emotional regulation and wisdom. *Front Psychol*. (2021) 12:648087. doi: 10.3389/fpsyg.2021.648087

63. Bakker D, Kazantzis N, Rickwood D, Rickard N. A randomized controlled trial of three smartphone apps for enhancing public mental health. *Behav Res Ther*. (2018) 109:75–83. doi: 10.1016/j.brat.2018.08.003

64. Bruehlman-Senecal E, Hook CJ, Pfeifer JH, FitzGerald C, Davis B, Delucchi KL, et al. Smartphone app to address loneliness among college students: pilot randomized controlled trial. *JMIR Ment Health*. (2020) 7(10):e21496. doi: 10.2196/21496

65. Bruhns A, Lüdtke T, Moritz S, Bücker L. A Mobile-based intervention to increase self-esteem in students with depressive symptoms: randomized controlled trial. *JMIR Mhealth Uhealth*. (2021) 9(7):e26498. doi: 10.2196/26498

66. Dagöö J, Asplund RP, Bsenko HA, Hjerling S, Holmberg A, Westh S, et al. Cognitive behavior therapy versus interpersonal psychotherapy for social anxiety disorder delivered via smartphone and computer: a randomized controlled trial. *J Anxiety Disord*. (2014) 28(4):410–7. doi: 10.1016/j.janxdis.2014.02.003

67. Dahne J, Collado A, Lejuez CW, Risco CM, Diaz VA, Coles L, et al. Pilot randomized controlled trial of a spanish-language behavioral activation mobile app (¡aptívate!) for the treatment of depressive symptoms among United States latinx adults with limited english proficiency. *J Affect Disord*. (2019) 250:210–7. doi: 10.1016/j.jad.2019.03.009

68. Dahne J, Lejuez CW, Diaz VA, Player MS, Kustanowitz J, Felton JW, et al. Pilot randomized trial of a self-help behavioral activation mobile app for utilization in primary care. *Behav Ther*. (2019) 50(4):817–27. doi: 10.1016/j.beth.2018.12.003

69. Depp CA, Perivoliotis D, Holden J, Dorr J, Granholm EL. Single-session mobile-augmented intervention in serious mental illness: a three-arm randomized controlled trial. *Schizophr Bull*. (2019) 45(4):752–62. doi: 10.1093/schbul/sby135

70. Greer JA, Jacobs J, Pensak N, MacDonald JJ, Fuh C-X, Perez GK, et al. Randomized trial of a tailored cognitive-behavioral therapy mobile application for anxiety in patients with incurable cancer. *Oncologist*. (2019) 24(8):1111–20. doi: 10.1634/theoncologist.2018-0536

71. Ham K, Chin S, Suh YJ, Rhee M, Yu E-S, Lee HJ, et al. Preliminary results from a randomized controlled study for an app-based cognitive behavioral therapy program for depression and anxiety in cancer patients. *Front Psychol*. (2019) 10:1592. doi: 10.3389/fpsyg.2019.01592

72. Hunt M, Miguez S, Dukas B, Onwude O, White S. Efficacy of zemedy, a mobile digital therapeutic for the self-management of irritable bowel syndrome: crossover randomized controlled trial. *JMIR Mhealth Uhealth*. (2021) 9(5):e26152. doi: 10.2196/26152

73. Hur J-W, Kim B, Park D, Choi S-W. A scenario-based cognitive behavioral therapy mobile app to reduce dysfunctional beliefs in individuals with depression: a randomized controlled trial. *Telemedicine and e-Health*. (2018) 24(9):710–6. doi: 10.1089/tmj.2017.0214

74. Imamura K, Tran TTT, Nguyen HT, Sasaki N, Kuribayashi K, Sakuraya A, et al. Effect of smartphone-based stress management programs on depression and anxiety of hospital nurses in Vietnam: a three-arm randomized controlled trial. *Sci Rep*. (2021) 11(1):11353. doi: 10.1038/s41598-021-90320-5

75. Sasaki N, Imamura K, Tran TTT, Nguyen HT, Kuribayashi K, Sakuraya A, et al. Effects of smartphone-based stress management on improving work engagement among nurses in Vietnam: secondary analysis of a three-arm randomized controlled trial. *J Med Internet Res*. (2021) 23(2):e20445. doi: 10.2196/20445

76. Jannati N, Mazhari S, Ahmadian L, Mirzaee M. Effectiveness of an app-based cognitive behavioral therapy program for postpartum depression in primary care: a randomized controlled trial. *Int J Med Inform*. (2020) 141:104145. doi: 10.1016/j.ijmedinf.2020.104145

77. Jarvis MA, Padmanabhanunni A, Chipps J. An evaluation of a low-intensity cognitive behavioral therapy mHealth-supported intervention to reduce loneliness

in older people. *Int J Environ Res Public Health*. (2019) 16(7):1305. doi: 10.3390/ijerph16071305

78. Liu H, Peng H, Song X, Xu C, Zhang M. Using AI chatbots to provide self-help depression interventions for university students: a randomized trial of effectiveness. *Internet Interv*. (2022) 27:100495. doi: 10.1016/j.invent.2022.100495

79. Lüdtke T, Westermann S, Pult LK, Schneider BC, Pfuhl G, Moritz S. Evaluation of a brief unguided psychological online intervention for depression: a controlled trial including exploratory moderator analyses. *Internet Interv*. (2018) 13:73–81. doi: 10.1016/j.invent.2018.06.004

80. Lukas CA, Eskofier B, Berking M. A gamified smartphone-based intervention for depression: randomized controlled pilot trial. *JMIR Ment Health*. (2021) 8(7):e16643. doi: 10.2196/16643

81. Mantani A, Kato T, Furukawa TA, Horikoshi M, Imai H, Hiroe T, et al. Smartphone cognitive behavioral therapy as an adjunct to pharmacotherapy for refractory depression: randomized controlled trial. *J Med Internet Res*. (2017) 19 (11):e373. doi: 10.2196/jmir.8602

82. Mantani A, Kato T, Furukawa TA, Horikoshi M, Imai H, Hiroe T, et al. Correction: smartphone cognitive behavioral therapy as an adjunct to pharmacotherapy for refractory depression: randomized controlled trial. *J Med Internet Res*. (2018) 20(8):e11702. doi: 10.2196/11702

83. Furukawa TA, Horikoshi M, Fujita H, Tsujino N, Jinnin R, Kako Y, et al. Cognitive and behavioral skills exercises completed by patients with major depression during smartphone cognitive behavioral therapy: secondary analysis of a randomized controlled trial. *JMIR Ment Health*. (2018) 5(1):e4. doi: 10.2196/mental.9092

84. Imai H, Yamada M, Inagaki M, Watanabe N, Chino B, Mantani A, et al. Behavioral activation contributed to the total reduction of depression symptoms in the smartphone-based cognitive behavioral therapy: a secondary analysis of a randomized, controlled trial. *Innov Clin Neurosci*. (2020) 17(7-9):21–5.

85. McCloud T, Jones R, Lewis G, Bell V, Tsakanikos E. Effectiveness of a mobile app intervention for anxiety and depression symptoms in university students: randomized controlled trial. *JMIR Mhealth Uhealth*. (2020) 8(7):e15418. doi: 10.2196/15418

86. Meyer B, Weiss M, Holtkamp M, Arnold S, Brückner K, Schröder J, et al. Effects of an epilepsy-specific internet intervention (emyna) on depression: results of the ENCODE randomized controlled trial. *Epilepsia*. (2019) 60(4):656–68. doi: 10.1111/epi.14673

87. Moberg C, Niles A, Beermann D. Guided self-help works: randomized waitlist controlled trial of Pacifica, a mobile app integrating cognitive behavioral therapy and mindfulness for stress, anxiety, and depression. *J Med Internet Res*. (2019) 21 (6):e12556. doi: 10.2196/12556

88. Newman MG, Jacobson NC, Rackoff GN, Bell MJ, Taylor CB. A randomized controlled trial of a smartphone-based application for the treatment of anxiety. *Psychother Res*. (2021) 31(4):443–54. doi: 10.1080/10503307.2020.1790688

89. Oh J, Jang S, Kim H, Kim J-J. Efficacy of mobile app-based interactive cognitive behavioral therapy using a chatbot for panic disorder. *Int J Med Inform*. (2020) 140:104171. doi: 10.1016/j.ijmedinf.2020.104171

90. Roepke AM, Jaffee SR, Riffle OM, McGonigal J, Broome R, Maxwell B. Randomized controlled trial of SuperBetter, a smartphone-based/internet-based self-help tool to reduce depressive symptoms. *Games Health J*. (2015) 4(3):235–46. doi: 10.1089/g4h.2014.0046

91. Röhr S, Jung FU, Pabst A, Grochtdreis T, Dams J, Nagl M, et al. A self-help app for Syrian refugees with posttraumatic stress (sanadak): randomized controlled trial. *JMIR Mhealth Uhealth*. (2021) 9(1):e24807. doi: 10.2196/24807

92. Stiles-Shields C, Montague E, Kwasny MJ, Mohr DC. Behavioral and cognitive intervention strategies delivered via coached apps for depression: pilot trial. *Psychol Serv*. (2019) 16(2):233–8. doi: 10.1037/ser0000261

93. Stolz T, Schulz A, Krieger T, Vincent A, Urech A, Moser C, et al. A mobile app for social anxiety disorder: a three-arm randomized controlled trial comparing mobile and PC-based guided self-help interventions. *J Consult Clin Psychol*. (2018) 86 (6):493–504. doi: 10.1037/ccp0000301

94. Zeng Y, Guo Y, Ho RTH, Zhu M, Zeng C, Monroe-Wise A, et al. Positive coping as a mediator of mobile health intervention effects on quality of life among people living with HIV: secondary analysis of the randomized controlled trial Run4Love. *J Med Internet Res*. (2022) 24(2):e25948. doi: 10.2196/25948

95. Zhu M, Cai W, Li L, Guo Y, Monroe-Wise A, Li Y, et al. Mediators of intervention effects on depressive symptoms among people living with HIV: secondary analysis of a mobile health randomized controlled trial using latent growth curve modeling. *JMIR Mhealth Uhealth*. (2019) 7(11):e15489. doi: 10.2196/15489

96. Watts S, Mackenzie A, Thomas C, Griskaitis A, Mewton L, Williams A, et al. CBT for depression: a pilot RCT comparing mobile phone vs. Computer. *BMC Psychiatry*. (2013) 13(1):49. doi: 10.1186/1471-244X-13-49

97. Christoforou M, Sáez Fonseca JA, Tsakanikos E. Two novel cognitive behavioral therapy–based Mobile apps for agoraphobia: randomized controlled trial. *J Med Internet Res*. (2017) 19(11):e398. doi: 10.2196/jmir.7747

98. Stalujanis E, Neufeld J, Glaus Stalder M, Belardi A, Tegethoff M, Meinlschmidt G. Induction of efficacy expectancies in an ambulatory smartphone-based digital placebo mental health intervention: randomized controlled trial. *JMIR Mhealth Uhealth*. (2021) 9(2):e20329. doi: 10.2196/20329

99. Schäfer SK, Kunzler AM, Lindner S, Broll J, Stoll M, Stoffers-Winterling J, et al. Transdiagnostic psychosocial interventions to promote mental health in forcibly displaced persons: a systematic review and meta-analysis. *Eur J Psychotraumatol*. (2023) 14(2):2196762. doi: 10.1080/20008066.2023.2196762

100. Singh V, Kumar A, Gupta S. Mental health prevention and promotion-A narrative review. *Front Psychiatry*. (2022) 13:898009. doi: 10.3389/fpsyt.2022.898009

101. Ehret AM, Joormann J, Berking M. Self-compassion is more effective than acceptance and reappraisal in decreasing depressed mood in currently and formerly depressed individuals. *J Affect Disord*. (2018) 226:220–6. doi: 10.1016/j.jad.2017.10.006

102. Krafft J, Haeger JA, Levin ME. Comparing cognitive fusion and cognitive reappraisal as predictors of college student mental health. *Cogn Behav Ther*. (2019) 48(3):241–52. doi: 10.1080/16506073.2018.1513556

frontiers | Frontiers in Digital Health

TYPE Systematic Review
PUBLISHED 07 November 2023
DOI 10.3389/fdgth.2023.1263488

OPEN ACCESS

EDITED BY
Edith Talina Luhanga,
Carnegie Mellon University Africa, Rwanda

REVIEWED BY
Ejemai Eboreime,
University of Alberta, Canada
Elissa Z. Faro,
University of Iowa Health Care, United States

*CORRESPONDENCE
Ogochukwu Udenigwe
ouden024@uottawa.ca

RECEIVED 19 July 2023
ACCEPTED 09 October 2023
PUBLISHED 07 November 2023

CITATION
Udenigwe O, Omonaiye O and Yaya S (2023) Gender transformative approaches in mHealth for maternal healthcare in sub-Saharan Africa: a systematic review.
Front. Digit. Health 5:1263488.
doi: 10.3389/fdgth.2023.1263488

Gender transformative approaches in mHealth for maternal healthcare in sub-Saharan Africa: a systematic review

Ogochukwu Udenigwe[1]*, Olumuyiwa Omonaiye[2,3] and Sanni Yaya[1,4]

[1]School of International Development and Global Studies, University of Ottawa, Ottawa, ON, Canada, [2]Centre for Quality and Patient Safety Research, Institute for Health Transformation, Deakin University, Burwood, VIC, Australia, [3]Centre for Quality and Patient Safety Research—Eastern Health Partnership, Eastern Health, Box Hill, VIC, Australia, [4]The George Institute for Global Health, Imperial College London, London, United Kingdom

Background: This review focuses on studies about digital health interventions in sub-Saharan Africa. Digital health interventions in sub-Saharan Africa are increasingly adopting gender-transformative approaches to address factors that derail women's access to maternal healthcare services. However, there remains a paucity of synthesized evidence on gender-transformative digital health programs for maternal healthcare and the corresponding research, program and policy implications. Therefore, this systematic review aims to synthesize evidence of approaches to transformative gender integration in digital health programs (specifically mHealth) for maternal health in sub-Saharan Africa.

Method: The following key terms "mobile health", "gender", "maternal health", "sub-Saharan Africa" were used to conduct electronic searches in the following databases: PsycInfo, EMBASE, Medline (OVID), CINAHL, and Global Health databases. The method and results are reported as consistent with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). Data synthesis followed a convergent approach for mixed-method systematic review recommended by the JBI (Joanna Briggs Institute).

Results: Of the 394 studies retrieved from the databases, 11 were included in the review. Out of these, six studies were qualitative in nature, three were randomized control trials, and two were mixed-method studies. Findings show that gender transformative programs addressed one or more of the following categories: (1) gender norms/roles/relations, (2) women's specific needs, (3) causes of gender-based health inequities, (4) ways to transform harmful gender norms, (5) promoting gender equality, (6) progressive changes in power relationships between women and men. The most common mHealth delivery system was text messages via short message service on mobile phones. The majority of mHealth programs for maternal healthcare were focused on reducing unintended pregnancies through the promotion of contraceptive use. The most employed gender transformative approach was a focus on women's specific needs.

Conclusion: Findings from gender transformative mHealth programs indicate positive results overall. Those reporting negative results indicated the need for a more explicit focus on gender in mHealth programs. Highlighting gender transformative approaches adds to discussions on how best to promote mHealth for maternal health through a gender transformative lens and provides evidence relevant to policy and research.

Systematic review registration: PROSPERO CRD42023346631.

KEYWORDS

digital health, mHealth, maternal health, gender integration, gender transformative, sub-Saharan Africa

Introduction

At the 71st World Health Assembly in 2018, resolutions on digital health underscored the need for digital health to not only enhance existing health service delivery models but to also contribute towards achieving health equity including gender equality (1). Precedents on gender integration in women's health were set in the 1990s and addressed the broad category of health issues that are unique to women such as maternal health and health issues that may manifest differently in women than men such as heart diseases. Significant global gatherings such as the International Conference on Population and Development and the World Conference on Women in Beijing recognized gender inequality as a critical factor influencing health, particularly for women who face disproportionate disadvantages in health outcomes (2, 3). Women face unequal access to healthcare resources and bear the burden of gender stereotypes that are perpetuated through health policies and programs, this had resulted in inadequate or inappropriate services for women (2–4). Targeting gender attitudes and norms is an important part of the broader strategy to achieve the sustainable development goals, but explicit attention to gender is often missing in health programming.

Aligned with sustainable development goal (SDG) 5 which aims to *Enhance the use of enabling technology, in particular information and communications technology, to promote the empowerment of women*, digital health is showing the potential to drive gender equality by reducing unequal access to and use of healthcare services (5). The field of digital health focuses on the use of information and communication technologies systems or channels, software, and data to improve health and wellness (6). While the healthcare transformations brought about by digital health are fundamentally technological, digital health also transforms the social, economic, and political context within which they occur (7, 8). Therefore, digital health programs must foreground the voices and realities of users, especially marginalized populations in their program design and delivery. Digital health has been incorporated across the pregnancy care continuum in efforts to address social determinants of health, improve the quality of care and ensure positive maternal health outcomes (9–11). However, to achieve meaningful impact, gender and digital inclusion must remain a priority in developing, implementing and evaluating digital health programs. Women, who are often the target groups for maternal health programs, are not homogenous. Their social, cultural and structural context will differ based on the relationships that govern their everyday lives (7). Barriers brought about by gender dynamics have demand and supply-side implications for women's participation in digital health for maternal healthcare. On the demand side, for example, the gender divide in mobile phone ownership or unfavourable community and cultural preferences, attitude, and norms around women's participation in digital health can impede women's participation in digital health and even exacerbate existing inequalities in access to digital health services (12, 13). On the supply side, breaches of confidentiality of women's health data on digital health programs or culturally insensitive digital health programs are detrimental to women's participation in digital health (14, 15). These implications demonstrate the importance of sex and gender considerations in digital health programming especially for maternal health. Considering gender in and of itself is not sufficient because some pathways to gender consideration in health can exploit or accommodate harmful gender norms rather than transform them. **Figure 1** depicts different types of gender inclusion strategies and serves as a guide for discussions on the implications of these strategies on gender equity outcomes.

Gender-blind health programs have no gender considerations, they ignore gender norms and relations and consequently risk reinforcing gender-based discrimination, biases, and stereotypes. Gender-aware programs, on the other hand, acknowledge gender norms and adopt an approach along a continuum as follows: First, gender exploitative approaches intentionally or unintentionally take advantage of gender inequalities to advance program outcomes thereby exacerbating gender inequalities; Second, gender-accommodating programs acknowledge but circumvent gender inequalities to achieve program outcomes; Third, gender transformative approaches in health programming aim to change gender power dynamics and/or reduce gender gaps in access to resources to achieve equitable gender norms and dynamics (16–18). Such gender-informed implications are integral to understanding how to approach health intervention efforts for maternal health.

The need for gender transformative approaches in health programming is increasingly highlighted in global health research, especially as it pertains to maternal healthcare (4, 19). This need recognizes gender as a key determinant of maternal health and acknowledges that women and girls are disproportionately disadvantaged in health outcomes. Gender

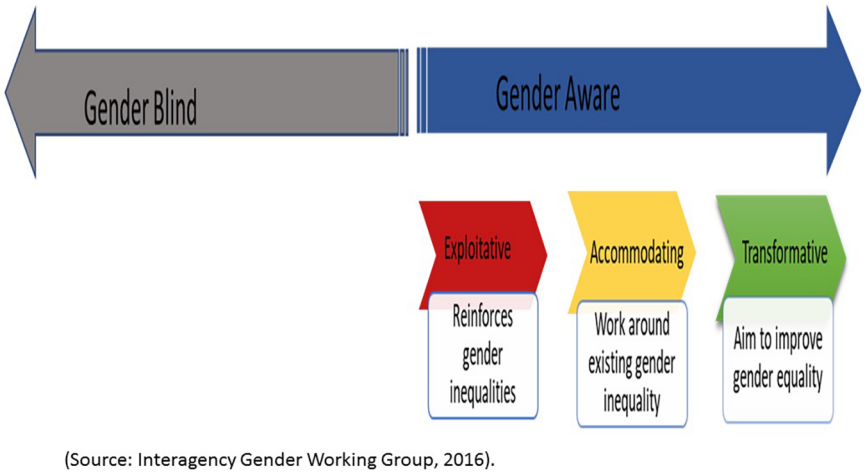


FIGURE 1
Gender integration continuum.

transformative approaches in non-digital health programs have been shown to be effective in improving maternal health. For example, integrating gender-specific differences in health promotion measures across sub-Saharan Africa led to shifts in gendered attitudes and behaviours which in turn improved maternal health outcomes (4). In Rwanda, an intervention that tackled inequitable power dynamics within heterosexual relationships saw increased modern contraceptive use among women and increased men's engagement in pregnancy care (20). In Uganda, a gender transformative approach to prevent violence against women and prevent HIV risk saw shifts in deeply entrenched attitudes on inter-partner violence among men and women (21).

Digital health interventions in sub-Saharan Africa are also adopting gender transformative approaches to address factors that derail women's access to maternal healthcare services. For instance, in Kenya, a mHealth program identified a digital divide within their target population and implemented strategies to improve women's digital access to quality maternal health services (22). Their strategies included the provision of inexpensive mobile phones, digital literacy for women, and working with men and the community to address negative social norms that restrict women's access to digital technology. In Nigeria, studies showed that addressing women's specific needs such as increased access to required technology improves women's participation in digital health programs and maternal health outcomes (23, 24).

Identifying such gender transformative approaches will inform policy and enhance best practices for gender integration in digital health. However, there remains a paucity of synthesized evidence on gender transformative considerations being made in digital health programs for maternal healthcare in sub-Saharan African contexts and the corresponding research, program and policy implications. Therefore, this systematic review aims to synthesize evidence of approaches to transformative gender integration that address gender inequality in mHealth for maternal health in sub-Saharan Africa. Addressing gender inequality in health programming is conceptualized as a gender-transformative approach. This review adopts the definition offered by the World Health Organization (WHO) and interprets a gender-transformative digital health program as one that "addresses the causes of gender-based health inequities through approaches that challenge and redress harmful and unequal gender norms, roles, and power relations that privilege men over women (18) (p. 136)".

This systematic review will address the following questions:

1. How are mHealth interventions for maternal health in sub-Saharan Africa adopting gender transformative approaches?
2. To what extent are gender-transformative interventions positively impacting maternal health outcomes?

Method

This review has been registered on PROSPERO with the registration ID CRD42023346631. A review protocol for this review was not prepared. The reporting of this review follows the guidelines outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) statements (25) (**Supplementary File S1**). We took a systematic approach to identify peer-reviewed articles where a mHealth intervention for maternal health was designed and implemented in a sub-Saharan African country. Studies were identified by searching for articles published between 2010 and 2021. Date limits were set in congruence with the widespread adoption of digital health foundations (such as programs, strategies and policies) across sub-Saharan African countries (26).

Eligibility criteria

We sought to identify studies reporting primary evidence regarding digital health for maternal healthcare, thus, we included research that examined the implementation, distribution and evaluation of digital health programs for maternal healthcare. We included peer-reviewed journal articles without

restrictions on the study type therefore quantitative, qualitative, and mixed-method studies were included. Maternal health refers to the health of women during pregnancy, childbirth and the postnatal period, therefore our targets were programs or interventions aiming to improve the uptake of services during pregnancy, childbirth, and post-partum follow-up which also reported gender transformative considerations such as consideration for gender roles. We also sought out studies that were conducted in a sub-Saharan country and limited the language to English due to the authors' language proficiencies.

We focused on studies that targeted women and/or men as end users, therefore studies targeting healthcare workers were excluded. We also excluded studies whereby mobile devices were only used for data collection purposes because we wanted the focus to be on devices used for intervention purposes. We did not include studies that were discussing the literature for the purpose of theory building or critique. The inclusion and exclusion table is available as a supplementary document (**Supplementary File S3**).

Search strategy

Five databases were searched from 2010 to September 2021. The databases are PsycInfo, EMBASE, Medline (OVID), CINAHL, and Global Health. We conducted test searches between September 2020-December 2020 and iteratively adjusted and refined the search strategy. We conducted initial searches in February 2021 and updated them in September 2021. Examples of search terms included "mobile health", "gender", "maternal health", and "sub-Saharan Africa". We also used synonyms, truncations, and wildcards. The electronic search strategy for the Medline (OVID) database is available as a supplementary document (**Supplementary File S2**).

Data extraction and appraisal

Studies included in this review were independently screened by two reviewers (OU and OO). The software Covidence (27) was used to organize and screen each study's title and abstract. The two reviewers subsequently conducted full text reviews of the selected studies. They assessed and resolved conflicts jointly or in consultation with the third author (SY). The two reviewers extracted the relevant data using a data extraction form developed purposely and piloted prior to review. Relevant data from quantitative and qualitative studies were collated and reported on the form. Relevant information included study design, type of mHealth intervention, study aim, intervention outcomes, findings, and limitations. We illustrated gender transformative approaches by adapting the definition of gender transformative programming into six categories as provided by WHO (18), they included ways in which programs; (1) consider gender norms/roles/relations, (2) consider women's specific needs, (3) address the causes of gender-based health inequities, (4) include ways to transform harmful gender norms, (5) seek to promote gender equality, (6) have strategies to foster progressive changes in power relationships between women and men. We were also open to including relevant data that did not fall within the WHO's definition, but we were able to align the extracted information with the predefined categories. See **Table 1** for gender consideration categories.

The reviewers appraised the quality of the manuscripts using the Mixed Methods Appraisal Tool (MMAT) (40). This tool enabled the appraisal of different classes of research including quantitative research, qualitative research, and mixed-method studies. In assessing the methodological quality of data, the tool examines the appropriateness of data collection methods, the concurrency between the study aims and data collection methods, the sample choice and the interpretation of results. We did not exclude articles based on quality scores alone because critically appraising mixed research studies remains controversial given the complexities involved (41–43). We, however, adhered to recommendations by Hong et al. (40), whereby studies not meeting the screen criteria (**Supplementary File S1** and **S2**) were not considered appropriate for appraisal. All 11 studies passed the screening. For each study design, scores were allotted percentages based on a methodological scoring system where possible items are divided by affirmative items (44, 45). Quality scores of each study were classified as weak (<50%), moderate (50%–80%) and strong (>80%). Screening questions were not allotted percentages. While Hong et al., (40) discourage an overall calculation of scores using the MMAT, we sought to provide a representation of ratings to inform the quality of studies included in this review.

Overall, the quality of the studies ranged from 0% (none of the criteria were met) to 100% (all the criteria were met). The qualitative studies were generally moderate to strong. One of the mixed-method studies was classified as weak for not meeting any of the criteria (37). The randomized control trial studies generally showed the risk of performance bias, this means that outcome assessors may have been aware of the applied intervention which could unconsciously or intentionally alter their assessment (46). Quality appraisals are detailed in **Table 2**.

Data analysis and synthesis

Data synthesis followed a convergent approach for mixed-method systematic review recommended by JBI (47). The review questions can be answered by both quantitative and qualitative studies therefore data synthesis involved data transformation by way of qualitizing. Qualitizing involves extracting data from quantitative studies and converting them to textual descriptions to allow integration with qualitative studies (47). Data were extracted using a data extraction form that collects information on the study design, type of mHealth intervention, study aims, intervention outcomes, findings, and limitations. The synthesized data are arranged in tabular forms to allow for comparison of the different approaches to gender transformative integration. The authors classified studies into subgroups according to the gender transformative categories identified within the studies. There is a global policy interest in addressing gender inequality in health programming (16). Highlighting the different

TABLE 1 Gender transformative considerations.

Author, Year	Considers gender norms/ roles/relations	Considers women's specific needs	Addresses the causes of gender-based health inequities	Includes ways to transform harmful gender norms	Seeks to promote gender equality	Strategies to foster progressive changes in power relationships between women and men
Ampt et al., (28)		The mHealth intervention was co-designed and tested with self-identified female sex workers from the target population.		The mHealth intervention acknowledges gender-based violence was likely as a result of participating in the mHealth intervention. To safe guard women, intervention ensured counseling, urgent medical treatment and protection by the community		
Dev et al., (29)		The authors identified women's limited knowledge on family planning (FP). The authors developed a FP decision aid designed to help prepare postpartum women to make personalized de-liberated contraceptive choices.			The FP program was designed to narrow the knowledge gap on family planning between men and women.	
Flax et al., (30)		The authors conducted a study apriori and identified that only 11% of women had phones. The mHealth program was designed to address cell phone gaps and enhance access to mHealth interventions by providing a group cell phone messaging intervention to promote optimal breastfeeding practices. Therefore women were able to participate even without individual phone ownership.			The mHealth program was offered as a multi-component program to improve women's financial independence through a microcredit program	
Harrington et al., (31, 32)			SMS messaging was designed to challenge personal subjective and social norms about postpartum pregnancy risk and contraceptive use.	The study was guided by women's emphasis on the need to educate men about FP in order to improve women's FP access.		The mHealth program took an innovative strategy to promote couple FP education and subsequently support couple decision-making through SMS messaging. Men provided feedback on the need to think beyond the woman-spouse dyad and include community-level engagement in FP.
Isler et al., (33)	The study considered gender norms such as the division of labour along gender lines resulting in domestic responsibilities for women. The intervention took	In evaluating the mHealth program, The authors planned data collection activities around cooking hours to allow for mothers to fulfil their household duties. Data collection took	mHealth showed that it is essential to involve male partners in mHealth maternal nutrition interventions as a means of facilitating the implementation of nutritional advice and			

(*Continued*)

TABLE 1 Continued

Author, Year	Considers gender norms/ roles/relations	Considers women's specific needs	Addresses the causes of gender-based health inequities	Includes ways to transform harmful gender norms	Seeks to promote gender equality	Strategies to foster progressive changes in power relationships between women and men
	tablets to women's door steps to show them educational videos on maternal and child nutrition.	place in nearby health centres or participants' homes to avoid mobility issues. In recognition of participants' childcare responsibilities, childcare provisions were made for women during focus group discussions as needed.	fostering constructive couple communication.			
Lund et al., (34)		The intervention design included women regardless of mobile phone and literacy status. This approach was chosen because the voucher component allowed all women, regardless of mobile phone status, access to emergency obstetric care, which the authors felt unethical to limit.				
Onono et al., (35)		mHealth Intervention provided decision-making support because the authors identified decision-making for pregnancy and childbirth service care-seeking as a complex behavior influenced by individual, family, societal, access, and health system factors.				
Parkes-Ratanshi et al., (36)			The study identified untreated men (partners to pregnant women) as primary drivers of Syphylis in pregnant women. This study aimed to increase the testing and treatment of pregnant women's male partners to reduce pregnant women's risk of syphilis.			
Schwartz et al., (37)		The content of the intervention ensured confidentiality by not disclosing women's HIV status. Messages were focused on counselling and support. Women that did not have partner support disclosed that the intervention was particularly important for them and met their needs.				

(*Continued*)

TABLE 1 Continued

Author, Year	Considers gender norms/ roles/relations	Considers women's specific needs	Addresses the causes of gender-based health inequities	Includes ways to transform harmful gender norms	Seeks to promote gender equality	Strategies to foster progressive changes in power relationships between women and men
Skinner et al., (38)		There was no cost for women to participate in the mHealth program. If a woman did not own a phone, messages were sent to another phone where she could read them.		Women indicated that the messages provided a base for discussion. The sharing of certain messages, such as around domestic violence, left the women feeling supported. Messages were shared with expectant fathers, close friends and colleagues.		
Trafford et al., (39)						Women participants attributed low levels of breastfeeding to social norms. The male gaze which indicated men's disapproval of women breastfeeding in public was cited as a reason for not breastfeeding. The messages from the mHealth program enabled women to resist pressure. Women also shared the messages with male relatives to prove the importance of breastfeeding.

approaches separately is important because it will add to the discussions on how best to promote mHealth for maternal health through a gender transformative lens and will provide evidence relevant to policy and research.

Results

Study characteristics

Figure 2 indicates a PRISMA study flow diagram describing how papers were selected for inclusion. **Table 3** provides a summary of the 11 studies that were appraised in this review. **Table 1** describes gender transformative considerations identified in the 11 studies. The studies are diverse in terms of sample size, sample population, study design and mHealth delivery system. Sample sizes ranged from 18 to 2,550 participants. Participants included pregnant and postpartum women.

Key finding 1: SMS-based services are the most common mHealth delivery system

Study designs included six qualitative studies (29, 31, 33, 35, 38, 39); three randomized control trials (28, 34, 36); and two mixed method studies (30, 37). The most common mHealth delivery system was text messages via standard short message service (SMS) on mobile phones (28, 30, 31, 33, 34, 36–39), the other approaches used interactive mobile apps (29, 35). Outcomes of interventions to improve maternal health varied across the studies. Three studies focused on reducing unintended pregnancies through the promotion of contraceptive use (28, 29, 31). Two studies focused on improving breastfeeding among postpartum mothers (30, 39), two studies aimed to increase women's access to skilled health personnel during pregnancy, childbirth and postpartum (34, 35). One study targeted improved and adequate nutrition among pregnant and breastfeeding mothers (33). Two studies aimed to prevent and manage sexually transmitted diseases among pregnant and postpartum women (36, 37).

Key finding 2: few studies substantively incorporated gender transformative dimensions in their study aims

Findings responding to the first research question indicate that all studies included at least one of the six gender transformative considerations but only three studies substantively incorporated gender transformative dimensions in the aim of their study (31, 33, 36). One study aimed to examine how gender impacts the content and delivery of a nutrition intervention focused on mothers (33). Another study aimed to involve men and women in

TABLE 2 (A) quality appraisal of qualitative studies. (B) Quality appraisal of quantitative studies. (C) Quality appraisal of mixed method studies.

Screen	Qualitative studies							Quality score
(A) Quality appraisal of qualitative studies								
	S1. Are there clear research questions?	S2. Do the collected data allow to address the research questions?	1.1. Is the qualitative approach appropriate to answer the research question?	1.2. Are the qualitative data collection methods adequate to address the research question?	1.3. Are the findings adequately derived from the data?	1.4. Is the interpretation of results sufficiently substantiated by data?	1.5. Is there coherence between qualitative data sources, collection, analysis and interpretation?	
Dev et al., (29)	✓	✓	✓	✓	✓	✓	✓	100%
Harrington et al., (31, 32)	✓	✓	✓	✓	✓	✗	✓	80%
Isler et al., (33)	✓	✓	✓	✓	✗	✗	✓	60%
Onono et al., (35)	✓	✓	✓	✓	✓	✓	✓	100%
Skinner et al., (38)	✓	✓	✓	✓	✓	✓	✓	100%
Trafford et al., (39)	✓	✓	✓	✓	✓	✓	✓	100%
Screen	Quantitative randomized control trials							Quality score
(B) Quality appraisal of quantitative studies								
	S1. Are there clear research questions?	S2. Do the collected data allow to address the research questions?	2.1. Is randomization appropriately performed?	2.2. Are the groups comparable at baseline?	2.3. Are there complete outcome data?	2.4. Are outcome assessors blinded to the intervention provided?	2.5 Did the participants adhere to the assigned intervention?	
Ampt et al., (28)	✓	✓	✓	✓	✓	✗	✓	100%
Lund et al., (34)	✓	✓	✓	✓	✓	✗	✗	60%
Parkes-Ratanshi et al., (36)	✓	✓	✓	✗	✓	✗	✗	40%
Screen	Mixed method studies							Quality score
(C) Quality appraisal of mixed method studies								
	S1. Are there clear research questions?	S2. Do the collected data allow to address the research questions?	5.1. Is there an adequate rationale for using a mixed methods design to address the research question?	5.2. Are the different components of the study effectively integrated to answer the research question?	5.3. Are the outputs of the integration of qualitative and quantitative components adequately interpreted?	5.4. Are divergences and inconsistencies between quantitative and qualitative results adequately addressed?	5.5. Do the different components of the study adhere to the quality criteria of each tradition of the methods involved	
Flax et al., (30)	✓	✓	✓	✗	✓	✗	✓	60%
Schwartz et al., (37)	✓	✓	✗	✗	✗	✗	✗	0%

discussions around family planning education and decision-making (31). Finally, one aimed to encourage men (partners of pregnant women) to get tested and treated for sexually transmitted infections (STI) to decrease incidences of STIs in women during pregnancy (36). For the rest of the studies, gender considerations were not explicitly stated but treated tangentially within the mHealth program's design, implementation, or evaluation.

Key finding 3: a common gender consideration was of women's specific needs

Two studies each included 3 gender transformative considerations (31, 33), four studies included 2 gender transformative considerations (28–30, 38), the rest of the studies only had one (34–37, 39). Two studies included strategies to promote gender equality (29, 30). Strategies to promote gender

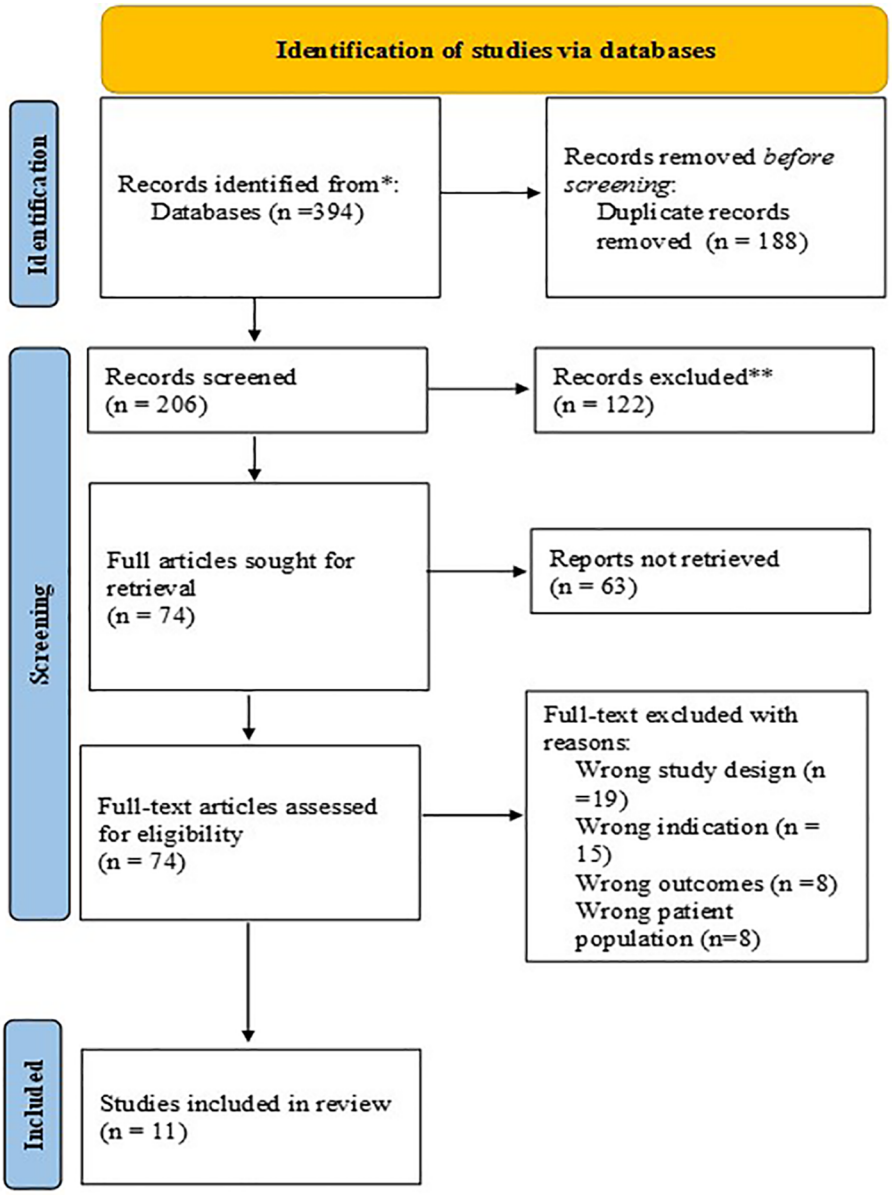


FIGURE 2
PRISMA flow diagram depicting the flow of information screened and reviewed.

equality included closing the knowledge gaps about family planning between men and women and improving women's financial stability through microcredit programs. Most of the gender considerations fell under the category of considering women's specific needs.

One study indicated consideration for women's specific needs by co-designing a mHealth program aimed at improving contraceptive knowledge and use with the target population, this approach enabled the researchers to integrate the needs of the women into their program (28). Four studies conducted preliminary research with their target population and designed mHealth programs based on identified needs; One study identified knowledge gaps as a barrier to women's decision-making about family planning and subsequently designed a mHealth program to educate women on contraceptive choices and enhance their decision-making regarding family planning (29). Another study was informed by formative studies that linked limited decision-making support to the use of pregnancy care. Consequently, the researchers designed a mHealth intervention to support women's decision-making (35). Two studies considered women's specific needs by first conducting a formative study that revealed a gender gap in women's access to mobile devices (30, 34). The researchers designed their mHealth program to optimize women's participation even without individual mobile phone ownership or with low literacy status. Women were able to participate either using group cell phones or shared cell phones. In addition, one study considered women's specific needs by being mindful of their schedules, domestic responsibilities, and transportation challenges during their study (33). The mHealth intervention, which involved

TABLE 3 Summary of studies selected for review.

Study authors, title, year	Country	Number of participants	Sample of population	Study design	mHealth delivery system
Ampt et al., (28) Effect of a mobile phone intervention for female sex workers on unintended pregnancy in Kenya	Kenya	786 women	−93 randomly selected sex-work venues in two sub-counties of Mombasa, Kenya (Kisauni and Changamwe) −401 participants from the intervention group from 47 venues −385 participants from the control group from 46 venues	A two-arm, cluster-randomized controlled trial study	SMS text messages
Dev et al., (29) Acceptability, feasibility, and utility of a mobile health family planning decision aid for postpartum women in Kenya	Kenya	25 postpartum women and 17 Family planning providers	Twelve (48%) postpartum women were from rural MCH clinics from the Nyanza region, and 13 (52%) were from urban clinics in Nairobi; 15 (60%) were adolescents and young women between the ages of 14–21.	A cross-sectional qualitative study	An interactive mobile application
Flax et al., (30) Group cell phones are feasible and acceptable for promoting optimal breastfeeding practices in a women's microcredit program in Nigeria.	Nigeria	375 postpartum women participants in total: Quantitative study $n = 195$ Qualitative study $n = 162$		Mixed method: interview and Focus group discussions	SMS text messages and voice message
Harrington et al., (31, 32) Engaging men in a mHealth approach to support postpartum family planning among couples in Kenya: a qualitative study.	Kenya	50 pregnant and postpartum women and men	35 men and 15 women Participants were chosen from two counties in the Nyanza region of Western Kenya. These hospitals serve a primarily low- to middle-income rural population, the majority of whom identify with the Luo ethnic group.	Qualitative study: Focus group discussions	SMS messages
Isler et al., (33) "If he sees it with his own eyes, he will understand": how gender informed the content and delivery of a maternal nutrition intervention in Burkina Faso.	Burkina Faso	78 pregnant or breastfeeding women and 8 Community Health Workers	The study sampled the catchment areas of two urban and four rural health centres. Healthcare providers identified pregnant and breastfeeding mothers to participate in the study	Qualitative study: focus group discussion	Video was shown on a tablet
Lund et al., (34) Mobile phones as a health communication tool to improve skilled attendance at delivery in Zanzibar: a cluster-randomized controlled trial.	Tanzania	2,550 pregnant women	1,311 women were allocated to the mHealth interventions and 1,239 women were allocated to standard care, i.e. no phone intervention. Participants were recruited from 24 primary healthcare facilities.	Cluster-randomized controlled trial with	SMS
Onono et al., (35) Narratives of women using a 24-h ride-hailing transport system to increase access and utilization of maternal and newborn health services in rural western Kenya: a qualitative study.	Kenya	18 postpartum women	Emphasis was placed on ensuring participants were from different socioeconomic backgrounds. Women were on average 27 years old, married and multiparous with over half having secondary education.	A qualitative study using in-depth interviews (IDIs) as the primary data collection method.	A mobile phone app

(*Continued*)

10.3389/fdgth.2023.1263488

TABLE 3 Continued

Study authors, title, year	Country	Number of participants	Sample of population	Study design	mHealth delivery system
Parkes-Ratanshi et al., (36) Low male partner attendance after syphilis screening in pregnant women leads to worse birth outcomes: The Syphilis Treatment of Partners (STOP) randomized control trial	Uganda	290 pregnant women	144 participants enrolled in SMS reminder, 146 enrolled in telephone call reminder	The sample was taken from the IDI clinic located within urban Kampala, where 220 pregnant women are seen at the ANC per month, 5.1% of whom tested positive for syphilis, and the Mulago Hospital ANC, where, on average, 2,000 pregnant women are seen monthly, with 2.4% testing positive for syphilis.	Text messages and telephone calls
Schwartz et al., (37) Acceptability and Feasibility of a Mobile Phone-Based Case Management Intervention to Retain Mothers and Infants from an Option B + Program in Postpartum HIV Care	South Africa	50 pregnant women	HIV-infected, pregnant women attending antenatal care (ANC) at Witkoppen Health and Welfare Centre (WHWC)	Quantitative cohort study	SMS text messages and phone calls
Skinner et al., (38) User assessments and the use of information from MomConnect, a mobile phone text-based information service, by pregnant women and new mothers in South Africa.	South Africa	46 women (pregnant women and new mothers)	Participants were purposively selected from 15 facilities in five provinces—Western Cape, KwaZulu-Natal, Free State, Gauteng and Mpumalanga—which were purposively selected to represent different language and cultural groups in South Africa. Within each province, three facilities were purposively selected among those serving large urban communities and those serving semirural communities or villages.	Qualitative study: interviews and focus groups.	Text-based information service
Trafford et al., (39) Reported infant feeding practices and contextual influences on breastfeeding: qualitative interviews with women registered to MomConnect in three South African provinces	South Africa	115 pregnant and postpartum women	Women over the age of 18 with an expected due date of delivery between April and October 2017 were recruited -using a series of SMS surveys sent out in English and the most common local language in each region—isi-Zulu (KwaZulu-Natal, Gauteng) or Sesotho	Qualitative study: included IDIs and FGDs.	Text messages via standard short message service (SMS) on mobile phones

Study authors, date	Intervention description	Study aims	Intervention outcomes	Major findings	Limitations
Ampt et al., (28)	WHISPER is an SMS intervention. Text messages in English are delivered two to three times per week for 12 months. Message content in the intervention group focused on promoting contraception, particularly long-acting reversible contraception. Message content in the control group focused on promoting nutritional knowledge and practices, including food safety, preparation, and purchasing.	This study aimed to assess the effectiveness of the intervention to reduce the incidence of unintended pregnancy among sex workers in Kenya.	Incidence of unintended pregnancy. Long-acting reversible contraceptive use. Dual contraceptive-method use. Contraceptive knowledge.	The intervention did not have a clinically significant effect on unintended pregnancies despite being developed in consultation with participants (sex workers). The intervention was associated with improved knowledge about contraception, particularly intrauterine devices, as well as increased use of dual-method contraception.	Short-lived measurement of outcomes The study did not account for selective loss of pregnancy among sex workers.

(*Continued*)

10.3389/fdgth.2023.1263488

TABLE 3 Continued

Study authors, date	Intervention description	Study aims	Intervention outcomes	Major findings	Limitations
Dev et al., (29)	iMACC is a client-facing mobile application designed to provide systematic, yet personalized, contraceptive counselling to postpartum women. iMACC combines images and text in a heuristic approach to provide tailored contraceptive counselling to women.	This study aimed to evaluate the acceptability and feasibility of the self-administered iMACC decision aid	Behaviour change	Participants indicated that the decision aid was easy to use but challenges arose as to low literacy levels posing a barrier to understanding and using the app. – The app was perceived as a confidential decision aid. – Improved knowledge of contraception with a personalized approach.	The evaluation study did not completely reflect the underlying view of participants towards the FP decision aid.
Flax et al., (30)	Weekly cell phone breastfeeding text and voice messages were delivered to women in groups during group breastfeeding learning sessions	This study aimed to examine the feasibility and acceptability of using group cell phones to deliver cell phone messages within a multi-component breastfeeding promotion intervention. The study also aimed to test the association between participation in small groups and reported breastfeeding practices. Methods	Increased the odds of exclusive breastfeeding. Shift social norms around breastfeeding practices.	Group phones were described as acceptable and functional. Participants reported improved status of the phone holder because she was seen as a leader. Some non-phone recipients wanted their own phones. Health information was passed along to family and friends.	Giving out cell phones to groups is not widely sustainable.
Harrington et al., (31, 32)	The intervention is an interactive SMS program that supports women and couples in contraceptive decision-making, method initiation and continuation.	The study aimed to explore men's and women's perspectives on using SMS to facilitate postpartum family planning counselling in Kenya. It also aimed to engage men in family planning decision-making.	Family planning initiation and continuation	Men strongly desire inclusion in FP decision-making. Men participants indicated that they were open to receiving family-panning related SMS, and discussed multiple advantages of using SMS dialogue for FP counselling. The majority of men and women in FGDs felt that receiving SMS about FP could promote improved communication with their partners.	There may have been selection bias of married women participants. Issues of access to maternal healthcare may be different for non-married women.
Isler et al., (33)	Healthcare providers visited pregnant and breastfeeding mothers at home to show them a set of maternal nutrition videos on a tablet.	To examine how gender affects the content and delivery of a nutrition-focused intervention in Burkina Faso.	Gender-sensitive nutrition intervention	Some mothers said that while they appreciated being shown the videos, they felt powerless to make nutritional changes without the support of their male partners. Healthy promotion in areas such as healthy nutrition, pregnancy and breastfeeding is largely considered a mother's domain. Mothers divulged feeling embarrassed at the prospect of discussing certain topics with men, including intra-household dynamics around nutrition financing.	The study did not interview fathers therefore, it lacks the perspectives of men and their role in maternal and child nutrition.

(*Continued*)

10.3389/fdgth.2023.1263488

TABLE 3 Continued

Study authors, date	Intervention description	Study aims	Intervention outcomes	Major findings	Limitations
Lund et al., (34)	An automated short messaging service (SMS) system providing mothers with unidirectional text messaging, and a mobile phone voucher system providing the possibility for direct two-way communication between mothers and their primary healthcare providers	The study aimed to evaluate the use of mobile phones to bridge the communication gap between pregnant women and health providers. The study also aimed to increase skilled delivery attendance in a setting with scarce resources.	Increased skilled delivery attendance amongst pregnant women.	60% of women in the intervention group delivered in the presence of a skilled birth attendant, compared to 47% of women in the control group. However, the intervention did not improve skilled birth attendance in rural areas. Not owning a mobile phone was associated with lower odds of skilled delivery attendance.	Methodological limitation in study design and external validity.
Onono et al., (35)	MAcess is a phone app with the following functions 1) A free SMS-based service that involves sending personalized trimester-based texts to mothers and reminders to use ante- and postnatal care services; 2) An interactive chat service known as m-convo; and 3) 24-h transport navigator system, through which women can request transport	The study aimed to explore how pregnant and postnatal women made decisions regarding care-seeking for pregnancy and childbirth; the processes of getting care from home to the hospital as well as their perceptions on how the MAccess intervention affected their pregnancy and childbirth care-seeking and utilization experience.	Reduction in maternal mortality through frequent communication with healthcare providers and access to reliable transport services.	The bidirectional text messages influenced the decision-making process for the mothers by increasing their knowledge of danger signs, individual birth plans, possible complications, and immediate post-delivery neonatal care. The bidirectional SMS system offered by the MAccess mitigated the challenge of inadequate information. MAccess linked mothers to reliable transport providers to transport them to hospitals. The mHealth innovation was found to be attractive and highly acceptable to women throughout pregnancy and childbirth and helped them navigate the complex and layered individual, infrastructural, and health system factors that put them at risk of adverse maternal and newborn outcomes.	The gatekeeping role of nurses who influenced participant selection might have introduced a bias.
Parkes-Ratanshi et al., (36)	Weekly text message reminders are sent out to women participants to encourage their partners to get tested for syphilis in clinics. This message lasts up to 8 weeks after women's diagnosis.	The aim of this study was to determine the effectiveness of three partner notification strategies, they include: standard of care (SOC) notifications vs. SOC plus text message reminders vs. SOC plus telephone call reminders.	The proportion of partners who presented at the clinic and received syphilis testing or treatment.	Post-enrollment partner attendance was 18%. This was considered low attendance. There were no significant differences between the methods used (SOC notification slip, telephone call or SMS).	The attendance rate recorded in this study is lower than the national average. Insufficient attention to cultural and systemic barriers faced by women to inform their partners of syphilis status and partner getting tested.

(*Continued*)

10.3389/fdgth.2023.1263488

TABLE 3 Continued

Study authors, date	Intervention description	Study aims	Intervention outcomes	Major findings	Limitations
Schwartz et al., (37)	This intervention provides support to HIV-infected women on highly active antiretroviral therapy (HAART) during late pregnancy and for six weeks postpartum. Participants received weekly pre-scripted messages from Case managers during pregnancy (36 weeks) and up to 8 weeks postpartum. Case managers made one pre-delivery and two post-delivery telephone calls to participants during follow-up to discuss delivery plans and postpartum care.	The objective of the study was to assess the acceptability, feasibility and potential for scale-up of the mobile phone-based case management intervention.	Retention of mothers in HIV care.	Findings showed that 98% of participants indicated that it was helpful to have a case manager assigned to support them during their pregnancy and postpartum, and all women indicated that they would recommend the case management program to a friend. The pilot intervention provided support to HIV-infected women on HAART during late pregnancy and for six weeks postpartum to be feasible and was highly acceptable. Post-partum contact at 6 weeks was maintained via cellphone with 88% of enrolled women and via telephone or clinic visits amongst 96% of women.	The study was not powered or designed to assess efficacy. The study only collected prospective data and completed questionnaires with women enrolled in the pilot intervention.
Skinner et al., (38)	MomConnect is a text service provided to pregnant women and new mothers. Women received text messages covering broad areas of child care and health.	The study aimed to describe the experiences of pregnant women and new mothers with MomConnect and assess participants' perceived value of MomConnect, and obtain suggestions for its further improvement.	Empowered mothers who have skills to care for their children.	The pregnant women and mothers uniformly were highly appreciative of MomConnect. First-time mothers felt that they had a lot to learn and drew support and confidence from the messages. MomConnect met the requirement of being responsive to the target group's needs.	
Trafford et al., (39)	Mom connect was introduced as a strategy to increase breastfeeding rates. Women registered with MomConnect receive health information through text messages and have access to a helpdesk which allows them to ask questions about health services received.	To provide evidence towards a richer understanding of women's practices around breastfeeding, key influences on their decision-making, and what might facilitate or prevent the enactment of their breastfeeding intentions.	Increased breastfeeding initiation	MomConnect messages (often used in conjunction with health workers' advice) were noted as a very valuable source of information that supported women's knowledge and decision-making. Women found MomConnect empowering and used the messages to teach their peers.	Many of the women were unemployed. Evidence shows an association between low rates of employment and high rates of breastfeeding.

10.3389/fdgth.2023.1263488

presenting nutrition information through an interactive video, was delivered to women at their doorstep. One study met the needs of pregnant and postpartum women with HIV through the provision of HIV counselling and support (37). This study also guaranteed women's privacy by protecting their sensitive health information. Another study delivered mHealth programs at no cost to low-income pregnant and new mothers (38).

Key finding 4: men have a pivotal role in maternal health

Turning to another category of gender transformative consideration, the aims or outcomes of three studies addressed causes of gender-based health inequities. One study indicated that SMS messages from the mHealth program challenged social norms around the use of contraceptives and pregnancy risk (31). Another mHealth program sought to enhance adequate nutrition among pregnant women by involving men who are often major decision-makers in maternal nutrition (33). Finally, one study identified untreated men partners as primary drivers of syphilis in women during pregnancy, therefore the program targeted pregnant women's partners to test for and treat syphilis symptoms (36). Three studies included considerations under the category of seeking to transform harmful gender norms. One acknowledged that women faced an increased likelihood of gender-based violence due to participating in the mHealth study (28). The authors arranged for the protection of women by providing urgent medical care where necessary and garnered support for and protection of women from community mobilizers. In another study, transforming gender norms also meant educating couples (men and women) about family planning through SMS text messages and supporting their joint decision-making (31). One study encouraged women to share text messages on pregnancy and child care with their spouses (38). Within these messages, the dangers of domestic violence were emphasized. Women reported a sense of support from receiving and sharing messages with their spouses.

Gender considerations in two studies indicated strategies to foster progressive changes in power relationships between women and men; One study engaged men in family planning education and decision-making support and also employed innovative strategies to go beyond couples or individual interventions but also community-level engagement to improve knowledge on family planning (31). Another study fostered progressive changes in power relationships between women and men by legitimizing the importance of breastfeeding through SMS text messages (39). Men's disapproval of breastfeeding deterred women from breastfeeding, however, women indicated that receiving and sharing the text messages from the mHealth program enabled them to resist pressure from men and encouraged breastfeeding.

Key finding 5: findings from gender transformative mHealth programs indicate positive results overall

Studies included in this review showed positive results overall. One mHealth program aimed at altering postpartum women's habits and behaviour toward contraceptive use saw improvements in women's knowledge of contraceptives (29). Another mHealth program aimed at increasing exclusive breastfeeding practice among postpartum women was described as acceptable and functional by the participants (30). Including men in a mHealth family planning program for postpartum women improved their communication with their women partners around contraceptive use (31). One study saw an increased rate of skilled delivery attendance amongst women participants as a result of the mHealth program (34). A mHealth program that provided a 24-hour transport navigator system reported improved maternal access to skilled pregnancy care including virtual communications with their healthcare providers (35). In another mHealth study, an intervention that aimed to retain and support HIV-infected mothers was perceived as helpful and supportive by participants (37). In two studies, participants demonstrated the positive impact of MomConnect, a mHealth program for pregnant and postpartum women. The mHealth program was responsive to the needs of new mothers and served as an empowering force toward positive breastfeeding practices for women (38, 39).

However, not all studies reported positive findings. One mHealth program was developed jointly with target participants in order to reduce incidents of unintended pregnancies, however, the program showed no clinically significant effect on unintended pregnancies among participants (28). Additionally, other studies indicated the need for a more explicit focus on gender consideration in a mHealth program's design or implementation. One study targeted pregnant and breastfeeding women to educate them on adequate nutrition during pregnancy (33). While participants improved their knowledge of appropriate nutrition during pregnancy and postpartum, they were powerless to make any nutritional changes without support from their male partners. Similarly, another study aimed at encouraging the testing and treatment of STIs among male partners of pregnant women indicated a limited or low effect of the program (36). The authors pointed to insufficient gender considerations in the mHealth design and implementation. In another mHealth study, limited considerations of intersecting domains of disadvantages, specifically gender and geographic location, led to the exclusion of the most vulnerable of women (34). In the study which aimed to improve women's access to skilled birth attendants, women were able to participate in the mHealth program regardless of phone ownership or literacy status. The study saw improvements in access to skilled birth attendants in urban areas but failed to reach rural women who were in dire need of skilled attendants during childbirth (34).

Key findings 6: gender considerations and maternal health outcomes

Furthermore, this review offered some evidence on how gender considerations influenced maternal health outcomes. In considering gender differences, one study identified crucial knowledge gaps that hampered post-partum women's use of modern contraceptives (29). Women's unmet need for contraception was exacerbated by their limited knowledge on contraceptives. Through the mHealth program, women showed

improved knowledge and more thorough understanding of contraceptives. The authors highlighted the potential of increased knowledge to improve contraceptive use among postpartum women (29). In another study, specific considerations for women's limited phone ownership increased their odds of exclusive breastfeeding for up to 6 months (30). Through the use of group cell phones, women received text messages that promoted optimal breastfeeding practices and were more likely to breastfeed exclusively for the first 6 months. In a similar study aiming to improve breastfeeding rates, text messages shared with women and their families targeted unfavourable social norms (39). Women felt empowered to make breastfeeding choices and to resist pressure against breastfeeding that was often brought about by patriarchal norms. Women in the study reported high rates of breastfeeding. In a study aimed at improving postpartum retention in HIV care, a mobile health program delivered health information and reminder text messages to women directly from their healthcare providers (37). Gender considerations ensured that women's HIV status was not disclosed in those text messages. This study showed improved communication between women and healthcare providers, especially among women who wanted to maintain the privacy of their health information. Overall, interactions with healthcare providers contributed to women's retention in HIV care (37).

Discussion

To our knowledge, this is the first systematic review reporting evidence on gender transformative approaches in mHealth programs for maternal healthcare in sub-Saharan Africa. The study highlights the various approaches to integrating gender transformative approaches in mHealth studies in line with the WHO's definition of a gender transformative approach to health programs (18). The findings indicate that while most of the evidence of transformative approaches centred on considering women's specific needs, there was a limited focus on advancing gender equality. No study covered the entire categories of gender transformative approaches and a few studies included approaches from a maximum of three categories.

Our findings with the most significant policy concern are the limited number of mHealth programs with an explicit focus on gender transformative considerations. Gender transformation was not necessarily central to most mHealth programs although they manifested during the study. This highlights the need for an explicit and intentional focus on gender considerations and the promotion of gender equality in mHealth programs for maternal healthcare. Our findings indicated that consideration of only one gender target is often insufficient to effect change. For instance, in a mHealth study to improve nutrition during pregnancy and early childhood, women who were target participants improved their knowledge of adequate nutrition, but improved awareness did not translate into appropriate action because men were not actively engaged in the program. Improving nutrition by targeting women alone presupposes their access to financial resources and decision-making power. In line with this insight are findings from studies in sub-Saharan Africa that illuminate the gender power dynamics inherent in the context of women's nutrition and health (48). The study highlighted the importance of considering women's broader social, cultural, and economic realities and involving men in health interventions.

Engaging men in and of itself is not a panacea as illustrated by another mHealth study in our review. The study observed that men were the predominant drivers of syphilis in pregnancy and encouraged women to recruit their men partners to test for and treat STIs (36). The study saw poor attendance from men and contended that gender-based barriers prevented effective communication between partners. A similar study in Congo highlighted the dangers of poorly designed mHealth programs for engaging men in maternal healthcare (49). The study, jointly targeting men and women, was designed to bridge the knowledge gap around modern contraceptives but instead saw higher participation among men than women. The study failed to account for the digital gap whereby men were often primary users of technology. As evidenced by our findings and the broader literature, engaging men in maternal health requires a deeper consideration of men's privilege and power over women (4). Men need to be engaged meaningfully in maternal health programs.

Encouragingly, most of the studies showed positive findings in advancing women's access to maternal healthcare services. Specifically, our findings show evidence of positive outcomes in multi-sectorial approaches to enhancing maternal health. One study from our review integrated breastfeeding promotion into a microcredit program for pregnant mothers in Nigeria. The aim was to improve women's financial stability while supporting breastfeeding through a mHealth program. Similar studies in the literature demonstrate how multi-pronged gender transformative programs for maternal health led to positive health outcomes. A mHealth program in Kenya empowered women in informal employment sectors to save for maternal health expenditure as well as improve their knowledge of maternal healthcare (50, 51). When financially empowered, women are more likely to seek and adhere to skilled maternal health care (50, 51). Similarly, programs to redress anemia in pregnant women in Burkina Faso and DRC went beyond nutrition-related activities to involve women in sanitation supply chain initiatives, enhance women's leadership in communities and shed light on gender-based violence (52). These examples show a recognition of the complex and interconnected factors that determine maternal health. They also highlight the potential of mHealth to facilitate a multisectoral approach to redress maternal mortality and morbidity.

Findings from our study illustrate the influences of gender considerations on maternal health outcomes. Our studies highlight the importance of gender considerations such as acknowledging that women are more likely than men to be digitally excluded. According to the Mobile Gender Gap Report 2021 published by the GSMA, the gender gap in mobile phone ownership in sub-Saharan Africa is at 13%, this translates to 74 million women who do not own a mobile phone (13). Studies

have shown that enhancing women's access to mobile devices enhances their participation in mHealth studies, increases their use of maternal healthcare services, and consequently improves maternal health outcomes (24, 53). Our findings show that women's perception of the security of their health information impacts their use of mHealth programs. A mHealth study that guaranteed women's privacy saw increased engagement with the program and subsequent improvement in maternal health outcomes (37). Similar to our findings, evidence from Tanzania shows that positive perceptions of personal privacy and security of a mHealth program enhances pregnant women's participation in the program (54). The study also showed enhanced relationships between women and their healthcare providers.

Policy implications

Implications of our findings for policy have been interspersed in the discussion. We draw further attention to privacy as a growing concern in digital health especially as it pertains to sensitive health information (55). Disclosing private health information puts women at increased risk of violence (37). Our findings showed strategies for circumventing privacy issues such as purposefully designed mHealth programs that deliver general messages on HIV without divulging women's HIV status. An additional strategy could be the integration of password-protected messages to ensure that only the intended recipient reads messages. Beyond program-level strategies, the Global Strategy on Digital Health advocates for country-level regulatory frameworks to enhance the protection and confidentiality of health data with the use of digital health (56). To address the challenges identified in our findings, gender considerations must be integrated into the planning and implementation of these frameworks. Our findings also indicate the need for improved digital access for women. Addressing issues related to affordability and literacy is key to enhancing women's access to and use of mobile health technologies (57). This will require cross-sectoral collaborations and an explicit focus on gender perspectives in policies and plans for digital health. For instance, subsidizing phones for women and girls and digital literacy programs can overcome gendered access barriers to mobile technologies (57).

An understanding of the existing gender ecosystem maximizes the potential of digital health innovations and minimizes risks particularly as it relates to engaging men in maternal health (58). Our findings show the need for men to be engaged meaningfully in maternal health programs. Strategies to enhance male engagement in sub-Saharan African countries have included the development of male engagement guidelines as evidenced in Tanzania (59). It is important to note that while well-intended, unintended consequences of these guidelines have been shown to present challenges for women. For instance, partner absence during antenatal care visits has resulted in delays in women seeking healthcare or refusal of care by healthcare workers (59). It is important to understand the existing gender ecosystem and assess the unintended consequences of strategies to engage men in maternal health.

The lack of programs that address all the gender-transformative categories indicates the need for a gender objective in each digital health program. Canada's International Development Research Centre (IDRC) recommends that digital health programs in underserved communities should include at least one research question or objective that aims to understand gender issues (58). This will address the noticeable risks of inadequate gender considerations as observed in some of the studies. In our findings, studies with explicit gender objectives also allowed gender to inform further actions in the research process such as data collection. Therefore, beyond having a clear objective to consider gender issues, it is important to maintain a commitment to adapting programs as gender issues become apparent during the course of a program.

Future research

As demonstrated in our review, there is limited research on gender transformative approaches in mHealth for maternal health in sub-Saharan Africa. Given the transformational potential of digital health, there is a need for more research on how digital health can reduce inequalities for end users, especially women and girls. Research is needed to investigate how gender inequalities shape assumptions, design and implementation of digital health tools. Studies also indicate the need to meaningfully engage gatekeepers in society who enforce gender power relations to enhance the success of digital health programs. Critical gaps identified in our study point to the need for methodologically strong gender transformative studies. There needs to be greater consistency in quality terminology and criteria that accommodate different study contexts. Future studies can investigate the adverse effects of enacting gender transformative approaches including familial tension because of changes in gender dynamics in relationships.

Strengths and limitations

This study reviewed evidence from both quantitative and qualitative studies thereby uncovering gender as presented from different perspectives. This approach allowed the authors to examine a robust pool of data while gaining insights into users' experiences of gender-transformative mHealth programming. These may not have been possible with only a quantitative or qualitative review of evidence. Despite the generally successful outcomes of gender transformative studies, these studies should be interpreted with caution in light of a few low-quality studies. Low methodological quality scores of studies are indicative of poorly designed studies, therefore, while they may include the relevant gender transformative dimensions, methodological gaps and low-quality studies may exaggerate result outcomes and lead to incorrect inferences. There is a need for more rigorous study designs, especially for mixed methods mHealth studies for maternal

healthcare. Furthermore, our analyses of findings indicate strong individual and community-level approaches to gender integration in mHealth programs. Similar approaches have been shown to transform gender norms and health-related outcomes in sub-Saharan Africa (60). However, previous research emphasizes gender transformative approaches at the structural level including legal or policy approaches (60, 61). These approaches have been shown to transform health challenges brought about by gender inequality and achieve effective and sustainable change.

Due to the language limitations of the authors, there was no non-English mHealth study included in this review. The authors may have missed other relevant studies that provide evidence on gender transformative approaches. Another limitation is that while the authors extracted the relevant data using the WHO definition as a guide, the gender transformative parameters were not explicitly stated in the studies. This calls attention to the need for clear reporting guidelines for gender considerations, especially in mHealth research. The literature shows a growing recognition of the importance of consistent standards for reporting gender considerations in health research, however, the deficiencies in the quality of reporting remain an issue (62, 63).

Conclusion

Digital health has been incorporated across the pregnancy care continuum in efforts to address social determinants of health, improve the quality of care and ensure positive maternal health outcomes. To achieve meaningful impact, gender and digital inclusion must remain a priority in developing, implementing and evaluating digital health programs. This study reviews gender transformative approaches to gender integration in mHealth for maternal health in sub-Saharan Africa. This review adopts the definition offered by the WHO and interprets a gender-transformative digital health program as one that "addresses the causes of gender-based health inequities through approaches that challenge and redress harmful and unequal gender norms, roles, and power relations that privilege men over women". Considering gender in and of itself is not sufficient because some pathways to gender consideration in health can exploit or accommodate harmful gender norms rather than transform them. While this review affirms that gender transformative approaches in digital health programs are advancing maternal healthcare outcomes, we noted that most programs were not substantively incorporating these considerations into their design, implementation, or evaluation. Implications of our study findings indicate the need for mHealth studies to explicitly acknowledge how power dynamics, values and norms impact maternal health and address these factors throughout the course of a mHealth program.

Data availability statement

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

Author contributions

OU: Conceptualization, Formal Analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. OO: Methodology, Supervision, Writing – review & editing. SY: Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

OU acknowledges the support from the Social Sciences and Humanities Research Council of Canada (SSHRC) through the Joseph-Armand Bombardier Graduate Scholarship for her doctoral studies.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2023.1263488/full#supplementary-material

References

1. WHO. *WHO guideline: Recommendations on digital interventions for health system strengthening*. Geneva: World Health Organization (2019).

2. UN. Beijing declaration and platform for action. Vol. 21, *Fourth World Conference on Women*. (1995). p. 1–2. Available at: http://www.un.org/womenwatch/daw/beijing/platform/health.htm

3. UNFPA. International Conference on Population and Development, 1994. *International Conference on Population and Development*. (2022). p. 1. Available at: https://www.unfpa.org/icpd (cited 2022 Nov 11).

4. Ruane-McAteer E, Amin A, Hanratty J, Lynn F, Corbijn Van Willenswaard K, Reid E, et al. Interventions addressing men, masculinities and gender equality in sexual and reproductive health and rights: an evidence and gap map and systematic review of reviews. *BMJ Glob Heal*. (2019) 4(5):e001634. doi: 10.1136/bmjgh-2019-001634

5. United Nations. Global indicator framework for the Sustainable Development Goals and targets of the 2030 Agenda for Sustainable Development. Work Stat Comm Pertain to 2030 Agenda Sustain Dev. (2020).1–21. Available at: https://unstats.un.org/sdgs/indicators/GlobalIndicator Framework after 2019 refinement_Eng.pdf%0A https://unstats.un.org/sdgs/indicators/Global Indicator Framework_A.RES.71.313 Annex.pdf

6. WHO. *Digital implementation investment guide (DIIG): Integrating digital interventions into health programmes*. Geneva, Switzerland: World Health Organization (2020). 1–182.

7. George AS, Morgan R, Larson E, LeFevre A. Gender dynamics in digital health: overcoming blind spots and biases to seize opportunities and responsibilities for transformative health systems. *J Public Health*. (2018) 40(suppl_2):ii6–ii11. doi: 10.1093/pubmed/fdy180

8. Figueroa CA, Luo T, Aguilera A, Lyles CR. The need for feminist intersectionality in digital health. *Lancet Digit Heal*. (2021) 3(8):e526–33. doi: 10.1016/S2589-7500(21)00118-7

9. Job N, Kuhnert K-L, Engelbrecht L. Mobile health and narrative entertainmenteducation to promote exclusive breastfeeding in western cape, South Africa. *South African J Clin Nutr*. (2021) 34(3 PG-54):54. doi: 10.1080/16070658.2021.1968126

10. Coleman J, Bohlin KC, Thorson A, Black V, Mechael P, Mangxaba J, et al. Effectiveness of an SMS-based maternal mHealth intervention to improve clinical outcomes of HIV-positive pregnant women. *AIDS Care*. (2017) 29(7):890–7. doi: 10.1080/09540121.2017.1280126

11. Omonaiye O, Nicholson P, Kusljic S, Manias E. A meta-analysis of effectiveness of interventions to improve adherence in pregnant women receiving antiretroviral therapy in sub-Saharan Africa. *Int J Infect Dis*. (2018) 74:71–82. doi: 10.1016/j.ijid.2018.07.004

12. Mlambo C, Sibanda K, Ntshangase B, Mvuyana B. ICT and Women's health : an examination of the impact of ICT on maternal health in SADC states. *Healthcare*. (2022) 10(802):1–16. doi: 10.3390/healthcare10050802

13. GSMA. GSMA Connected Women—The Mobile Gender Gap Report 2021 (2021). Available at: www.gsmaintelligence.com (Accessed January 2, 2022).

14. Phillips TK, Bonnet K, Myer L, Buthelezi S, Rini Z, Bassett J, et al. Acceptability of interventions to improve engagement in HIV care among pregnant and postpartum women at two urban clinics in South Africa. *Matern Child Heal J*. (2019) 23(9):1260–70. doi: 10.1007/s10995-019-02766-9

15. Udenigwe O, Okonofua FE, Ntoimo LFC, Yaya S. Enablers and barriers to the acceptability of mHealth for maternal healthcare in rural Edo, Nigeria. *Dialogues Heal*. (2022) 1:100067. doi: 10.1016/j.dialog.2022.100067

16. Rutgers. Adopting the gender transformative approach in sexual and reproductive health and rights, and gender-based violence programmes. Utrecht: Rutgers (2018). Available at: https://rutgers.international/resources/rutgers-gta-toolkit-module-1/

17. Interagency Gender Working Group. Understanding and applying the gender equality continuum. (2016).

18. WHO. *Gender mainstreaming for program managers; A practical approach*. Geneva: World Health Organization (2011). Available at http://apps.who.int/iris/bitstream/10665/44516/1/9789241501071_eng.pdf

19. Morgan R, George A, Ssali S, Hawkins K, Molyneux S, Theobald S. How to do (or not to do)…gender analysis in health systems research. *Health Policy Plan*. (2016) 31(8):1069–78. doi: 10.1093/heapol/czw037

20. Doyle K, Levtov R, Barker G, Bastian GG, Bingenheimer JB, Kazimbaya S, et al. Gender-transformative bandebereho couples' intervention to promote male engagement in reproductive and maternal health and violence prevention in Rwanda: findings from a randomized controlled trial. *PLoS One*. (2018) 13(4):e0192756. doi: 10.1371/journal.pone.0192756

21. Abramsky T, Devries K, Kiss L, Nakuti J, Kyegombe N, Starmann E, et al. Findings from the SASA! study: a cluster randomized controlled trial to assess the impact of a community mobilization intervention to prevent violence against women and reduce HIV risk in Kampala, Uganda. *BMC Med*. (2014) 12(1):15–7. doi: 10.1186/s12916-014-0122-5

22. MSD. *How maternal health innovations are closing the digital divide a merck for mothers report 2*. Rahway, NJ: Merck for Mothers (2021).

23. Udenigwe O, Okonofua FE. Understanding gender dynamics in mHealth interventions can enhance the sustainability of benefits of digital technology for maternal healthcare in rural. *Front Glob Women's Heal*. (2022) 3:1002970. doi: 10.3389/fgwh.2022.1002970

24. Oyeyemi SO, Wynn R. Giving cell phones to pregnant women and improving services may increase primary health facility utilization: a case-control study of a Nigerian project. *Reprod Health*. (2014) 11(1):1–8. doi: 10.1186/1742-4755-11-8

25. Page MJ, Mckenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement : an updated guideline for reporting systematic reviews. *Int J Surg*. (2021) 88(March):1–9. doi: 10.1016/j.ijsu.2021.105906

26. WHO. Atlas of eHealth country profiles: the use of eHealth in support of universal health coverage: based on the findings of the third global survey on eHealth 2015. *Global observatory for eHealth*. Geneva, Switzerland: WHO. (2016). pp. 1–392. Available at: https://apps.who.int/iris/bitstream/handle/10665/204523/9789241565219_eng.pdf?sequence=1

27. Covidence. Covidence systematic review software. *Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia*. Covidence (2022). www.covidence.org (Accessed January 1, 2022).

28. Ampt FH, Lim MSC, Agius PA, L'Engle K, Manguro G, Gichuki C, et al. Effect of a mobile phone intervention for female sex workers on unintended pregnancy in Kenya (WHISPER or SHOUT): a cluster-randomised controlled trial. *Lancet Glob Heal*. (2020) 8(12):e1534–45. doi: 10.1016/S2214-109X(20)30389-2

29. Dev R, Woods NF, Unger JA, Kinuthia J, Matemo D, Farid S, et al. Acceptability, feasibility and utility of a mobile health family planning decision aid for postpartum women in Kenya. *Reprod Health*. (2019) 16(97):1–11. doi: 10.1186/s12978-019-0767-9

30. Flax VL, Ibrahim AU, Negerie M, Yakubu D, Leatherman S, Bentley ME. Group cell phones are feasible and acceptable for promoting optimal breastfeeding practices in a women's Microcredit program in Nigeria. *Matern Child Nutr*. (2017) 13(1 PG-e12261):e12261. doi: 10.1111/mcn.12261

31. Harrington EK, McCoy EE, Drake AL, Matemo D, John-Stewart G, Kinuthia J, et al. Engaging men in an mHealth approach to support postpartum family planning among couples in Kenya: a qualitative study. *Reprod Health*. (2019) 16:17. doi: 10.1186/s12978-019-0669-x (PG-(11-February 2019)):(11-February 2019).

32. Harrington EK, McCoy EE, Drake AL, Matemo D, John-Stewart G, Kinuthia J, et al. Engaging men in an mHealth approach to support postpartum family planning among couples in Kenya: a qualitative study. *Reprod Health*. (2019) 16(1):1–11. doi: 10.1186/s12978-019-0669-x

33. Isler J, Sawadogo NH, Harling G, Bärnighausen T, Adam M, Sié A, et al. "If he sees it with his own eyes, he will understand": how gender informed the content and delivery of a maternal nutrition intervention in Burkina Faso. *Health Policy Plan*. (2020) 35(5 PG-536–545):536–45. doi: 10.1093/heapol/czaa012

34. Lund S, Hemed M, Nielsen BB, Said A, Said K, Makungu MH, et al. Mobile phones as a health communication tool to improve skilled attendance at delivery in Zanzibar: a cluster-randomised controlled trial. *BJOG An Int J Obstet Gynaecol*. (2012) 119(10):1256–64. doi: 10.1111/j.1471-0528.2012.03413.x

35. Onono M, Odhiambo GO, Congo O, Waguma LW, Serem T, Owenga MA, et al. Narratives of women using a 24-hour ride-hailing transport system to increase access and utilization of maternal and newborn health services in rural western Kenya: a qualitative study. *Am J Trop Med Hyg*. (2019) 101(5 PG-1000–1008):1000–8. doi: 10.4269/ajtmh.19-0132

36. Parkes-Ratanshi R, Mbazira Kimeze J, Nakku-Joloba E, Hamill MM, Namawejje M, Kiragga A, et al. Low male partner attendance after syphilis screening in pregnant women leads to worse birth outcomes: the syphilis treatment of partners (STOP) randomised control trial. *Sex Health*. (2020) 17(3):214–22. doi: 10.1071/SH19092

37. Schwartz SR, Clouse K, Yende N, Rie AV, Basset J, Ratshefola M, et al. Acceptability and feasibility of a mobile phone-based case management intervention to retain mothers and infants from an option B+ program in postpartum HIV care Sheree. *Matern Child Heal J*. (2015) 19(9):2029–37. doi: 10.1007/s10995-015-1715-0

38. Skinner D, Delobelle P, Pappin M, Pieterse D, Esterhuizen TM, Barron P, et al. User assessments and the use of information from MomConnect, a mobile phone text-based information service, by pregnant women and new mothers in South Africa. *BMJ Glob Heal*. (2018) 3(Suppl. 2 PG-e000561):e000561. doi: 10.1136/bmjgh-2017-000561

39. Trafford Z, Swartz A, Jewett S, Barron P, Winch PJ, LeFevre AE, et al. Reported infant feeding practices and contextual influences on breastfeeding: qualitative interviews with women registered to MomConnect in three South African provinces. *Int Breastfeed J*. (2020) 15(1 PG-81):81. doi: 10.1186/s13006-020-00315-7

40. Hong QN, Pluye P, Fabregues S, Bartlett G, Boardman F, Cargo M, et al. *Mixed methods appraisal tool (MMAT) version 2018*. Canada: Canadian Intellectual Property Office; Copyright (#1148552) (2018). 1–11.

41. Pace R, Pluye P, Bartlett G, Macaulay AC, Salsberg J, Jagosh J, et al. Testing the reliability and efficiency of the pilot mixed methods appraisal tool (MMAT) for systematic mixed studies review. *Int J Nurs Stud*. (2012) 49(1):47–53. doi: 10.1016/j.ijnurstu.2011.07.002

42. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Res Methodol*. (2008) 8:1–10. doi: 10.1186/1471-2288-8-45

43. Fàbregues S, Molina-Azorín JF. Addressing quality in mixed methods research: a review and recommendations for a future agenda. *Qual Quant*. (2017) 51(6):2847–63. doi: 10.1007/s11135-016-0449-4

44. De Vet HCW, De Bie RA, Van Der Heijden GJMG, Verhagen AP, Sijpkes P, Knipschild PG. Systematic reviews on the basis of methodological criteria. *Physiotherapy*. (1997) 83(6):284–9. doi: 10.1016/S0031-9406(05)66175-5

45. Renwick L, Pedley R, Johnson I, Bell V, Lovell K, Bee P, et al. Mental health literacy in children and adolescents in low- and middle-income countries: a mixed studies systematic review and narrative synthesis. *Eur Child Adolesc Psychiatry*. (2022) (0123456789):1–25. doi: 10.1007/s00787-022-01997-6

46. Probst P, Grummich K, Heger P, Zaschke S, Knebel P, Ulrich A, et al. Blinding in randomized controlled trials in general and abdominal surgery: protocol for a systematic review and empirical study. *Syst Rev*. (2016) 5(1):1–6. doi: 10.1186/s13643-016-0226-4

47. Stern C, Lizarondo L, Carrier J, Godfrey C, Rieger K, Kirkpatrick P, et al. Methodological guidance for the conduct of mixed methods systematic reviews. *JBI Methodol*. (2020) 19(2):120–9. doi: 10.1097/XEB.0000000000000282

48. Mwangome M, Prentice A, Plugge E, Nweneka C. Determinants of appropriate child health and nutrition practices among women in rural Gambia. *J Heal Popul Nutr*. (2010) 28(2):167–72. doi: 10.3329/jhpn.v28i2.4887

49. Corker J. "Ligne verte" toll-free hotline : using cell phones to increase access to family planning information in the democratic Republic of Congo peer-reviewed case study. *Public Heal Commun Mark*. (2010) IV:23–37. Available at: www.casesjournal.org/volume4

50. Pembe AB, Mbekenga CK, Olsson P, Darj E. Why do women not adhere to advice on maternal referral in rural Tanzania? Narratives of women and their family members. *Glob Health Action*. (2017) 10(1):1364888. doi: 10.1080/16549716.2017.1364888

51. Abajobir A, de Groot R, Wainaina C, Njeri A, Maina D, Njoki S, et al. The impact of i-PUSH on maternal and child health care utilization, health outcomes, and financial protection: study protocol for a cluster randomized controlled trial based on financial and health diaries data. *Trials*. (2021) 22(1):1–13. doi: 10.1186/s13063-021-05598-7

52. UNICEF. *A review of gender-transformative approaches and promising practices in health, nutrition and HIV programming in Africa*. New York, NY: UNICEF (2022). Available at: https://www.unicef.org/media/127671/file/Assessment%Report%(English).pdf

53. Jennings L, Omoni A, Akerele A, Ibrahim Y, Ekanem E. Disparities in mobile phone access and maternal health service utilization in Nigeria: a population-based survey. *Int J Med Inform*. (2015) 84(5):341–8. doi: 10.1016/j.ijmedinf.2015.01.016

54. Hackett KM, Kazemi M, Sellen DW. Social science & medicine keeping secrets in the cloud : mobile phones, data security and privacy within the context of pregnancy and childbirth in Tanzania. *Soc Sci Med*. (2018) 211(June):190–7. doi: 10.1016/j.socscimed.2018.06.014

55. Clouse K, Phillips TK, Noholoza S, Mogoba P, Myer L, Camlin C, et al. Careconekta: study protocol for a randomized controlled trial of a mobile health intervention to improve engagement in postpartum HIV care in South Africa. *Trials*. (2020) 21(1 PG-258):258. doi: 10.1186/s13063-020-4190-x

56. WHO. *Global strategy on digital health 2020–2025*. Geneva, Switzerland: World Health Organization (2021). Available at: http://apps.who.int/bookorders (cited 2023 Aug 29).

57. UN women. *Power on: how we can supercharge an equitable digital future*. UN Women (2023). p. 1. Available at: https://www.unwomen.org/en/news-stories/explainer/2023/02/power-on-how-we-can-supercharge-an-equitable-digital-future (Accessed August 29, 2023).

58. Flynn-Dapaah K, Rashid AT. Gender digital equality in ICT interventions in health: evidence from IDRC supported projects in developing countries. *J Commun Inform*. (2008) 5(3):2019. Available at: http://ci-journal.net/index.php/ciej/article/view/526/512

59. Osaki H, Sao SS, Kisigo GA, Coleman JN, Mwamba RN, Renju J, et al. Male engagement guidelines in antenatal care: unintended consequences for pregnant women in Tanzania. *BMC Pregnancy Childbirth*. (2021) 21(1):1–10. doi: 10.1186/s12884-021-04141-5

60. Heymann J, Levy JK, Bose B, Ríos-Salas V, Mekonen Y, Swaminathan H, et al. Improving health with programmatic, legal, and policy approaches to reduce gender inequality and change restrictive gender norms. *Lancet*. (2019) 393(10190):2522–34. doi: 10.1016/S0140-6736(19)30656-7

61. Fisher J, Makleff S. Advances in gender-transformative approaches to health promotion. *Annu Rev Public Health*. (2022) 43:1–17. doi: 10.1146/annurev-publhealth-121019-053834

62. Kalenga CZ, Parsons Leigh J, Griffith J, Wolf DC, Dumanski SM, Desjarlais A, et al. Sex and gender considerations in health research: a trainee and allied research personnel perspective. *Humanit Soc Sci Commun*. (2020) 7(1):1–7. doi: 10.1057/s41599-020-00643-3

63. Day S, Mason R, Tannenbaum C, Rochon PA. Essential metrics for assessing sex & gender integration in health research proposals involving human participants. *PLoS One*. (2017) 12(8):1–13. doi: 10.1371/journal.pone.0182812

frontiers | Frontiers in Digital Health

TYPE Study Protocol
PUBLISHED 30 November 2023
DOI 10.3389/fdgth.2023.1159744

OPEN ACCESS

EDITED BY
Nikki Holliday,
Coventry University, United Kingdom

REVIEWED BY
Eugenio Santoro,
Mario Negri Institute for Pharmacological Research (IRCCS), Italy
Edith Talina Luhanga,
Carnegie Mellon University Africa, Rwanda

*CORRESPONDENCE
Simona Bertoli
✉ simona.bertoli@unimi.it

†These authors have contributed equally to this work

RECEIVED 27 March 2023
ACCEPTED 13 November 2023
PUBLISHED 30 November 2023

CITATION
Castelnuovo G, Capodaglio P, De Amicis R, Gilardini L, Mambrini SP, Pietrabissa G, Cavaggioni L, Piazzolla G, Galeone C, Garavaglia G, Bertoli S and the DEMETRA Study Group (2023) Study protocol of a clinical randomized controlled trial on the efficacy of an innovative Digital thErapy to proMote wEighT loss in patients with obesity by incReasing their Adherence to treatment: the DEMETRA study.
Front. Digit. Health 5:1159744.
doi: 10.3389/fdgth.2023.1159744

Study protocol of a clinical randomized controlled trial on the efficacy of an innovative Digital thErapy to proMote wEighT loss in patients with obesity by incReasing their Adherence to treatment: the DEMETRA study

Gianluca Castelnuovo[1,2†], Paolo Capodaglio[3,4†], Ramona De Amicis[5,6], Luisa Gilardini[†6], Sara Paola Mambrini[5,7], Giada Pietrabissa[1,2], Luca Cavaggioni[6,8], Giuseppina Piazzolla[9], Carlotta Galeone[10,11], Giacomo Garavaglia[12], Simona Bertoli[5,6]* and the DEMETRA Study Group

[1]Clinical Psychology Lab, IRCCS Istituto Auxologico Italiano, Milan, Italy, [2]Department of Psychology, Catholic University of Milan, Milan, Italy, [3]Orthopaedic Rehabilitation Unit and Research Lab in Biomechanics, Rehabilitation and Ergonomics, IRCCS Istituto Auxologico Italiano, Piancavallo, Italy, [4]Department of Surgical Sciences, Physical and Rehabilitation Medicine, University of Turin, Turin, Italy, [5]International Center for the Assessment of Nutritional Status and the Development of Dietary Intervention Strategies (ICANS-DIS), Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, Milan, Italy, [6]Obesity Unit and Laboratory of Nutrition and Obesity Research, Department of Endocrine and Metabolic Diseases, IRCCS Istituto Auxologico Italiano, Milan, Italy, [7]Division of Nutritional Rehabilitation, IRCCS Istituto Auxologico Italiano, Piancavallo, Italy, [8]Department of Biomedical Sciences for Health, Università degli Studi di Milano, Milan, Italy, [9]Interdisciplinary Department of Medicine, School of Medicine, University of Bari "Aldo Moro", Bari, Italy, [10]Bicocca Applied Statistics Center (B-ASC), Università degli Studi di Milano-Bicocca, Milan, Italy, [11]Biostatistics & Outcome Research, Statinfo, Milan, Italy, [12]Advice Pharma Group S.r.l., Information Technology Unit, Milan, Italy

Despite the increasing importance of innovative medications and bariatric surgery for the treatment of obesity, lifestyle interventions (diet and physical activity) remain the first-line therapy for this disease. The use of digital devices in healthcare aims to respond to the patient's needs, in order to make obesity treatment more accessible, so our study aims to assess the safety and efficacy of a Digital Therapy for Obesity App (DTxO) for achieving weight loss and its maintenance in patients affected with obesity undergoing an experimental non-pharmacological treatment. Here we present the study protocol of a prospective, multicenter, pragmatic, randomized, double-arm, placebo-controlled, parallel, single-blind study on obese patients who will be treated with a new digital therapy to obtain an improvement in their disease condition through the application of different simultaneous strategies (a dietary regimen and personalized advice program, a tailored physical exercise program, a cognitive–behavioural assessment and program, alerts and reminders, dedicated section on prescribed drugs intake, and chat and online visits with clinical professionals). We believe that DTxO will offer a promising intervention channel and self-regulation tool holding the potentiality to decrease treatment burden and treat more patients thanks to the partial replacement of traditional medical consultation with digital or telephone management, improving self-engagement and reducing the high demands the "obesity pandemic" for both patients and national health services in terms of time, cost, and effort.

Clinical trial registration: clinicaltrials.gov, identifier, NCT05394779.

KEYWORDS
obesity, digital therapy, weight loss, diet, physical activity, cognitive–behaviour therapy, randomized control trial

1. Introduction

Obesity is defined as a multifactorial disease characterized by an excess of body fat which is associated with a significantly higher risk of multiple chronic diseases, for example diabetes mellitus, cardiovascular disease (CVD), depressive disorder, and different types of neoplasia (1–3). Moreover, obesity is associated with significant impairments in patients' quality of life and even a reduction of life expectancy ranging from 5 to 20 years (4, 5).

Its prevalence has been globally increasing in the last 50 years to the extent that it is nowadays defined as an "obesity pandemic", with the Global Burden of Disease (GDB) recently reporting that almost one third of world population can be listed as overweight or obese (5). Also in Italy obesity prevalence is rising, increasing by almost 30% in adult population in the last 30 years (6). Despite the increasing employment of medications and bariatric surgery for the treatment of obesity (7), lifestyle interventions (hypocaloric diet, physical activity, and cognitive-behavioral therapy) remain the first-line treatment (8). However, long term weight loss maintenance can be difficult both for metabolic readjustment and for the arduousness in lifestyle changes compliance, and patients often suffer weight regain following initial weight loss (9–11). Moreover, although drastic lifestyle changes usually lead to clinically significant and durable weight loss, they often require multiple in-person sessions with high demands for patients and national health services in terms of time, cost, and effort (12). Therefore, behavioral weight loss still has the unsolved issue of how to reduce healthcare intensity not affecting the regular social support, reliability, and feedback essentials to improve and preserve the "patient empowerment". This recent concept has been proposed for managing diabetes (13) and is based on the concept that the patient is more encouraged to follow self-chosen behavioral adjustments than changes prescribed by others. Self-education with extensive lifestyle modifications and systematic self-monitoring provides positive and long-term effects on metabolic profile, together with quality of life, knowledge, and healthy behavior improvements (14).

It is in this context that digital health tries to respond to the patient's needs and makes the management of this condition more accessible (15). Scientific research has investigated different options of weight loss programs (computer, TV programs, smartphone applications), in order to meet patient's necessities and make obesity treatment more available (15). Smartphones indeed can be a valid option becoming a self-education device holding the capacity to reduce healthcare's workload and treat more people through the partial replacement of traditional medical consultation with digital or telephone management (16, 17). This approach seems the most convenient and accessible, considering the widespread use of smartphones among the population and the fact that weight-loss apps (Apps) can help reinforce dietary plan compliance, physical activity, and weight monitoring with the expedient of progress awareness through goals' achievement (18, 19). Consistent with these findings, our study aims to assess the safety and performance of a Digital Therapy for Obesity App (DTxO) for weight loss and its maintenance in patients affected with obesity undergoing an experimental non-pharmacological treatment. The primary objective is to evaluate the weight loss in patients using DTXO compared to control patients (Placebo App patients DTXO after a period of 6 and 12 months of DTXO use. The secondary objectives are:

1. To assess the degree of dietary recommendation adherence (Mediterranean dietary pattern questionnaire) and of improvement of food and nutrition knowledge (Food Knowledge Moynihan Questionnaire) of the two groups during our study
2. To evaluate the reduction of physical inactivity (reduction) (IPAQ Questionnaire), lower limbs physical function improvements (SPPB test) randomized in the two randomized group.
3. To measure the reduction of stress (DASS-21 scale), the increase in goodwill to approach therapy (URICA-28 scale), the increase in the (various) self-esteem in coping with life's challenges (GSES-10), the measure of Binge Eating Behaviour (BES-16 scale), the measure of emotion regulation problem (DERS-36 scale), the increase of quality of life (PGWBI questionnaire), the measure of sleep habits (PSQI) in each randomized group
4. To identify the level of acceptability in the use of DTxO after 12 months

2. Material and methods

2.1. Study design

This is a prospective, multicenter, pragmatic, randomized, double-arm, placebo-controlled, parallel, single-blind study on obese patients treated with a new digital therapy (DTX Obesity) to obtain an improvement in their disease condition through the application of different simultaneous strategies (a dietary regimen and personalized advice program, a tailored physical exercise schedule, a cognitive-behavioural assessment and program, alerts and reminders, dedicated section on prescribed drugs intake, and chat and online visits with clinical professionals) (see **Figure 1**).

2.2. Study setting

The clinical centers involved in the trial are Istituto Auxologico Italiano and Azienda Ospedaliero-Universitaria

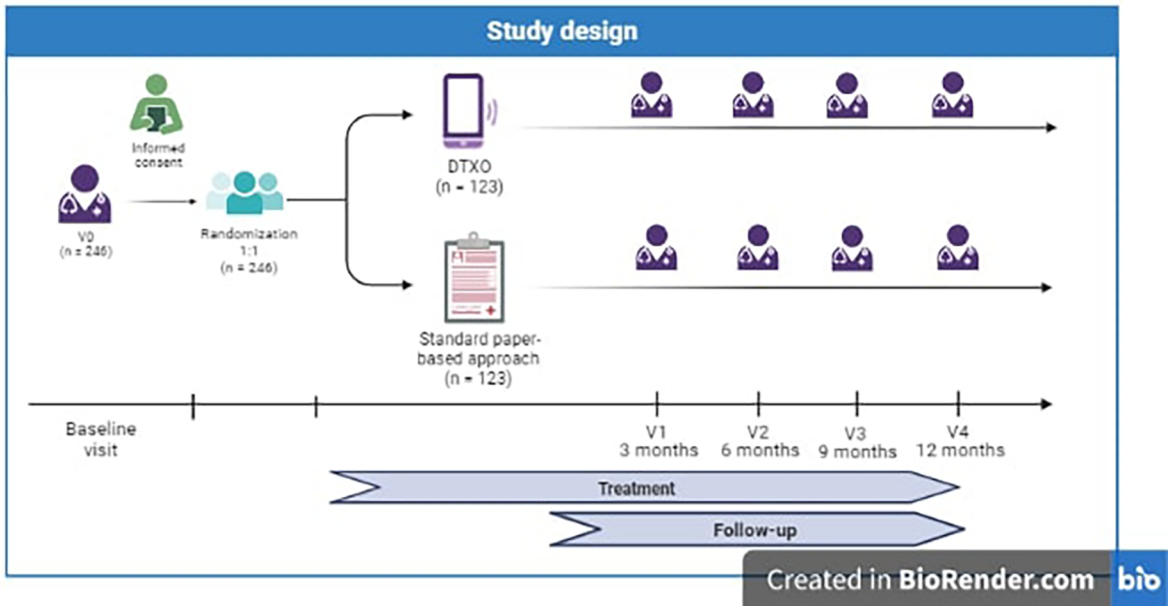


FIGURE 1
Study design of the DEMETRA study.

Consorziale Policlinico Di Bari, specialized centers in obesity care. Both centers fully take part in trial activities, in the enrolling and take in charge of participants, implementation of interventions, and assessment of results. Heterogeneity is guaranteed both by the high number of patients, all enrolled in a comprehensive obesity outpatient clinic, and by the fact that the two centers are respectively in North and South of Italy, areas with different lifestyle and food habits.

2.3. Study duration

The estimated study duration is 16 months. Recruitment will last approximately 4 months. Each patient will be followed for approximately 12 months after the baseline visit.

2.4. Eligibility criteria

2.4.1. Inclusion criteria

- Participants must be between 18 and 65 years of age at the time of informed consent.
- Informed consent form (Capability of giving informed consent, including compliance to requirements and restrictions listed in).
- BMI between 30.0 kg/m2 and 45 kg/m2.
- Participants must own, be prone to use technology of mobile Apps (tech-savvy) and be able to download the App described in the protocol.
- Italian language native speakers or foreign patients who fully understand Italian language, as the information and instructions for health programs will be in Italian.

2.4.2. Exclusion criteria

- Diseases: recent cardiovascular event (<6 months), heart failure (class >II), GFR <60 ml/min, type 1 diabetes, cancer within the first 5 years; drug or alcohol abuse, active eating disorder or history of bulimia and anorexia nervosa, Binge Eating Scale score >27, psychiatric disorders not compensated, score ≥2 in the "Depression," "Anxiety" and "Psychoticism" subscales of the Symptom Checklist-90-R; visual or vision impairments; secondary obesity, Edmonton Obesity Staging System score >4; bariatric surgery in the previous 2 years; referred pain to lower limb joints on the Numeric Rating Scale ≥5; recent (<6 months) weight loss ≥10%.
- Prior/Concomitant Therapies and other exclusions: Changed in pharmacological therapy with drugs that may affect eating behavior and/or energy metabolism during the last 3–6 months; participation in other weight-loss programs or in other clinical trials; known child-bearing women; any condition could the results of the study. Potentially interfering with the results of the study according to the investigator.

The presence of inclusion criteria and absence of exclusion criteria is investigated during the visit.

2.5. Recruitment

This study will include consecutive obese adult patients who self-referred to obesity centers seeking weight-loss treatment at the centers involved in the trial: patient enrolment will occur during a visit with an endocrinologist and a dietitian. Before any screening procedure, all patients must sign a written informed consent for the study.

2.6. Training

The training and experience necessary to use the device subject to pre-market clinical investigation (a device with pending CE marking and used in the investigation according to its intended use) were evaluated during the risk assessment of the device. It has been shown that the provision of information for safe use (User Manual) is sufficient for the healthcare professional user. Additionally, specific onboarding procedures for the patient have been determined as per the onboarding experience. Users will also be instructed to report any software malfunctions to the assistance service by emailing a dedicated assistance portal.

2.7. Randomization

Patients will be randomly assigned 1:1 to the intervention or comparison group according to a pre- defined, centralized randomization list. Randomization is indeed globally guaranteed between the two centers. To maintain an overall balancing between groups, we will perform block randomization (using random block sizes of 8 patients) ("plan").

2.8. Blinding and unblinding procedures

Patients in the control arm will be provided with a placebo App to maintain patients' blindness to treatment allocation. Breaking the blind during the study (for a single subject) should be considered only when knowledge of the treatment assignment is deemed essential by the investigator due to immediate safety concerns or is considered essential for the immediate management of the subject, and should be discussed with the sponsor beforehand, if possible. It is the responsibility of the investigator to promptly document and explain any unblinding to the sponsor. Any unblinding will result in the discontinuation of subject participation from the study.

2.9. Study intervention

DTxO is an investigational therapeutic intervention for obese patients, under clinical validation with a randomized and placebo-controlled clinical trial for confirmatory purposes. DTxO is downloadable software for mobile devices (smartphone, tablet), classified as Class IIa Medical Device (MD), available for patients as an application (App) that integrates different non-pharmacological approaches. It is intended to improve weight loss, weight-loss maintenance, and overall health in patients with obesity by increasing their self-engagement, self-monitoring and adherence to dietary/exercise and behavioral programs. The App integrates different non-pharmacological approaches, engaging the patient through monitoring of her/his/them non-vital parameters, monitoring of patient diet and exercise, monitoring of patient psychological status, prescription of exercise and diet in a weight-loss program, and configurable data presentation charts for provision of additional information to professional users.

The Investigational Medical Device (IMD) consists of a Medical Web Application, also called Medical Dashboard, integrated into the electronic Case Report Form (eCRF), where the physician will collect patient's data necessary for enrolment and profile creation; and a Patient App, which will allow interacting with the patient to guide their behaviors, record therapeutic outcomes and improve therapeutic adherence.

The DTxO App will provide a dietary regimen with a personalized advice program, a tailored physical exercise program, a cognitive–behavioural assessment and support program, alerts and reminders, dedicated section on prescribed drugs intake and dietary and exercise program, chat and online visits with clinical professionals, and trophies to improve patient engagement. All these tools aim not only to weight loss, but they also help patients to improve their health and ameliorate their lifestyle.

The user downloads the App, creates their own account and starts to use it at baseline visit, by answering to psychological questionnaires and inserting food preferences. Thereafter, the app is synchronized with physician platform, so that a dietary regimen and a physical activity program is developed. The user then, helped by alerts and reminders, should daily use the app for consulting and filling in dietary, physical activity, and cognitive-behavioural sections, follow thee app's advice, be encouraged and self-monitor the situation. The App sends feedback and changes its advices, customizing them on the patient and going hand in hand with the constant use by the patient. The experimental arm will be compared with a control arm in which dietary and exercise programs will be delivered through a standard paper-based approach. The control arm will also be equipped with a placebo App, not customized on the habits and feedback of the patient, with no alerts and reminders nor chat or online visits.

2.10. Measurements and procedures

2.10.1. Baseline visit (V0)

During the first visit, the patient will sign the informed consent form and will be evaluated for inclusion/exclusion criteria. Once the enrolment is confirmed, the patient is randomized 1:1 to either the experimental arm (DTx Obesity App) or control arm. The physician will explain to the patient how to download, access, and use the assigned App.

2.10.2. Follow-up visit

After the baseline visit, the patient will follow the dietary, physical and psycho- behavioural programs suggested by the physician and will use the DTXO/placebo App as explained by the physician. Face-to-face follow-up visits are foreseen at 3 months (V1), 6 months (V2), 9 months (V3) and 12 months (V4) after enrolment. During each visit, the following assessments are performed:

- Physical examination
- Assessment of vital signs
- Laboratory blood tests
- Short Physical Performance Battery

- Assessment of patient's responses to psycho-behavioural questionnaires by a psychologist. If any psycho-pathological problem is detected, the patient is referred to a specialist and exits the study.
- Recording of concomitant medications
- Safety assessment (adverse events investigation)

2.10.3. Telemedicine/remote contact

The Telemedicine visit and the remote contact between the patient and the physician will occur in specific situations related to weight- loss changes. This remote contact will allow the physician to understand the entity of the patient's symptoms and to manage them promptly. These utilities are part of the digital therapy: together with DTXO App, they will help patients to achieve treatment goals and assist them. To guarantee a proper assistance to all patients, control group was provided with email address and telephone number of enrolment outpatient clinic in case of need.

2.11. Study control

The efficacy of the DTXO App will be assessed with the comparison of results with a group of patients treated with a standard of care approach (paper-based, in-person) and a placebo App.

The placebo App will include only the non-medical modules of the digital therapy, for example, patient onboarding data and access, while it will not include any medical module; for example, the digital placebo App will not allow the input of the patient's reported clinical data nor provide indications or reminders about the therapy or diagnosis.

The placebo App has been introduced to make the experience of subjects in the control arm more similar to that of subjects in the experimental arm; moreover, data collection through the placebo app (questionnaires) will streamline the data collection process, avoiding manual data entry and reducing the possibility of entry errors. Through the placebo App, the patient will be asked to complete weekly the Diet Adherence questionnaire (NRS scale). No trend graph will be provided to the patient to monitor their performance, nor any personalized content will be displayed. standard clinical practice. The placebo App will be used only as a data entry tool for questionnaires.

Patients in the control group will be provided with a dietary plan and nutritional education and behavioral contents in paper format.

2.12. Statistical analysis

2.12.1. Calculation of sample numerosity

Considering the primary endpoint, an 80% power and a two-sided type I error of 0.05 would be reached with a global sample size of 172 patients (86 per group) in order to be able to measure a 1.5 kg weight change after 6 months between the two groups, assuming an SD of 3.5 kg and a drop-out rate of 30%. We hence established a target of 246 patients considering a drop-out rate of 30%.

2.12.2. Data analysis

We will perform descriptive analyses (at each visit) by computing categorical variables (percentages and frequencies) and continuous variables (mean and median values, standard deviations, quartiles, and extreme values). If useful, descriptive data will also be examined graphically through histograms and box-and-whisker plots. Appropriate statistical methods for comparisons between- group (i.e., interventional vs. control group) and within-group will be applied according to the nature of the variables (continuous or categorical variables), as specified in the following sections. % CI will not be finalized before the. Data analyses will be conducted using SAS version 9.4 (SAS Institute, Cary, NC, USA) statistical software.

3. Results

The advantage of the study is the possibility to provide simultaneous treatment strategies to improve the health status of obese patients. We expect that the concurrent application of these strategies and the possibility of customizing patients' activities based on their habits and feedback will allow greater benefits than the application of separate, no customized programs and a standard paper-based in- person approach. Both treatment arms will receive potential benefits by participating in this study.

The control arm will receive the current standard of care (paper-based approach and face-to-face follow-up visits) for obese patients with the addition of a placebo app. However, greater benefits are hypothesized in the experimental arm due to the possibility of customization of physical activity in accordance with the perceptions and feedbacks of the patient, the availability of notifications and reminders for drug intake (dosage, frequency) to improve treatment adherence, the availability of contents to support and favor compliance with the prescribed diet, the possibility to schedule an online visit or communicate with the clinical specialist via chat through the App, and the availability of multimedia psycho-behavioral content (motivational exercises, self-acceptance exercises, mindfulness exercises, interactive emotional eating exercises, self-efficacy exercise).

Differences between the DTxO and the App placebo are reported in **Table 1**.

4. Discussion

Considering the "obesity pandemic", the benefits of lifestyle changing and the widespread use of smartphones among the population, the DTxO offers a promising intervention channel and self- regulation tool holding the potentiality to decrease treatment burden and treat more patients thanks to the partial replacement of traditional medical consultation with digital or telephone management. This approach seems the most convenient and accessible, since weight-loss apps (Apps) can help reinforce dietary plan compliance, physical activity, and weight monitoring with the expedient of progress awareness through goals' achievement (18, 19).

TABLE 1 Differences between the dTxO and the App placebo.

	DTXO	App placebo
Diet		
Dietary plan and personalized suggestions' alerts	X	
Dietary questionnaires		
Mediterranean-style eating questionnaire	X	X
Questionnaire food knowledge Moynihan questionnaire	X	X
Dietary program self- assessment		
Diet adherence (NRS scale)	X	X
Diet enjoyment (NRS scale)	X	
Hunger/satiety scale (NRS scale)	X	
Physical activity		
Physical activity plan	X	
Physical activity program self- assessment		
Level of physical activity questionnaire (IPAQ)	X	X
Fatigue perceived (NRS scale), Pain perceived (NRS scale)	X	X
Psycho-behavioural assignment		
Psychobehavioural self- assessment	X	X
DASS-21, URICA-28, GSES-10, BES-16, DERS-36, PGWBI, PSQI		
Recording of concomitant medications	X	
Telemedicine/remote contact	X	

Different randomized controlled trials investigated the effects of digital health tools in weight-loss interventions in overweight and obese patients, showing the potential efficacy of self-monitoring of diet and physical activity with mobile phones combined with behavioural counselling or a newly developed, multifactorial, and daily-based personalized program for cognitive–behavioural therapy (CBT) conducted by a psychologist (20–22). One study also showed that the use of the digital health care service together with the support of a therapist online led to better results in anthropometric measures, for example body composition and body weight, and physiological indices and obesity-related psychological factors compared to the use of the app combined with self-care and no therapist intervention, thus suggesting that programs based on technological devices should be multidimensional and are undoubtedly most effective in combination with human support (23). Interestingly, literature data suggest that subjects using Apps related to health to record diet, physical activity, and health-related behaviours, gain significantly better results when the three dependent variables (i.e., lifestyle, physical activity, and eating behaviour) are combined compared with non-users (24).

Nevertheless, it was clear that app-based recording devices can be very appealing, considering the ascertained relationship between lifestyle and chronic diseases, since they can be very effective in preventative healthcare measures, increasing self-engagement, self- monitoring and adherence to dietary/exercise and behavioral programs. The App integrates different non-pharmacological approaches, engaging the patient through monitoring of her/his nonvital parameters, monitoring of patient diet and exercise, monitoring of patient psychological status, prescription of exercise and diet in a weight-loss program, and configurable data presentation charts for provision of additional information to professional users. Apps can be considered promising, since they can be harnessed for monitoring and motivation of patients to persevere in a healthy lifestyle (24).

Ethics statement

This study was approved by the Ethical Committee of the Istituto Auxologico Italiano (research code: 2022_05_17_22; acronym: DEMETRA) and the Azienda Ospedaliero-Universitaria Consorziale Policlinico Di Bari (research code: 2022_10_07_7392). An informed consent from all the participants involved in our study will be collected.

Author contributions

Conceptualization, SB; methodology, SB GC, PC, RD; formal analysis, CG; investigation, LG, DEMETRA Group Study; writing–original draft preparation, RD; writing–review and editing, all authors; supervision, SB; project administration, SB; funding acquisition, SB. All authors contributed to the article and approved the submitted version.

Members of the DEMETRA study group

Amalia Bruno, Antonina Orlando, Alessio Genovese, Federica Sileo, Marta Pellizzari, Marina Croci, Silvia Martinelli.

Funding

This research is funded by Advice Pharma Group S.r.l., and Theras Group. The funders were not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

Conflict of interest

GG was employed by the company Advice Pharma Group. CG was employed by Statinfo.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Kahn SE, Hull RL, Utzschneider KM. Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature*. (2006) 444:840–6. doi: 10.1038/nature05482

2. Neeland IJ, Poirier P, Després J-P. Cardiovascular and metabolic heterogeneity of obesity. *Circulation*. (2018) 137:1391–406. doi: 10.1161/circulationaha.117.029617

3. Smith SC. Multiple risk factors for cardiovascular disease and diabetes Mellitus. *Am J Med*. (2007) 120:3–11. doi: 10.1016/j.amjmed.2007.01.002

4. Taylor VH, Forhan M, Vigod SN, McIntyre RS, Morrison KM. The impact of obesity on quality of life. *Best Pract Res Clin Endocrinol Metab*. (2013) 27:139–46. doi: 10.1016/j.beem.2013.04.004

5. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the global burden of disease study 2013. *Lancet*. (2014) 384:766–81. doi: 10.1016/S0140-6736(14)60460-8

6. d'Errico M, Pavlova M, Spandonaro F. The economic burden of obesity in Italy: a cost- of-illness study. *Eur J Heal Econ*. (2022) 23:177–92. doi: 10.1007/s10198-021-01358-1

7. Pi-Sunyer X, Astrup A, Fujioka K, Greenway F, Halpern A, Krempf M, et al. A randomized, controlled trial of 3.0 mg of liraglutide in weight management. *N Engl J Med*. (2015) 373:11–22. doi: 10.1056/nejmoa1411892

8. Yumuk V, Tsigos C, Fried M, Schindler K, Busetto L, Micic D, et al. European guidelines for obesity management in adults. *Obes Facts*. (2015) 8:402–24. doi: 10.1159/000442721

9. Leibel RL, Seeley RJ, Darsow T, Berg EG, Smith SR, Ratner R. Biologic responses to weight loss and weight regain: report from an American diabetes association research symposium. *Diabetes*. (2015) 64:2299–309. doi: 10.2337/db15-0004

10. Lemstra M, Bird Y, Nwankwo C, Rogers M, Moraros J. Weight loss intervention adherence and factors promoting adherence: a meta-analysis. *Patient Prefer Adherence*. (2016) 10:1547–59. doi: 10.2147/PPA.S103649

11. Kroeger CM, Hoddy KK, Varady KA. Impact of weight regain on metabolic disease risk: a review of human trials. *J Obes*. (2014) 2014. doi: 10.1155/2014/614519

12. Steele RM, Mummery WK, Dwyer T. A comparison of face-to-face or internet-delivered physical activity intervention on targeted determinants. *Heal Educ Behav*. (2009) 36:1051–64. doi: 10.1177/1090198109335802

13. Wong CKH, Wong WCW, Lam CLK, Wan YF, Wong WHT, Chung KL, et al. Effects of patient empowerment programme (PEP) on clinical outcomes and health service utilization in type 2 diabetes mellitus in primary care: an observational matched cohort study. *PLoS One*. (2014) 9:e95328. doi: 10.1371/journal.pone.0095328

14. Trento M, Gamba S, Gentile L, Grassi G, Miselli V, Morone G, et al. Rethink organization to IMprove education and outcomes (ROMEO): a multicenter randomized trial of lifestyle intervention by group care to manage type 2 diabetes. *Diabetes Care*. (2010) 33:745–7. doi: 10.2337/dc09-2024

15. Klasnja P, Pratt W. Healthcare in the pocket. *J Biomed Interv*. (2012) 45:1–7. doi: 10.1016/j.jbi.2011.08.017.Healthcare

16. Coons MJ, DeMott A, Buscemi J, Duncan JM, Pellegrini CA, Steglitz J, et al. Technology interventions to curb obesity: a systematic review of the current literature. *Curr Cardiovasc Risk Rep*. (2012) 6:120–34. doi: 10.1007/s12170-012-0222-8

17. Papazian L, Forel J-M, Gacouin A, Penot-Ragon C, Perrin G, Loundou A, et al. Neuromuscular blockers in early acute respiratory distress syndrome. *N Engl J Med*. (2010) 363(12):1107–16. doi: 10.1056/NEJMoa1005372

18. Spring B, Schneider K, McFadden HG, Vaughn J, Kozak AT, Smith M, et al. Multiple behavior changes in diet and activity: a randomized controlled trial using mobile technology. *Arch Intern Med*. (2012) 172:789–96. doi: 10.1001/archinternmed.2012.1044

19. Spring B, Duncan JM, Psy D, Janke EA, Andrea T, Mcfadden HG, et al. *NIH Public Access*. (2013) 173:105–11. doi: 10.1001/jamainternmed.2013.1221. Integrating

20. Turner-McGrievy GM, Beets MW, Moore JB, Kaczynski AT, Barr-Anderson DJ, Tate DF. Comparison of traditional versus mobile app self-monitoring of physical activity and dietary intake among overweight adults participating in an MHealth weight loss program. *J Am Med Informatics Assoc*. (2013) 20:513–8. doi: 10.1136/amiajnl-2012-001510

21. Allen JK, Stephens J, Dennison Himmelfarb CR, Stewart KJ, Hauck S. Randomized controlled pilot study testing use of smartphone technology for obesity treatment. *J Obes*. (2013) 2013. doi: 10.1155/2013/151597

22. Spring B, Pellegrini C, McFadden HG, Pfammatter AF, Stump TK, Siddique J, et al. Multicomponent MHealth intervention for large, sustained change in multiple diet and activity risk behaviors: the make better choices 2 randomized controlled trial. *J Med Internet Res*. (2018) 20:1–13. doi: 10.2196/10528

23. Kim M, Kim Y, Go Y, Lee S, Na M, Lee Y, Choi S, Choi H.J. Multidimensional cognitive behavioral therapy for obesity applied by psychologists using a digital platform: open-label randomized controlled trial. *J MIR Mhealth Uhealth*. (2020) 8 (4):e14817. doi: 10.2196/14817

24. Sarcona A, Kovacs L, Wright J, Williams C. Differences in eating behavior, physical activity, and health-related lifestyle choices between users and nonusers of Mobile health apps. *Am J Heal Educ*. (2017) 48:298–305. doi: 10.1080/19325037.2017.1335630

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 08 December 2023
DOI 10.3389/fpubh.2023.1284566

OPEN ACCESS

EDITED BY
Nikki Holliday,
Coventry University, United Kingdom

REVIEWED BY
Shweta Pathak,
University of North Carolina at Chapel Hill, United States
Carlos Laranjeira,
Polytechnic Institute of Leiria, Portugal

*CORRESPONDENCE
Jacob Cookey
✉ Jacob.Cookey@virtualhallway.ca

RECEIVED 29 August 2023
ACCEPTED 16 November 2023
PUBLISHED 08 December 2023

CITATION
Tennankore K, Jones J, Miller A, Adib A, Mathew S, Rasic D and Cookey J (2023) Transforming healthcare delivery: a descriptive study of a novel provider-to-provider virtual care platform.
Front. Public Health 11:1284566.
doi: 10.3389/fpubh.2023.1284566

Transforming healthcare delivery: a descriptive study of a novel provider-to-provider virtual care platform

Karthik Tennankore[1], Jennifer Jones[2], Ashley Miller[3], Ashfaq Adib[4], Shan Mathew[5], Daniel Rasic[4] and Jacob Cookey[4]*

[1]Division of Nephrology, QEII - Dickson Building, Halifax, NS, Canada, [2]Division of Digestive Care and Endoscopy, QEII - Victoria Building, Halifax, NS, Canada, [3]Division of General Internal Medicine, QEII - Bethune Building, Halifax, NS, Canada, [4]Virtual Hallway Consults Inc., Halifax, NS, Canada, [5]Pleasant Street Medical Group, Dartmouth, NS, Canada

Introduction: Addressing challenges in access to specialty care, particularly long wait times and geographic disparities, is a pressing issue in the Canadian healthcare system. This study aimed to evaluate the impact and feasibility of provider-to-provider phone consultations between primary care providers (PCPs) and specialists using a novel virtual care platform in Nova Scotia (Virtual Hallway).

Methods: We conducted a cross-sectional survey over 5 months, involving 211 PCPs and 34 specialists across Nova Scotia. The survey assessed the need for formal in-person referrals as well as clinician satisfaction. Statistical methods included descriptive statistics and the one-sample *t*-test.

Results: We found that 84% of provider-to-provider phone consultations negated the need for an in-person specialist referral. It was also reported that 90% of patients that did require in-person consultation had enhanced care while they awaited an in-person appointment with a specialist. Very high levels of satisfaction were reported among both PCPs and specialists, and there was a noticeable increase in billing volumes related to these consultations as measured by provincial billing codes.

Conclusion: The findings indicate that provider-to-provider phone consultations are feasible, well-accepted and also effective in reducing the need for in-person specialist visits. This approach offers a promising avenue for alleviating waitlist burdens, enhancing the quality of care, and improving the overall efficiency of healthcare delivery.

KEYWORDS

virtual care solutions, Hallway medicine, healthcare solution, digital healthcare, peer to peer communications, healthcare technologies, synchronous communication technology, eConsult

Introduction

Access to specialty care remains a pivotal concern for healthcare systems around the world. This paper explores an innovative approach to enhancing this access through synchronous provider-to-provider communication between primary care providers and specialists, with a focus on the Canadian healthcare context. The central contribution of this study is to provide

an empirical analysis of how digital platforms, specifically using a platform called Virtual Hallway, can streamline the referral process, reduce wait times, and democratize access to specialty healthcare services.

Background

The Canadian healthcare system faces substantial hurdles in specialty care provision, with patients enduring lengthy wait times that span an average of 26 weeks, though this can vary between provinces (1). The crux of the issue lies not only in these protracted delays but also in the uneven distribution of healthcare specialists, who are predominantly located in urban areas. This geographic inequality necessitates often burdensome travel for rural inhabitants and disproportionately impacts those with limited resources (2). Adverse health and economic consequences are well-documented, with delays leading to severe outcomes such as increased morbidity and mortality (3, 4).

One potential solution to improve access to specialty care is to reduce the barriers to engaging in peer-to-peer consultation between primary care providers (PCPs) and specialists. Primary care physicians and nurse practitioners serve as the entry point for patients to access healthcare services, including specialty care. In most situations, a patient must be referred by a PCP to access specialty care. Referrals must be reviewed for appropriateness and completeness, triaged to determine urgency, and then patient appointments are subsequently booked based on information provided in a referral. However, it has been observed that many formal referrals and specialist-patient consultations could be avoided if the PCP could consult directly with the specialist, thereby reducing wait times and potentially improving patient outcomes (5).

Electronic consultations (eConsults) can facilitate PCPs seeking specialist advice digitally, in writing, and without the need for an in-person referral (6). Another promising avenue is the utilization of synchronous peer-to-peer consultations, whereby the PCP arranges a phone consultation with a specialist, receives verbal advice, and subsequently implements the recommended care plan. This approach has the potential to optimize access to specialty care and reduce waiting times for patients who do not require formal specialist consultation through specialist-PCP phone consults.

To date, the evidence for peer-to-peer communication in medical care has been limited by heterogeneity (i.e., differences in program type, outcome measures) as well as a paucity of studies specifically focusing on synchronous provider-to-provider phone consults (7).

Purpose

This descriptive study sought to answer the following questions: how do primary care providers (PCPs) and specialists utilize Virtual Hallway (a novel phone consultation platform) for peer-to-peer communication? What impact does the platform have on in person referrals? What impact does it have on billing code usage? Is this novel virtual care platform satisfactory?

Methods

Setting

The pilot project was conducted in partnership with Nova Scotia Health and the Coordinated Accessible National (CAN) Health Network over a five-month period starting in May 2022. Nova Scotia has a population of just over 1 million people spread across approximately 55, 284 square kilometers. The participants involved nurse practitioners, family medicine specialists, and medical specialists. Physicians (both referring and consulting) have been eligible for compensation for synchronous provider-to-provider consultation since April 2017 as described in the Nova Scotia Medical Services Insurance Interim Fee Guide.[1]

Virtual Hallway platform

Virtual Hallway is an online platform that facilitates provider-to-provider patient-focused virtual consultation via synchronous telephone conversations. To initiate a phone consult request, a requesting provider (usually a primary care provider but occasionally a specialist) logs onto the Virtual Hallway system and completes an electronic form for a patient-specific question, with an option to attach any relevant patient documents (e.g., laboratory results, images). Primary care providers submitting the request for consult can book a phone consultation with a specific specialist of their choosing. The service is offered at no cost to patients and providers, and fee-for-service specialists and family physicians are reimbursed using existing provincial billing codes. The encounter consists of a brief phone call (typically about 10 min) between the providers that occurs on a date and time specified by each provider. At the conclusion of each phone consult, the specialist completes a consult report summarizing the advice given. The platform complies with all applicable Canadian healthcare privacy legislation including PIPEDA and provincial privacy acts (PHIA, PHIPA, HIA, etc.). It also secures patient data using Medstack which is a data security compliance platform for digital health applications that adheres to all standard healthcare security frameworks including: HIPAA, SOC2 and ISO 27001.

Study design

This was a cross-sectional survey of all healthcare providers using the Virtual Hallway platform from July 14 to November 10, 2022 to determine the acceptability and feasibility of peer-to-peer phone consultations among healthcare providers.

A population-based sampling approach was used. This study was conducted in Nova Scotia, Canada between July 14, 2022 and November 10, 2022. This study conforms to the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for reporting cross-sectional, observational studies.

1 https://msi.medavie.bluecross.ca/wp-content/uploads/sites/3/2023/08/Interim-Fee-Reference-Guide.pdf

The survey was developed using an iterative participatory design including the investigators and users of the platform to ensure it aligned with the needs of both the clinician users and the local health system (8). The questions were designed to assess the acceptability of the platform as well as the need for and quality of subsequent in-person referral. At the close of each consultation, the specialist and the requesting provider were requested to complete closeout surveys embedded within the Virtual Hallway platform related to the experience of the completed consult. The questions are found in supplement A.

Recruitment and respondent characteristics

Participants were recruited using a convenience sample of the entire population of interest. All providers on the Virtual Hallway platform interested in conducting phone consults and licensed to practice medicine in the province of Nova Scotia were eligible to create an account in the system and participate in synchronous provider-to-provider virtual phone consults. Physicians, NPs, and specialists were informed of the service through a combination of email communications, fax communication, and general information available on the website. All providers that were already registered at the time of study initiation were able to complete the questionnaire each time they completed a phone advice call. There were a limited number of volunteer participants which mitigated potential risks from this research and data collection ensured participant privacy by de-identifying the information collected. An exemption letter was obtained from the Nova Scotia Health Research Ethics Board as a program evaluation study in accordance with Chapter 2 of the Tri-Council Policy Statement guidelines.

Data collection and analysis

The digital questionnaire was administered through the platform at the end of each virtual consultation. Characteristics of respondents were described using appropriate univariable statistical approaches for continuous and categorical data. Likert scale responses to survey questions were reported as medians with interquartile ranges and the proportion responding to each category were also described.

Questions such as, level of satisfaction with consult experience were captured in a 5-point Likert-scale ranked from 1 (strongly disagree) to 5 (strongly agree), with 3 being neutral. These responses were analyzed by computing descriptive statistics (mean, median, mode and standard deviation). Mean Likert scale values were interpreted as the overall agreement toward a variable in the questionnaire. To determine the statistical significance of the difference of the mean from the neutral value, we performed the one-sample *t*-test (alpha level = 0.05).

To analyze dichotomous variable responses (yes/no), such as referral avoidance due to phone consultation, we measured the frequency distribution of the overall responses and by specialty groups to determine response variance over different specialties.

Results

Response rate and respondent characteristics

Two-hundred and eleven PCPs from across Nova Scotia participated in the pilot study by completing at least one or more consults along with the associated closeout survey(s). This accounted for approximately 15% (211/1357) of the active PCPs in the province. The survey response rate by PCPs was 81.8%, whereas for specialists the response rate was 72.1%. We received 654 closeout survey responses from specialists; for data completeness we excluded incomplete responses, leaving a total 632 responses. There were 34 specialists who participated across 17 specialty areas (mean number of cases per provider = 17.7, sd 27.3 range 2–156). We received 614 closeout survey responses from PCPs. There were 608 after excluding incomplete responses. A total of 181 PCPs participated (mean number of cases per provider = 3.33, sd 4.8, range 1–45).

The most consulted specialties were general internal medicine (39%) and psychiatry (18%), followed by rheumatology (6%) and obstetrics and gynecology (5%) (Figure 1).

Impact on need for formal consultation

PCP survey results indicated that 84% (511/608) of phone consults resulted in avoidance of the need for an in-person referral (Table 1). Among the remaining 16% of phone consults that did not eliminate the need for a referral, the PCP respondents indicated that none (0%) of those were initially intended to avoid an in-person referral. PCPs indicated that for 87% of the phone advice calls that ended up requiring a formal consultation, that the phone advice improved the quality of the in-person referral. Ninety percent of PCP consultations were found to have enhanced the patient's care while they awaited an in-person appointment with a specialist.

Stratified analysis by specialty revealed variability in the percentage of cases avoiding a formal in-person consultation. PCPs reported the avoidance of formal in-person specialist consultation in 93.7% (194/207) of phone advice calls with internal medicine followed by 87.8% (101/115) with psychiatry—with at least 40% of in-person referrals being avoided across all specialties. All responses by PCPs on avoiding a formal in-person consultation by specialty is provided in Table 2.

A total of 608 responses were received from specialists for the post-phone consult survey (Table 3). Specialists reported that 42% of the cases reviewed would have been at least somewhat necessary for an in-person consultation had the referral been made through traditional in-person routes. In contrast, in-person consultations were deemed unnecessary for 58% of the cases.

User satisfaction

Nearly all PCPs reported a high level of satisfaction with their consult experience, with 99% indicating they were either "very satisfied" or "satisfied" (Figure 2) (mean 4.93, sd 0.29, $p < 0.01$). Ninety-six percent of specialists were either "very satisfied" or

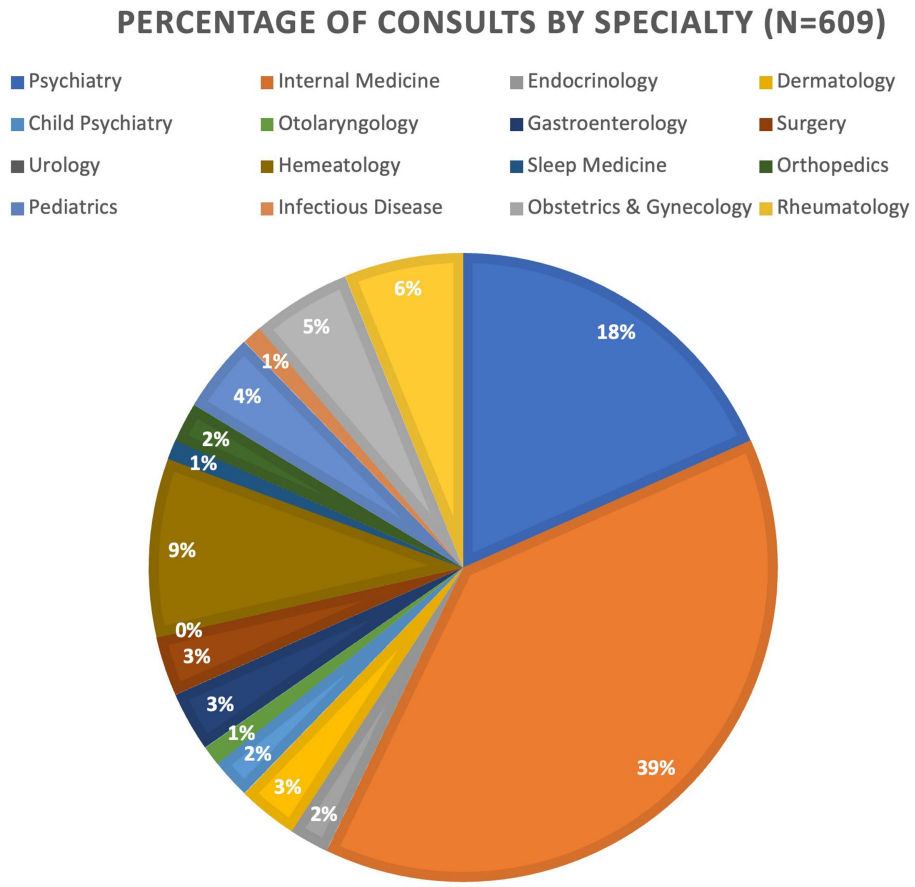


FIGURE 1
Percentage of total consults by specialty during pilot study.

TABLE 1 Result of the closeout survey of referring physicians and nurse practitioners on closing the phone consult.

Did this Virtual Hallway consultation avoid the need for an in-person referral?	
	Cases reviewed between [2022-07-14] and [2022-11-22] n = 608
Yes	84%
No	16%
Subset of survey to those referring physicians and practitioners that answered No. (n = 97)	
Was this consultation intended to avoid referral?	
Yes	0%
No	100%
Did this consult improve the quality of your referral?	
Yes	87%
No	13%
Did this consultation improve the patient's care while they wait for an in-person referral?	
Yes	90%
No	10%

"satisfied" with the consult experience (Figure 3) (mean 4.75, sd 0.54, $p < 0.01$).

Billing volumes

The number of synchronous provider-to-provider virtual consultations are reflected in billing volumes for codes 03.09 K (billed by specialist provider completing consultation) and 03.09 L (billed by referring provider requesting consultation). Although volumes have increased progressively since the introduction of these billing codes in 2017, there is a notable deflection in volumes from 2021 to 2022 which correlates to the time period of this study (Figure 4).

Discussion

This study examined the acceptability and impact of synchronous provider-to-provider communication for improving specialty care access within the Canadian healthcare landscape using a novel digital platform. The aims of the study were to determine to what extent synchronous provider-to-provider consultations reduce the need for in-person specialist appointments, determine the acceptability of the platform, and to describe utilization patterns of these consults by healthcare providers.

10.3389/fpubh.2023.1284566

TABLE 2 Result of the closeout survey of PCP consultants on closing the phone consult by specialty of the consultation.

Did this Virtual Hallway consultation avoid the need for an in-person referral?		
	Total consults	Answer yes: *n* (%)
Internal Medicine	207	194 (93.72%)
Psychiatry	115	101 (87.83%)
Infectious Disease	8	7 (87.5%)
Endocrinology	42	36 (85.71%)
Rheumatology	34	27 (79.41%)
Hematology	43	34 (79.07%)
Pediatrics	19	15 (78.95%)
Dermatology	18	13 (72.22%)
Obstetrics and Gynecology	37	26 (70.27%)
General Surgery	13	9 (69.23%)
Gastroenterology	19	13 (68.42%)
Child and Adolescent Psychiatry	12	8 (66.67%)
Orthopedics	12	8 (66.67%)
Otolaryngology	8	5 (62.5%)
Urology	6	3 (50%)
Pain Medicine	5	2 (40%)

TABLE 3 Result of the closeout survey of specialist consultants on closing the phone consult.

If this patient had been referred directly to your clinic, would the referral have been:	
Answer	Cases reviewed between [2022-07-14] and [2022-11-10] (*n* = 608)
Necessary	26%
Somewhat necessary	16%
Somewhat unnecessary	10%
Unnecessary	48%

Our findings demonstrate a higher-than-average physician response rate and that phone consultations significantly reduce the need for in-person referrals—with higher referral avoidance compared to other modalities of peer-to-peer communication found in the literature (e.g., eConsults). We also found that even those consults that go on to in-person referrals are enhanced through phone consultation. Of note, there was discrepancy found between PCPs and specialists regarding perceived need for in-person referrals following phone consults, although there were uniformly high satisfaction ratings for these phone interactions.

The primary finding is that PCPs identified that 84% of all synchronous provider-to-provider consultations avoided the need for an in-person referral. In Canada, wait times for specialist appointments are an all-time high (1). As a result, these synchronous provider-to-provider consultations could have significant impacts on alleviating the burden on waitlists and improving access to specialty care. If this finding continues to be replicated, then this model of consultation offers potential time and cost advantages compared to in-person referrals. Consequently, synchronous provider-to-provider consultations may provide a more efficient and affordable approach to specialty care access, reducing the burden on healthcare systems and patients alike (7).

In addition to referral avoidance, this study revealed that in cases where the synchronous consultations did not eliminate the need for an in-person referral, they provided valuable benefits through an improvement in the quality of the subsequent in-person referral or an optimization of patient management while on the waitlist. Interestingly, our study identified a discrepancy between the perceptions of specialists and PCPs regarding the necessity of referrals. Specialists indicated that over half of the phone consults (58%) would have been unnecessary or somewhat unnecessary if they had been referred in-person, whereas PCPs reported that 84% of cases would have been referred for in-person consultations if not for the phone consult. This incongruence highlights the need for improved communication and understanding between specialists and PCPs to optimize the referral process and resource allocation in the healthcare system (9).

The potential to reduce waitlists is a significant advantage of synchronous provider-to-provider consultations. Our findings are consistent with previous research on electronic consultations (eConsults), which have been shown to improve access to specialty care and reduce wait times; with up to 65% of eConsults avoiding the need for in-person specialist referral (6, 10, 11). The current study found potential referral avoidance beyond the upper range of these studies. One possibility for the disparity between these two peer-to-peer consultation methods is that phone consults allow a synchronous dynamic conversation to take place, being able to clarify and ask questions, which may allow a greater scope of consultation. As described from one of the physician focus groups in the Cook et al. (12) study: "I find the value of communicating on the phone because it's two ways, back and forth, and then I get my answers right away" (12).

By decreasing the number of unnecessary in-person referrals, phone consultations could contribute to more efficient resource allocation and streamlined access to specialty care for patients in need.

The satisfaction ratings reported among healthcare providers suggests a high level of acceptability of this model of synchronous provider-to-provider consultations. Provincial billing volumes demonstrated a progressive increase in uptake of provider-to-provider synchronous consultation since compensation was first introduced in 2017. There is a notable deflection point in the volume data from 2021 to 2022, which correlates to the significant expansion of the platform users within the province. The platform supported increased uptake of phone consultation relative to the status quo of unsupported booking, documentation, and billing. Although compensation for this form of care is available independent of the Virtual Hallway platform, there are significant administrative barriers related to booking times when both referring, and specialist, physicians are available. This is one of the major advantages of the platform, as it facilitates physician-to-physician communication without administrative staff, or the complex scheduling previously required among clinics. Additional benefits of the platform include ease of documentation and

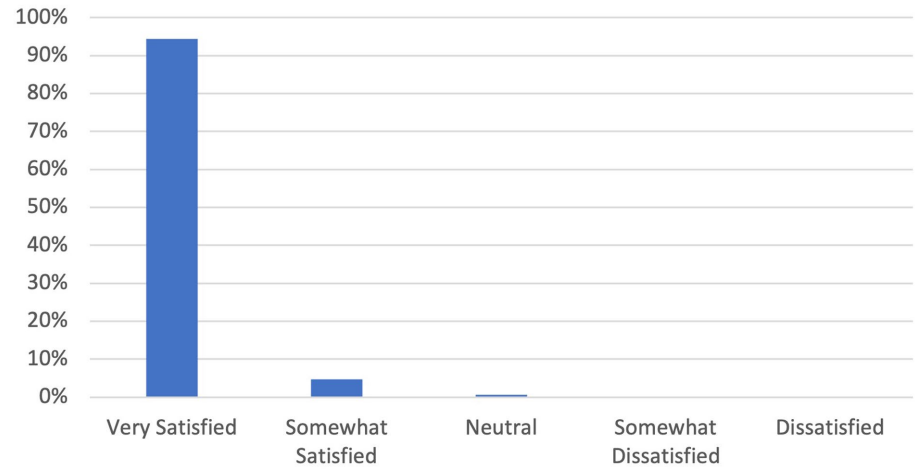


FIGURE 2
PCP satisfaction survey.

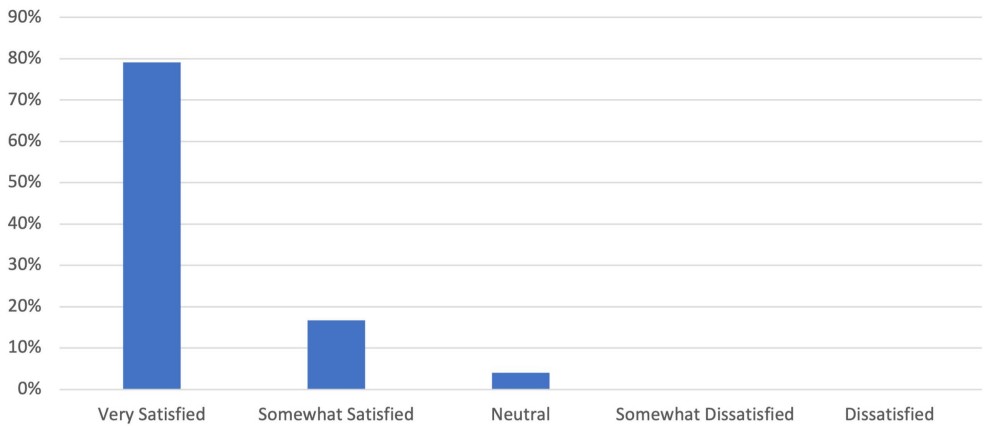


FIGURE 3
Specialist survey.

automated billing functions, all of which minimize the administrative burden associated with care delivery.

One of the strengths of this study was having a high response rate (72% for specialists and 82% for PCPs) given the typically low historical response rates for physician surveys (typically well below 50%) (13, 14). This may indicate that this system, with immediate, real-time feedback, provides a unique and effective way of surveying physicians compared to alternative strategies such as email reminders, financial incentives or even personalization, all of which have uncertain benefits and significant variability in the literature (14). The high satisfaction ratings by both PCPs and specialists also suggests high feasibility for implementation into clinical practice across primary care and specialty areas.

Limitations of our study include the use of a convenience sample and a focus on a single Canadian province, which may limit the generalizability of the results. Additionally, the study did not assess the long-term impact of phone consultations on patient outcomes or healthcare system performance. Another limitation includes the survey questions which may introduce potential bias toward positive responses. Furthermore, this was a pilot study and may not have had a sample size sufficient to detect a true effect and may be subject to selection bias. Results should not be interpreted as providing conclusive evidence, but rather as guidance for future research. Future research should utilize objective measures of healthcare quality, outcome, and utilization to evaluate the effectiveness of phone consultations in other healthcare settings and populations. Comparative studies between phone consults and other forms of peer-to-peer communication should be conducted to understand the relative benefits of different modalities.

In conclusion, this pilot study demonstrates the potential of provider-to-provider synchronous virtual care to address the pressing issues of waitlists and access to specialty care in the Canadian healthcare system and the benefit of a novel digital platform in supporting uptake of this care modality. The findings suggest that phone consultations are well-accepted among healthcare providers and can avoid a significant proportion of in-person referrals.

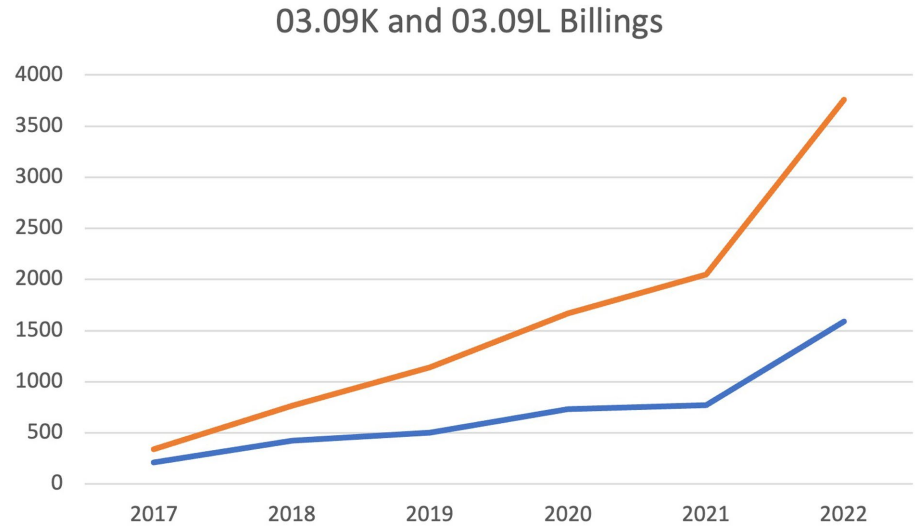


FIGURE 4
Total number of Nova Scotia province-wide 03.09 K and 03.09 L billing submissions by year. Deflection point reflects timepoint when Virtual Hallway was initiated for this study.

Moreover, phone consultations may provide a more time-efficient and cost-effective alternative to in-person referrals, with the potential to reduce waitlists, improve patient outcomes, and enhance the overall quality of care.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Nova Scotia Health Research Ethics Board. The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional review board waived the requirement of written informed consent for participation from the participants or the participants' legal guardians/next of kin because research relies exclusively on secondary use of anonymous information, with the process of data recording and dissemination of results not generating identifiable information.

Author contributions

KT: Data curation, Formal analysis, Writing – original draft, Writing – review & editing. JJ: Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing. AM: Formal analysis, Writing – review & editing. AA: Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. SM: Validation, Writing – review & editing. DR: Conceptualization, Methodology, Writing – original draft, Writing – review & editing. JC: Conceptualization, Methodology, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

The authors would like to thank the Nova Scotia Health Innovation Hub for their support.

Conflict of interest

DR and JC are co-founders of Virtual Hallway Consults Inc., which developed the provider-to-provider platform that is under investigation in this study. AA is currently employed by Virtual Hallway.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Moir Mackenzie, Barua Bacchus. (2022). Waiting your turn wait times for Health Care in Canada, 2022 report. Available at: https://www.fraserinstitute.org/sites/default/files/waiting-your-turn-2022.pdf.

2. Bosco C, Oandasan I. *Review of family medicine within rural and remote Canada: Education, practice, and policy*. Mississauga, ON: College of Family Physicians of Canada. (2016). Available at: http://www.cfpc.ca/ARFM.

3. Bacchus B, Nadeem E, Taylor J. The effect of wait times on mortality in Canada. Fraser Institute. (2014). Available at: https://www.fraserinstitute.org

4. Day B. The consequences of waiting In: *Reducing wait times for Health care: What Canada can learn from theory and international experience, edited by Steven Globerman* (2013). 43–75.

5. Keely E, Liddy C, Afkham A. Utilization, benefits, and impact of an e-consultation service across diverse specialties and Primary care providers. *Telemed E-Health*. (2013) 19:733–8. doi: 10.1089/tmj.2013.0007

6. Liddy C, Moroz I, Afkham A, Keely E. Sustainability of a Primary care-driven eConsult service. *Ann Fam Med*. (2018) 16:120–6. doi: 10.1370/afm.2177

7. Tian PG, Jaminal JR, Harris HS, Chambers T, Alvarado S, Eurich D. Characteristics and outcomes of physician-to-physician telephone consultation programs: Environmental scan. *JMIR Form Res*. (2021) 5:e17672. doi: 10.2196/17672

8. Simonsen J, Hertzum M. Iterative participatory design In: J Simonsen, JO Bærenholdt, M Büscher and JD Scheuer, editors. *Design research: Synergies from interdisciplinary perspectives*. London: Routledge (2010). 16–32.

9. Donohoe MT, Kravitz RL, Wheeler DB, Chandra R, Chen A, Humphries N, et al. Reasons for outpatient referrals from generalists to specialists. *J Gen Intern Med*. (1999) 14:281–6. doi: 10.1046/j.1525-1497.1999.00324.x

10. Joschko J, Keely E, Grant R, Moroz I, Graveline M, Drimer N, et al. Electronic consultation services worldwide: environmental scan. *J Med Internet Res*. (2018) 20:e11112–0. doi: 10.2196/11112

11. Liddy C, Drosinis P, Armstrong CD, Mckellips F, Afkham A, Erin Keely CT, et al. What are the cost savings associated with providing access to specialist care through the Champlain BASE EConsult service? A costing evaluation. *BMJ Open*. (2016) 6:e010920. doi: 10.1136/bmjopen-2015

12. Cook DA, Sorensen KJ, Wilkinson JM. Value and process of curbside consultations in clinical practice: a grounded theory study. *Mayo Clin Proc*. (2014) 89:602–14. doi: 10.1016/j.mayocp.2014.01.015

13. Barnhart BJ, Reddy SG, Arnold GK. Remind me again: physician response to web surveys: the effect of email reminders across 11 opinion survey efforts at the American Board of Internal Medicine from 2017 to 2019. *Eval Health Prof*. (2021) 44:245–59. doi: 10.1177/01632787211019445

14. Cunningham CT, Quan H, Hemmelgarn B, Noseworthy T, Beck CA, Dixon E, et al. Exploring physician specialist response rates to web-based surveys. *BMC Med Res Methodol*. (2015) 15:32. doi: 10.1186/s12874-015-0016-z

frontiers | Frontiers in Digital Health

TYPE Original Research
PUBLISHED 22 December 2023
DOI 10.3389/fdgth.2023.1289373

OPEN ACCESS

EDITED BY
Edith Talina Luhanga,
Carnegie Mellon University Africa, Rwanda

REVIEWED BY
Joe Wherton,
University of Oxford, United Kingdom
Makuochi Nkwo,
University of Greenwich, United Kingdom

*CORRESPONDENCE
S. F. Lakha
slakha@torontomu.ca

RECEIVED 05 September 2023
ACCEPTED 05 December 2023
PUBLISHED 22 December 2023

CITATION
Lakha SF, Sohail SF, Holtzman CB, Akkok ZA, Khandwala A, Suhanic W, Pennefather P and Fels DI (2023) Power of narrative: a case study about documenting private insightful experiences while dealing with pain and associated disability.
Front. Digit. Health 5:1289373.
doi: 10.3389/fdgth.2023.1289373

Power of narrative: a case study about documenting private insightful experiences while dealing with pain and associated disability

S. F. Lakha[1,2]*, S. F. Sohail[1], C. B. Holtzman[1], Z. A. Akkok[1], A. Khandwala[3], W. Suhanic[4], P. Pennefather[2,4] and D. I. Fels[1]

[1]Inclusive Media and Design Centre, Ted Rogers School of information Technology Management, Toronto Metropolitan University, Toronto, ON, Canada, [2]Institute of Medical Sciences, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada, [3]School of Administrative Studies, York University, Toronto, ON, Canada, [4]gDial Inc., Toronto, ON, Canada

Objective: People adjusting to living with a chronic disability, such as chronic pain, seek support and resources from societal systems, including health systems, to help them cope with this reality. This case study describes the use of a digital health platform designed to help in that quest.
Method: MyHealthMyRecord (MHMR), is being developed to record, register and curate personal private experiences of a chronic condition. MHMR allows users to record and log short (30–90s) personal and private audio-videos of their accommodation-seeking journey in a way that can be encrypted, registered, curated and shared privately. This case study describes the use of a prototype version of the platform by a participant co-designer who experienced a sudden onset of a chronic pain condition, of undetermined origin. System use began three months after the onset of the condition and just after being discharged from several months of hospitalization without any definitive diagnosis.
Result: During a three-month period, 65 short unstructured contributions were authored and logged. This paper presents a qualitative analysis of that content. The clips used various communication styles that documented experiences, concerns, issues, positive and negative interactions and pain episodes. Using thematic analysis with open coding, three domains (person-facing, accessibility and system-facing) and eight themes (pain, joy, therapy, environmental, recommendations, technical, culture and communication) were identified. Comments about pain, stress, etc., were the most common and occurred in 75% of all videos while technical and therapy/physio related comments were the fewest and occurred in 3 and 9% of the videos, respectively.
Conclusion: We conclude that it is possible to create recordings of events, thoughts, reflections and issues on different aspects affecting an individual's health and well-being impact, including effects of the chronic condition as well as tangential outcomes such as accessibility (or lack of it), using MHMR over a longer period of time. The next steps will be to develop functionality to annotate the recordings, automatically analyze and summarize collections of recordings to make them consumable, useful and understandable to the individual and others, and then to share those analyses and summaries with others. In addition, evaluate this functionality longitudinally with more users.

KEYWORDS

chronic pain, audio-video recordings, digital data, personal health record, qualitative evidence

Introduction

People who suffer the sudden onset of a disabling chronic condition experience a shock to their routine and a disruption of their life-course, as well as to their medical and social situations. They need support in learning to cope and engage with making sense of their new reality (1). That support can take the form of therapeutic/clinical medications and casual/formal social practices involving both professional support from counsellors and support workers and social support from a broader circle of care. Living with chronic conditions typically involves active negotiation of episodically disabling situations (2) in a manner that requires active engagement by the affected person rather than just their passive acceptance of recommended coping strategies and simple compliance with imposed therapeutic/rehabilitation interventions (3). That engagement requires communication of evidence of private, personal needs and wants to require accommodations in order to advocate for accommodations that often are mandated by law or institutional policies (3, 4).

Increasingly, that communication involves digital technology. However, human factors considerations such as diverse orientations and false expectations have been shown to limit positive effects from implementing digital health technology for supporting person-centred approaches and self-advocacy (4). Person-centred approaches to healthcare are intended to promote agency, engagement, empowerment and self-efficacy in the person receiving the care (1, 5). We propose that having a patient register their experiences with care, technologies, and interventions can lead to the development of important person-centred healthcare aims of autonomy and independence as discussed by (6).

We are developing a system for documenting, registering, and sharing human factors such as agency, engagement, empowerment and self-efficacy by users of healthcare, rehabilitation and accessibility programs. This is carried out through enabling the generation of a personally owned and curated digital repository ("scrapbook") of short video recordings of personal observations concerning an individual's care/recovery/accommodation journey following the sudden onset of a chronic condition. We believe that the video nature of this record will make it more accessible and meaningful to the author and those with whom they decide to share their record(s). It will also set the stage for creating a corpus of voice/video evidence for directing care that are generated and owned by patients and are ready for regulated precision reporting regimes.

The MyHealthMyRecord (MHMR) platform is designed to document, register and share a person's private experiences with a chronic condition for which they seek care, support, and accommodations (7, 8). The idea is to allow them to assert ownership over person-centered evidence used to guide interventions designed to assist them in coping with their condition.

The aim of this case study was to explore how digital scrapbooking of short audio-video (AV) commentaries may offer an accessible and inclusive method for recording and anchoring person-centred documentation of their personal healthcare journeys. The MHMR platform described here is based on video guestbook technology developed earlier to help people living with disabilities to comment on hospitality services they experienced (9). MHMR is being designed around the concept of curating a digital scrapbook made up of an unstructured collection of short-duration, first-person AV's that register commentary on concerns, issues, successes, and failures of care and support the user has experienced. Although designed first to be a tool for self-reflection, we are confident that with time, user-curated repositories of such AVs can be used to create a registered representation of the patient's voice useful more generally in advocating for their needs and wants as people with lived/living experiences, and those of others like them. MHMR technology is designed to highlight the user's experience, perspective, and identity as these relate to the care support and accommodations that they access for their chronic condition. The short format makes them amenable to quick review and analysis by both human and machine entities trained to recognize patterns pointing to available solutions. MHMR provides a counter-narrative to the official, system-serving record of care. In this paper we illustrate how the well-known method of thematic analysis of the qualitative MHMR content can be used to structure the unstructured person-reported observations (PROs).

This paper illustrates how the MHMR technology can allow the person receiving support and care for a chronic health condition to express his or her needs, ideas, experiences, concerns, and coping strategies in an easily generated and accessible manner. Such evidence is now recognized as vital to patient self-management necessary for effective community-based care (10). In addition, this paper comments on how such data can fill a need to accommodate human factors such as diverse orientations and false expectations that have been shown to limit positive effects from implementing approaches that seek to register patient experiences (4).

The research questions addressed in this paper are thus: (1) what is the feasibility of recording short-duration first-person perspective narrative reflections on condition-associated challenges? and, (2) what are the experiences, expressions, and attitudes expressed over time using the platform by a participant co-designer whose life was disrupted after the sudden onset chronic pain of undefined causes? In order to preserve the privacy and anonymity of the participant, gender pronouns are alternated in this paper, and no institutional names are provided.

Background

This literature review provides a brief overview of the need for people-centred applied science and design, highlighting the mostly unexplored potential of qualitative information to enhance communication between individuals with illness and health professionals. It examines individual engagement and empowerment through curated health data to facilitate sharing with health professionals and care circles.

Individuals with chronic illnesses must manage their conditions while also engaging with the healthcare system. While programs exist to assist in medical management and holistic care, there is an apparent paucity of resources and research addressing

the thorough tracking, integration, and communication of activities individuals perform to cope with their illnesses between healthcare specialist appointments (11). For example, someone may use naturopathic and medical treatments, but this is not necessarily conveyed during appointments due to reasons such as memory, significance, priorities, embarrassment, and readiness to discuss at the time of their visit.

Despite extensive advocacy for person/patient-centred care by governments, healthcare organisations, and individuals, its implementation has often proven limited, unsustainable, or unsuccessful (12). Additionally, the absence of a reliable and trusted method for registering, recording and tracking patient perspective within the medical record so as to demonstrate benefits of patient involvement in setting care priorities persists. While considerable efforts have been made to measure quantitative patient data, these processes are primarily driven by clinicians' needs (13), but there is a high chance of data abandonment rates (14). While clinicians acknowledge the usefulness of person-centred care, there is no standardized methodology for its implementation or integration. For instance, patients may gain access to their institutional health record items and lab results (15), but this access does not extend to documenting a patient's complete care journey (16).

Ongoing progress and likelihood of future innovations enabled by existing natural language processing system and, data science, promise to enhance communication between patients and healthcare providers in ways that promote voice based sensemaking (8). However, effective communication also requires an understanding of cognitive factors, such as the social embodiment on how information is perceived and communicated (17).

Research and design methodologies derived from human factors, co-design (18), and human-computer communication/interaction paradigms provide a pathway to mobilize diverse perspectives and ways of knowing in the co-creation of actionable meaning, aligning with the WHO Framework on integrated people-centered health services (19). The distinction between person-centered and patient-centered healthcare lies in the contrast between building meaning and building function (20). The evaluation of our tool's impact on increasing acceptability, reducing communication barriers, and enhancing trust between healthcare providers and chronic pain patients is central to our research, contributing to the construction of meaning-in-care. However, it is important to consider the willingness and feasibility of successfully producing and capturing qualitative information on a long-term basis. In addition, what topics and ideas constitute these data must also be assessed to determine whether they can be used in a meaningful way to support documentation, self-assessment, reflection and communication.

The Structure-Process-Output framework is derived from the Donnebedian model of clinical activity (reference) that has been used for more than 50 years. It has been shown to help physicians make sense of and construct a self-narrative about how the clinical care structures they interact with to deliver therapeutic processes can lead to desired clinical outcomes (21). Recently, Lakha et al. (22, 23) have proposed a hybridization of the Donnebedian model, with a more generalized and evaluative "theory-of-change" represented as a logic model (24). The idea of the proposed Donnebedian-Logic Hybrid (DLH) evaluation model is to focus on more proximal consequences of structure and process interactions that can be directly appreciated by the reflective evaluator with respect to expected outputs. Accordingly, a "theory-of-change" or reflections on why the change is, or is not, occurring as expected can be documented and reported as evidence within a cause-and-effect framework. This allows all individuals involved in the care process, including patients and their circle of care, to augment the medical record in a manner that is still meaningful for both patients and doctors while still useful for efficient translation of processes into desired outcomes. This will be a necessary improvement in health system record keeping as it transitions from a clinician-centred approach to a more patient-centred one.

The DLH evaluation model is adapted in this paper to design a way for patients/system-users to gather evidence, in the form of short video recordings of opinions and reflections, for making sense of appreciated or disappointing outputs associated with their ongoing interaction with system structures and processes associated with health system services.

The goal of MHMR is to create a patient-owned repository of experiential evidence organized within a Structure-Process-Output framework (SPO) that can guide the interpretation of this patient generated evidence. The platform is being designed to help the user to document observable model elements of patient-centered Structure/Inputs and Process/Activities, and then process them into observable/actionable Outputs.

MHMR system design

The MHMR platform is being developed to support collecting and assessing qualitative data generated by patients. It may also be possible in future development to integrate more quantitative data from the health system that is increasingly accessible to patients through existing patient portals (10) and emerging health data-sharing platforms (25).

The current embodiment of the MHMR platform consists of a tablet-based video recording application that accepts short-duration audio/video user input (**Figure 1**). The videos are then stored on a server for secure transfer, processing, and storage of video records. In the embodiment reported here, the primary functionality is to serve as a digital video scrapbook which documents and bundles a person's experiences with their disabling condition.

As with the video-guestbook technology from which it was derived (9), the MHMR platform was designed to be inclusive of users with different needs (e.g., sign language users) by allowing them to record short duration audio and/or videos on the topic of their choice at their leisure. The aim was to provide users with a safe way to record, not only their physical state but also their emotional state, without sacrificing their independence or privacy. Privacy was protected through data indexing and encryption strategies that are not a focus of this paper. The participant was advised that the content would not be shared with anyone beyond the research teams.

The framework goals guiding this initial design work on an MHMR platform are illustrated in **Figure 1**. Patient Curation of

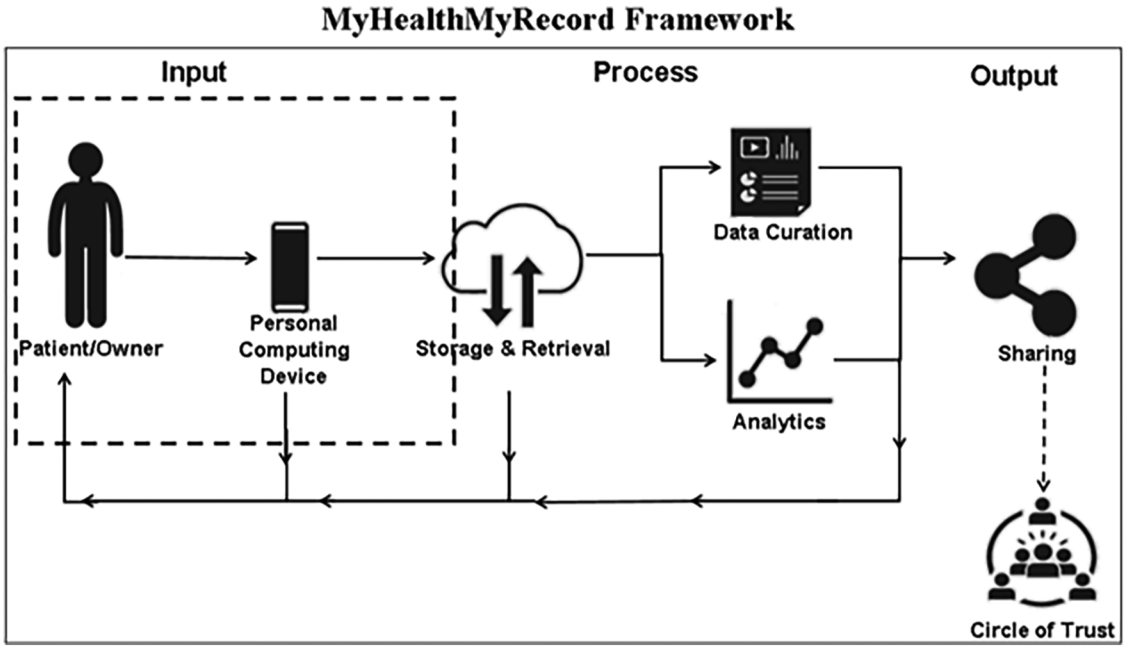


FIGURE 1
Myhealthmyrecord framework.

Private Data (created as short duration first-person videos, which can contain private opinion and reflections-qualitative data, as well as measured vital statistics-quantitative data—that are owned and curated by the patient) leads to Process/Activities: Private Data Driven Option Identification/Evaluation (Patient generated data used by patient with possible input from trusted counsellors including Health Care Providers (HCPs) as evidence to develop knowledge, attitudes, and beliefs concerning care and accommodations options accessible through the patient's circle of care). These Input and Process/Activities may ultimately lead to Output: Acceptable Collaborative Care (supporting bilateral communication and collaboration between patient, HCPs, and circle of care in making decisions around care use options). Our model attempts to illustrate how Inputs from individual experiences can affect Processes and lead to positive and negative Outputs. The dotted line in **Figure 1** represents the focus of this paper.

Methodology

This study was approved by the Toronto Metropolitan University's (formerly, Ryerson University) human research ethics board (Protocol # REB 2016-150). The study followed all necessary procedures to ensure the confidentiality of a participant. Prior to the study, the participant co-designer completed and signed a written informed consent document. Personal information was then gathered.

The methodology employed in this research is a combination of an exploratory case study method (26) and the participatory design method (27, 28). The first stage of the system under investigation is the creation and storage of short-duration videos created by an individual with later stages being to support the communication and knowledge transfer between HCP, circle of care and the individual enabled by the MHMR application. The exploratory component arises as no specific outcome is expected. Rather, we want to explore how the individual will use the system, what topics are addressed and the content of the videos.

It is expected that the trajectory of results will identify meaningful and measurable outputs over time and can be instrumental in generating outcomes as the study proceeds. The research is carried out in natural environments where the participant passed time at home, school (workplace) and social settings. All recordings made by PX are converted into MP3 files and distributed among the research team for coding into themes and major domains.

Participant Co-Designer and System Use (identified as PX in this document with no relationship to the individual's actual initials or gender identity). As a first step in the system design, we recruited a co-designer participant to assist in developing a usable platform. PX played a pivotal role in shaping the MHMR design through active participation in the development process. This encompassed contributions to brainstorming, ongoing feedback, validation of design choices, user-interface problem-solving, and direct involvement in usability testing, all of which enriched the project.

In 2015, PX experienced the sudden onset of an arthritis-like inflammatory condition with mysterious origins. This resulted in multiple motoric disabilities, including limited use of legs and hands. After a 3-month hospitalisation that was unsuccessful in diagnosing a cause and providing a treatment strategy, PX was discharged home to other formal caregivers (physiologists, rheumatologists, etc.) and, with the support of PX's family,

expected to find ways to cope with this new situation. PX's family was known by one of the researchers and was contacted to invite PX to join the project as a participant co-designer. PX agreed to use the MHMR application on a tablet version over a three-month period to help us understand the usability of our initial implementation of the MHMR platform and the extent to which interacting with it was experienced as meaningful and useful. PX began using the system during a 3-week business/pleasure excursion soon after leaving the hospital. Once returned, PX used the system at school for 11 weeks.

In total, PX recorded 65 videos with an average length of 36 s (s) (ranging from 12 to 148 s). Apart from being informed that the ultimate purpose of the technology was to allow people living with chronic conditions to document, reflect and comment upon their condition, PX was not given any direction on topics or times for creating his videos. Nevertheless, she focused on physical pain, difficulties or successes in daily activities, overall frustrations or other events that required attention. During the three-month study, PX provided continuous feedback and ideation to the team via email or in person with unstructured comments of usability issues and on how the system might be improved. During this time, we also noticed that PX stayed engaged in the effort, preventing the project from being abandoned.

Qualitative analysis

The Consensual Qualitative Research (CQR) methodology was selected to analyze the qualitative data as it has been shown to be effective to better understand complex data and reducing bias that can result with a single researcher (29). The CQR updated method proposed by Hill, Thompson and William (29) is an inductive, open-ended research method for investigating unobservable internal experiences carried out through consensus between two invested researchers. It allows researchers to remain open to the data's revelations while focusing on and discussing participant narratives and descriptions together. This method allows for in-depth investigation of occasional events and emergent outcomes, making it useful when knowledge in a specific area is limited. It promotes diverse perspectives in order to have a thorough picture of the phenomenon (29). Additionally, team members were able to discuss disagreements respectfully while agreeing on themes that best captured the meaning of the data. The videos generated by PX were open-coded for relevant and recurring themes (see **Table 1**).

The first step in the analysis was to formulate domains and themes. Two independent reviewers reviewed the videos without communication with the other team members. These reviewers coded blocks of video data (phrases, sentences) into themes, and labelled and defined them. Upon individual completion, all team members rejoined and discussed the themes as peer reviewers. Through discussion and consensual agreement, three domains and eight themes were created to capture all data. **Tables 1**, **2** provide the domain and theme labels, their definitions, and examples from the data of each theme. Furthermore, **Table 1** shows how each domain fits within the SPO framework.

Validity and reliability of analysis

As recommended (30, 31), several quality criteria were used to ensure the validity of the results and their interpretation: two independent coders with different backgrounds coded 20% of the data-set to ensure the reliability of the themes and definitions A Cohen's kappa coefficient was calculated. If kappa was below 0.6, an iterative process was then performed to understand and mitigate the differences. All of the themes and definitions contained in **Table 1** had Cohen's kappa greater than 0.88. The remaining data were then coded by a single rater.

Results and discussion

In general, PX was satisfied with the technology even though there were some technical difficulties associated with the tablet-based implementation. He was able to record short-duration, first-person videos, in several locations, that captured her ideas and topics. In addition, although she recorded videos about numerous topics, these tended to be centred around her pain experiences and related disabilities. There were no videos related to other topics such as family, friends, restaurant reviews, etc., that were independent of his pain or related disabilities despite explicit instruction to use the technology for whatever purposes

TABLE 1 Mapping of domains, themes, and definitions that evolved from PX's video content.

Domain, themes	Definition
Person facing: (outputs)	Physical and mental experiences related to the individual.
Pain, exhaustion, stress, Irritated, frustrated	The individual is in a state of physical and mental pain, distress, and suffering
Joy	Delight or happiness that the individual feels about an experience
Therapy/physio	Training or therapy to help rehabilitate and allow the individual to cope
Accessibility experiences: (processes)	Positive or negative comments and emotions, actions taken by others, or systems related to accessibility.
Environmental	Tangible or concrete situations or factors that affect the individual
Recommendation	The individual expresses their thought process and observations or comments about the incidents
Technical	Comments about technology and the MHMR application
System facing: (structure)	Descriptions/comments about physical, communication and attitudinal experiences and barriers experienced in various settings.
Culture	Ideas, customs, and social behaviour of a society different from that of the individuals
Communication	Professional interactions with people

TABLE 2 Themes and quotations.

Themes	Quotation
Pain, exhaustion, Stress, irritated, frustrated	A "Today was the first time (Physiotherapy) had a 9 am session….worst decision ever, my body was so stiff and I got drained out, it was too difficult for me" B "along with school another thing is very stressful is a recruitment process as it's my final year so it's so much pressure and I have to apply everywhere and have to get the job offer as its a last kick of the can."
Joy	A "visit the conference… Great sets of lectures especially D… it was so amazing, so inspiring, motivational, charismatic and great to listen to him whereas others were fine" B This morning I went swimming……., and it's pretty cool because it was hot water, the hot tub was very relaxing, especially for muscles, …"
Therapy/physio	A "I went to physio yesterday and my physiotherapist was quite impressed by my progress, ……., slow and steady,…"
Environmental	A "The opening ceremony was great stuff and the people were very accommodating…the whole day I was accommodated which was very fun…. I got around pretty nicely……" B "The weather in Denver is horrible in the sense that its the same as Toronto, it snowed a lot, and no one cleaned it up. Snow on the sidewalk the road everywhere so hard to walk I was slipping and off balance, I didn't expect that,."
Recommendation	A "I wanted to talk about my evening class where we are talking about the wheelchair manufacturer industry and my professor kept on putting me on the spot, which was not nice, …….he doesn't have to put me on the spot like what does PX feel like…I didn't get a chance to speak to him but for sure going to follow up to him on email for this."
Technical	A "…This is the third time I am recording this, hopefully, the camera doesn't shut off and this gets posted."
Culture	A "Today is Thursday the learning journey was pretty cool and went to a local school in Hanoi, Vietnam and donated about … bucks, ……, students didn't see outsider (often), all the students were handpicked across Hanoi public school and it was pretty cool,.." B "So Sydney has been treating me great, love it, very accessible, very accommodating, but too many hills and it's tough getting around ….. ……"
Communication	A "So it was a pretty interesting event in yesterday class, so my professor was talking,, explaining and this was in the beginning of class so I respectfully had my hand up for a while then he's like you have your hand up for a while and there's nothing to be discussed, …… and I m just like yeah…' can you just speak louder' and, he said 'I can't that is the loudest I can speak and I said ' I cannot hear you', so he said 'come to the front' pretty much, …I was shocked….., and I said 'but I can't' and he said 'well I guess we're stuck we can't do much about it' " B "So I got a job offer with (company) and they have acceptance party, and the best part about this is that they called me up and asked me if I'm comfortable with the venue, and if not they will change the entire venue…"

seemed appropriate. This reflected perhaps his understanding that the ultimate goal of this research was to create a medium for sharing personal experiences related to the disability between interactions with members of the user's circle of care. In addition, the participant was enthusiastic about the research goal of finding ways of using this evidence to improve the experience of others facing disabilities due to a chronic condition of sudden onset. The videos are unlike the types of public contributions typically seen on social media and were more private and reflective in nature. There were no indications that PX found the process onerous or boring, and use actually began while PX was travelling on an extended multi-week trip. The total number of videos per domain were: Person Facing (43), Accessibility Experience (29) and System Facing (26) for a total number of videos equalling 65. The number of statements and videos containing specific themes were illustrated in **Table 3** (some videos contained more than one theme).

Domain 1: Person Facing:

(1) Pain, Exhaustion, Stress, Irritated, Frustrated: This theme had an overwhelming majority of comments and videos containing expressions from this theme among all of the themes in the person-facing domain. PX described or commented 57 times of the total 82 comments in the person-facing theme (70%) in 25 of 43 (58%) person-facing videos and 49 of the 65 total videos (75%). For example, PX describes the physical pain during a conference: "I just couldn't take it going the whole day. It's very painful." PX expressed and characterize her pain in several negative and positive emotional dimensions such as lack of sleep, anger, frustration, as well as hope or joy (see Theme 2). For example, "I'm not getting enough sleep for the past two-three days and, on top, have pain in the knees and ankles sore…"

TABLE 3 Frequency of statements and appearance in number of videos for each theme.

Domain	Themes	Number of statements (% of comments for domain)	Number of videos containing comments from theme (% of total # of videos)
Person facing (82 total comments)	Pain, exhaustion, stress, irritated, frustrated	57 (70)	49 (75)
	Joy	19 (23)	18 (28)
	Therapy/physio	6 (7)	6 (9)
Accessibility experiences (76 total comments)	Environmental	46 (61)	33 (51)
	Recommendation	28 (37)	28 (43)
	Technical	2 (3)	2 (3)
System facing (52 total comments)	Culture	14 (26)	9 (14)
	Communication	38 (63)	30 (46)

This kind of feeling and emotion characterization is consistent with other studies (32). In addition, it is not surprising that the majority of video scraps are related to this theme. The idea that pain can lead to feelings of frustration, worry, anxiety, and depression seems obvious, particularly if it is of a chronic nature (32, 33). It should be noted that chronic pain can also lead to long-lasting emotional disturbances often referred to as a "secondary pain affect" (34). Negative emotional responses or low mood states, such as anger or depression, can intensify the pain experience (35–37).

In particular, PX showed that he often expressed: feelings of confusion and worry by the pain, and the social unpleasantness of living with pain. Corbett et al (38), also found that people living with pain often express despair related to the lack of empathy by "others." Examples from PX include: "In the evening, I didn't really do much, I was just bored in the room, my roommate went out."

Although PX received no instruction or direction on topics for videos, ones related to pain were dominant. This may indicate that the documentation of pain experiences was a priority for PX or that she was comfortable sharing such videos with the design and development team. Identifying and tracking instances and frequency of pain, exhaustion, stress, irritation or frustration may allow users to reflect on when these emotions occur and examine trends of sustained negative emotional states. This may, in turn, lead to new ideas, solutions or coping strategies.

Keeping pain journals or using pain scales on a long-term basis is usually encouraged by HCPs to document and track pain experiences (39), but often, these are abandoned or seen as not useful to the individual or the HCPs (40). In addition, there can be considerable quantities of data generated, which must then be synthesized and organized into meaningful information. Having more creative and non-text outlets may improve people's ability and interest in long-term documentation and tracking.

(2) Joy: Not all expressions related to pain experiences were negative particularly when PX was hopeful of change although there were considerably fewer expressions of joy than pain. Of the 19 expressions of joy (23% of the total person-facing comments), 9, 7, 3 occurred when he was involved in personal activity, conference and classroom respectively. Prior research on positive expressions of happiness, inspiration and emotional well-being is limited as most of the documentation processes employed in pain have been associated with measuring or document pain exclusively (41) rather than the wide range of emotions that are seen in PX's data. Positive affect, according to numerous theorists, facilitates approach behaviour or continued action (42). From this perspective, experiences of positive affect could prompt individuals to engage more or earlier with their environments and partake in activities.

For example, PX shared his activities,

> "This morning I went swimming with my brother, there was no else in the swimming pool and it's pretty cool because it was hot water, the hot tub was very relaxing, especially for muscles, … I had the first time at the rehab, so this was a good start and relaxing after the whole week of hectic travelling".

> "First day back to school was amazing because the professor teaching us the audit was just amazing, he simplified everything to the take made it like, he called it the grandma language if you can explain it to your grandma then it makes sense otherwise it's not. So it was so fun, (able to engage) it was nice, mind-opening and mind-blowing."

He was able to record that his ability to participate in pool activities, the classroom and the conference demonstrated that an individual's own mental state such as joy, relaxing and positive anticipation could be a factor in the experience of pain. A next step would be to link factors such as the frequency and intensity of negative and positive emotions with pain or related experiences as well as their proximity to each other. For example, how many positive experiences are required to offset the negative emotions elicited from pain, and how does the time interval between positive and negative emotions mitigate the experience of pain and/or associated issues such as long-lasting emotional disturbances?

(3) Therapy /Physio: There were only 6 comments of the 82 person-facing comments (7%) with 9% (6 of 65) of the total number of videos containing comments from this theme. For successful therapeutic encounters, PX described feeling a sense of security and belonging, or expressed a sense of empathy and engagement. For example, "I went to physio yesterday and my physiotherapist was quite impressed by my progress…and that's why he still kept me on his caseload."

PX gave positive accounts of his experiences as a patient, which often contextualized the service as an interpersonal relationship in primary or tertiary services. Blockley et al. (43) found that supportive interpersonal relationships reduce patient vulnerability and that nurses play a key role in the development and maintenance of these relationships. Research conducted on the experience of physiotherapy in rehabilitation services reported finding one main theme of personal interactions, and five sub-themes (empathetic and caring physiotherapist, socialisation with other patients, alleviated boredom, changed perceptions of the weekend, and contentment with the amount of therapy) emerging from their data. Patients valued interacting with physiotherapists and other patients (44).

For example,

> "Today I have done too much physio lately [with the assistance of the staff] and my knees hurt it was a different pain now, they seem weak but yeah it hurts same old …..I have to figure it out, pretty much."

In addition, having a feeling of mutual understanding and recognition by staff enabled PX to move toward acceptance of his pain as something that was part of his daily living. These factors (personal relationship and understanding) appear to be

more important to an individual with pain than the amount of therapy received (44).

Domain 2: Accessibility Experiences:

(4) Environmental: The environmental theme in the accessibility domain constituted the next most discussed theme after pain in the entire video set. Fifty-one percent of all videos (33 of 65 total videos) contained comments related to the environmental theme.

In the environmental theme, comments were related to physical infrastructure, weather, the physical location of people around her, and the accessibility services. Positive comments included enjoying his time at an event, feeling accommodated, being independent and respected and enjoying the company of others.

For example, during a conference PX recorded a positive video about the physical accessibility offered by the hotel "after registration we got the room keys and I got the accessible room, which is pretty big and spacious, with a double bed and shower chair. Very nice, very accessible." A second example concerned a positive experience with the availability of accessible airport services: Denver airport was pretty good we landed late night at 12 am something and it was a long walk, so they offered the golf car and a wheelchair was also available, there were no long lines (for me, for immigration) …not to wait so it was great."

Poor accessibility for PX in the environment also occurred and resulted in negative expressions including feeling out of control, stressful, angry, alone, dependent, or disorganized. In particular, PX expressed frustrations with inaccessible physical environments or lack of accessible resources especially those that made her reliant on others for assistance. E.g.,: "Oh my gosh! Z (institution) needs to fix all the doors, those automatic doors are not working like you know first of all they're so heavy so I cannot push them and the people are so inconsiderate they see me struggling and I am kind of pushing it, but they still don't come by to help me, I don't know which world they're living in, but like that's pretty rude I would say but I guess they just don't care, so I got to deal with it. It sucks when it's nighttime after my 10 pm class there is like no one, the whole campus is dead so there is no one to help me."

In the social model of disability, the three main barriers experienced by people with disabilities (PWD) are environment, attitudinal and systemic in their environment (45). Physical barriers refer to obstacles in the built environment (e.g., architectural, transportation, communication, services, and physical infrastructure) (46), attitudinal to attitudes towards PWD by others (47), and systemic barriers introduced by a system that would include barriers from policies, procedures, legislation, cultural values (e.g., by government, and organizations (48). PX identified and experienced all of these barriers and, similar to comments in the pain and therapy themes, realized that he was facing a need to "deal with it on my own." As a result of PX's sudden onset of a chronic condition, her experience with barriers was also sudden and novel. Neither he nor his parents had any experience with the disability community, self-advocating, or available resources. Feelings of being alone and needing to address barriers independently are common occurrences among PWD (49). The process of learning to self-advocate and find resources can be challenging and is often frustrating, particularly for people with newly acquired disability (50, 51). Being able to record and reflect on experiences of barriers and challenges may expedite this learning process because evidence can be collected, reviewed and presented to other parties as support for requests, complaints, recommendations, or accommodations.

(5) Recommendation: Twenty-eight of 76 comments (37%) in the Accessibility domain were related to the Recommendation theme. They consisted of opinions, positive feedback, and offering suggestions about coping or accommodation strategies for institutions, individuals and events. These recommendations were particularly evident in comments about PX's educational institution. An example of a recommendation for an individual was:

> " In my evening class where we were talking about the wheelchair manufacturing industry and my professor put me on the spot, which was not nice……he can generalize and make statements, he doesn't have to put me on the spot … I have tried to speak to him but didn't get the chance I will follow up on the email."

An example of a recommendation for an event was: "I have a presentation in a week and I am just thinking about how it will go because there are stairs to get to the front, so even though someone clicks the slides would it still be the same thing if I speak from the back of the class, would that be the same as presenting from the front and how would that affect my presentation, and my marks".

Finally, an example of a suggestion for an institution was: "So I am actually debating, arguing with another professor too about another mark, participation mark. The way he marked, 'participation mark' is so weird that I am just like this doesn't make sense, so we're going back and forth, ……there is no point in arguing with him. So I am just waiting to get my final participation marks and take it to the director of the school or the UPD and discuss it with him."

It is well known that with some exceptions (e.g., McCloy and DeClou 2013) (52), few recent publications consider educational accessibility for a diverse spectrum of students who are identified as disabled in Canadian institutions (53). In addition, Rosenzweig (54), suggests that special training in which faculty, general, and accessibility staff must be educated about behaviours and disabilities strategies that could be used to support more inclusive classrooms and education. However, as discovered by PX, few instructors seem to have knowledge of these strategies and how to apply them in the classroom. Being able to record a situation and discuss the resulting impact on the individual may assist individuals in presenting evidence to concerned parties. It may also allow an individual to reflect on their reaction to a situation and consider appropriate actions at a later date or based on their own suggestions.

(6) Technical: There were only two comments (3%) in the Accessibility domain about the recording system and technology. Both comments related to the recording system failing to record or stop, and having to re-record the session. In this case study, we focused mainly on the first stage of the journey (building confidence in the documentation) while being mindful of the full journey. Consequently, the system used by PX was an early prototype where we wanted to first understand the process of and ability to record short videos. We also wanted to understand the types of topics and commentary that were generated. Despite experiencing some technical difficulties, PX was keen to document. For example: "Oh my gosh! okay this is the third time I am recording this and I want to talk about my evening class and its incident."

The technical difficulties experienced by PX assisted the research team in correcting errors and improving the system.

Domain 3: System Facing:

(7) Culture: This theme represents comments about customs and traditions in other countries related to access; these include comments on transportation, scheduling, out-of-pocket costs, and resources. In the system-facing theme, there are 14 of 52 (23%) comments). Attitudes towards disability are not always uniform within a region or even within a country. Different groups or individuals may have beliefs about disability that vary from those held by the wider society. Beliefs may also vary even within small communities and even within families.

PX experienced support (or lack of) from the same or different systems, often expressing positive and negative emotions associated with those experiences. For example, in Japan,

> "Japan was just amazing we got escorted out the doors,……The washrooms were very accessible, has a bed there as well… when we took the train, the train staff would escort me to the platform and took out a ramp and got me in.. And the best part……they already sent a message and the person was standing there, just waiting, it was mind-blowing and so accommodating….. streets were amazing flat straight, we went out into the city, it was very nice, loved it,". " I'm trying to board for Hanoi from Japan, which was pretty frustrating because they didn't allow to me to take the wheelchair to the gate, there like I am going to get my wheelchair with the luggage pickup, that was pretty annoying."

Since the United Nations adopted the Convention on the Rights of Persons with Disabilities, ratified by 168 countries, there has been progress as well as stubborn obstacles (55). PX encountered the results of efforts to improve accessibility as well as the obstacles that have yet to be rectified. Having a way of capturing the instances of those barriers may offer individuals like PX with support in advocating for improvements on a small scale. Collectively, if many narratives can be gathered together by more than one person, PWD or their associates can provide evidence to justify demands to communities, governments and/or organisations for improvements and resolutions for change.

Communication

The final theme captures PX's desire to express his personal stories of the interaction in hope of it positively influencing others. PX articulated the need for ethical and inclusive environments at the institution. However, most of the 38 comments (14) in this domain were negative where PX expressed disappointment, frustration, struggle and a desire to be treated fairly and with dignity.

e.g., "I got an assignment back and got a 70% which I think I had done better or would've done much better but I wasn't able to attend one of the classes due to an [medical] appointment. I told the professor and he didn't have any slides for that class. He tried to gather some notes and by the time he gave me the notes it was the day when the assignment was due and I submitted and then I saw the notes so it was too late to incorporate them…so I emailed the prof and let's see if he gets back to me. Technically speaking I should've been accommodated if someone has an appointment and does not have the slides just that day… no one takes notes in that class.

For e.g., "Oh my god! So it's more than a week and since I have reported to maintenance for the doors (to repair) and still haven't fixed it".

While these examples and most of PX's comments in this study are common and congruent with others (e.g., 45) they were new for PX. Being able to document these experiences may offer personal benefits such as having a mechanism for "venting", reflection and, perhaps, mediation. In addition, finding ways to automatically analyze the video material may offer opportunities for discovering patterns, topic threads or specific issues. Finally, if these videos can be collected together with others, they may be used to affect systemic change by providing evidence of patterns and ongoing, common issues.

Summary discussion

PX was able to successfully make recordings and capture her successes, failures, issues and concerns on a number of different but related topics using the MHMR technology. The MHMR interface appeared to have helped PX's need and want for chronicling an assortment of topics including ongoing barriers, hindrances experienced, dissatisfaction, pain, and excitement. Despite no enforced video length cut-off, brief recordings were recommended and seemed to be preferred by PX. Constraining recordings to as short as one minute would be a possible interface for the MHMR system and may simplify any large-scale application of the approach because it is easier to store and manage short videos compared with larger ones.

Another interface issue was allowing the deletion and re-recording of a video related to a specific domain. PX suggested that a video generated in a single take was more genuine. It is possible that others may find it difficult to make a point in a single take, particularly for novice users who must learn how to use the medium. However, it would be up to the user to decide

what is shared regardless. It is unclear whether PX was suggesting that his recordings became inauthentic as a result of multiple takes or whether others may judge her recordings as inauthentic from a more polished appearance and flow as a result of the practice offered by multiple recordings. Future research will examine the effect of multiple recordings on the assessment of trustworthiness, believability, and authenticity by the individual as well as those with whom the recordings would be shared. There is likely to be a trade-off between the view of authenticity, genuineness and the clarity and efficiency that comes with practice. Further research about restricting or reducing the number of takes and the effect on authenticity and exertion is required.

The second research question addressed the possible topics that were recorded by PX and whether those reflected his current life situation. No instructions or direction was provided to PX about which topics to use in order to give him control and flexibility over them. However, PX was recruited by the research team early on so he was aware of the purpose of the project which may have indirectly influenced topic selection. While there was a wide variety of topics that were broached by PX, the majority of them were related to her pain and experiences as a person with a disability. In addition, these topics could be categorized into domains related to the self or inward-looking, those related to a system (e.g., education, transportation, and government) or outward-looking, and to accessibility or barriers experienced using a standard qualitative evaluation method. Technical aspects of the MHMR system was the only topic not related to pain or disability.

PX had newly acquired the status of a person living with a major chronic disability. That may have influenced his perspective and experiences as they were all new to her. Involving people with either congenital disabilities or who have been living with a disability for a long time may show different results regarding the level or number of comments related to accessibility or barriers. However, people can experience disabling conditions suddenly at any point in their lives. MHMR can allow people to document those experiences for reflection or to serve as a simple method of collecting evidence for future treatment, advocacy or service acquisition.

Limitations

There are a number of limitations of this study; first and foremost is that it presents results from a single case study involving a self-selected, educated individual with newly acquired chronic pain and associated disabilities. A larger sample with a variety of people with chronic illness and pain will need to be recruited in order to make generalisations about the use and usefulness of MHMR. However, it was important to determine whether it would even be feasible to launch, support and use MHMR in the "wild" in a longitudinal context as part of the formative work that is necessary to justify future work.

A second limitation was that all files were encrypted once PX clicked the save button, for security purposes, and all date-related information was lost. This reflected using open encryption platforms and their default privacy configurations compatible with the Android devices used for this simple usability phase of the design process, rather than the comprehensive data registration service developed by gDial Inc. that is now being adapted to operate on Android systems. As described by Pennefather et al (8), there is a need to save as much contextualization information concerning the videos as possible to increase their value. This will require the creation of personally controlled data accounts to serve as private data repositories whose entries can be made selectively public for specified limited purposes (8). As such, we could not determine when specific recordings were made to examine any novelty effect of the MHMR system use. However, the frequency of MHMR recordings may not be related to novelty but instead to the salience of specific situations. The current iteration of MHMR has resolved this issue.

Another limitation was that we only examined the input stage of the MHMR system. We did not examine the curatorial process or output/sharing components of MHMR. It was important to determine whether the data-collecting task was feasible over the long term before efforts were made to complete the system. Written journals and diaries are common methods for collecting a patient's thoughts and feelings about their health situations (56). However, the abandonment rate is high (40) and many studies do not carry out longitudinal studies. We were interested in determining the feasibility of using the MHMR system to capture a patient's thoughts and feelings over a longitudinal time-frame as a first step in designing and developing the entire system described in **Figure 1**. Furthermore, the study does not explore topics such as digital literacy, inclusivity, and socioeconomic difficulties in the setting of chronic disease as potential influences on the outcome, however, these topics will be a focus of future research with additional participants.

Although there are limitations in this formative research, we believe that this study adds to the growing body of knowledge on digital and visual reporting and how it may play a role in addressing the needs of individuals living with chronic pain or illness. The project hypothesis, that communicating the experience of illness through recording and registering video commentary in a video scrapbook format may provide positive psycho-social benefits to some patients with chronic pain or illness, appears to be supported.

Conclusion

The paper presented two related notions: (1) to describe the experiences and insights of an individual with chronic illness in their day-to-day life, and (2) how technology could assist in addressing or overcoming the disability associated with that experience. A single perspective was explored: a person with a new but chronic pain condition who was learning to self-manage his condition and document that process with the technology provided. Using MHMR as a personal self-reporting diary/repository over a three-month period showed that it was feasible to generate ongoing and valuable insights over the longer term, as evidenced in the pilot case study. Moreover, themes could be identified that usefully could be distinguished in terms of patient facing (perceived consequences/outputs).

The next steps are to build a seamless and efficient communication channel that allows individuals with chronic illness,

healthcare professionals, and other stakeholders to access, evaluate, and act on acquired data, fostering a collaborative approach to care. We aim to employ artificial intelligence and data visualization techniques to analyse, summarize and illustrate the collections of the recorded qualitative data. This will involve developing interactive visual representations that summarize the data into meaningful collections and interpretations of the subjective/qualitative information, and then provide means to share those summaries with an individual's circle of care. It is also important to understand and evaluate how communication between a patient and their circle of care can be facilitated and whether that communication facilitates improved health and well-being outcomes.

To enable the appropriate use of such data in healthcare, we envision a system in which the collected information is securely and easily accessible to care practitioners and other stakeholders. Implementing a safe data-sharing method that conforms with all relevant data privacy standards will be a part of our recommended solution. Future studies will require testing the application with a larger group and integration of qualitative data with quantitative data available from other sources like wearable devices or patient portals.

Data availability statement

The datasets presented in this article are not readily available because the data will not be available in a public repository due to the participant's privacy and confidentiality issues. Requests to access the datasets should be directed to SFL.

Ethics statement

The studies involving humans were approved by Toronto Metropolitan University (formerly, Ryerson University). The studies were conducted in accordance with the local legislation and institutional requirements. A participant provided his/her written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

SFL, DF and PP researched literature and conceived the study with input from WS, and AK. SFL PP and DF led protocol development, gained ethical approval, and patient recruitment. SFL, CBH and SFS did the data analysis. ZAA provided guidance on the analysis. SFL wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article.

Funding was generously provided by the Canadian Institute for Health Research (CIHR-PEG-!456215).

Acknowledgments

We also thank the university staff, system developers Raza Al Miri and Vera Parilova, and gDial for supporting this project. Finally, we thank PX for participating and persisting with this project. gDial was not involved in the study design, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

Conflict of interest

WS and PP were employed by gDial Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Pennefather P, Seaborn K, Fels DI. Inclusive systemic design for health system flourishment. In: Jones P, editor. *Systemic design*. Tokyo: Springer (2018). p. 69–85.

2. Richards NC, Gouda HN, Durham J, Rampatige R, Rodney A, Whittaker M. Disability, noncommunicable disease and health information. *Bull W H O*. (2016) 94(3):230. doi: 10.2471/BLT.15.156869

3. Taha J, Czaja SJ, Sharit J, Morrow DG. Factors affecting usage of a personal health record (PHR) to manage health. *Psychol Aging*. (2013) 28(4):1124. doi: 10.1037/a0033911

4. Butler DS, Moseley GL. *Explain pain*, 2nd ed. Adelaide: No igroup publications (2013).

5. Basch E. Patient-reported outcomes—harnessing patients' voices to improve clinical care. *N Engl J Med*. (2017) 376(2):105–8. doi: 10.1056/NEJMp1611252

6. Friedman CP, Rubin JC, Sullivan KJ. Toward an information infrastructure for global health improvement. *Yearb Med Inform*. (2017) 26(01):16–23. doi: 10.15265/IY-2017-004

7. Fels DI, Lakha SF, Holtzman CB, Khandwala A, Whitfield M, Suhanic W, Pennefather P. *Videobooking: a person-centred record of living with chronic*

conditions and associated disabilities. In: *International conference on computers helping people with special needs*. Cham: Springer (2018). p. 489–92.

8. Pennefather P, Suhanic W, Lakha F, Fels D. Publishing authentic, private, personal data about service quality of healthcare for pain. In: ELPUB (2018). 2018 Jun 22.

9. Turco J, Gibbs C, Fels DI. User generated video reviews by hotel guests. (2017). To be presented at ENTER2017, Rome. doi: 10.1145/503376.503378

10. Rigby M, Georgiou A, Hyppönen H, Ammenwerth E, de Keizer N, Magrabi F, et al. Patient portals as a means of information and communication technology support to patient-centric care coordination–the missing evidence and the challenges of evaluation. *Yearb Med Inform*. (2015) 24(01):148–59. doi: 10.15265/IY-2015-007

11. Bhatia A, Kara J, Janmohamed T, Prabhu A, Lebovic G, Katz J, et al. User engagement and clinical impact of the manage my pain app in patients with chronic pain: a real-world, multi-site trial. *JMIR Mhealth Uhealth*. (2021) 9(3): e26528. doi: 10.2196/26528

12. Najafizada M, Rahman A, Oxford K. Analyzing models of patient-centered care in Canada through a scoping review and environmental scan. *Z Gesundh Wiss*. (2023) 31(3):355–67. doi: 10.1007/s10389-021-01528-8

13. Ehrenstein V, Kharrazi H, Lehmann H, Taylor CO. Obtaining data from electronic health records. In: Gliklich RE, Leavy MB, Dreyer NA, editors. *Tools and technologies for registry interoperability, registries for evaluating patient outcomes: A user's guide*, 3rd ed, Addendum 2. Rockville, MD: Agency for Healthcare Research and Quality (US) (2019). Chapter 4. p. 62. Available at: https://www.ncbi.nlm.nih.gov/books/NBK551878/

14. Greenhalgh T, Hinder S, Stramer K, Bratan T, Russell J. Adoption, non-adoption, and abandonment of a personal electronic health record: case study of HealthSpace. *Br Med J*. (2010) 341:c5814. doi: 10.1136/bmj.c5814

15. Dendere R, Slade C, Burton-Jones A, Sullivan C, Staib A, Janda M. Patient portals facilitating engagement with inpatient electronic medical records: a systematic review. *J Med Internet Res*. (2019) 21(4):e12779. doi: 10.2196/12779

16. Tsai CH, Eghdam A, Davoody N, Wright G, Flowerday S, Koch S. Effects of electronic health record implementation and barriers to adoption and use: a scoping review and qualitative analysis of the content. *Life*. (2020) 10(12):327. doi: 10.3390/life10120327

17. Maier AM, Kreimeyer M, Lindemann U, Clarkson PJ. Reflecting communication: a key factor for successful collaboration between embodiment design and simulation. *J Eng Des*. (2009) 20(3):265–87. doi: 10.1080/09544820701864402

18. Harrington C, Dillahunt TR. *Eliciting tech futures among black young adults: a case study of remote speculative co-design*. In: *Proceedings of the 2021 CHI conference on human factors in computing systems*; 2021 May 6. p. 1–15.

19. World Health Organization. *Global patient safety action plan 2021-2030: towards eliminating avoidable harm in health care*. Geneva: World Health Organization (2021).

20. Eklund JH, Holmström IK, Kumlin T, Kaminsky E, Skoglund K, Höglander J, et al. "Same same or different?" A review of reviews of person-centered and patient-centered care. *Patient Educ Couns*. (2019) 102(1):3–11. doi: 10.1016/j.pec.2018.08.029

21. Donabedian A. Evaluating the quality of medical care. *Milbank Q*. (2005) 83 (4):691–729. doi: 10.1111/j.1468-0009.2005.00397.x

22. Lakha SF. A Method for Evaluation of the Management of Chronic Non-Cancer Pain in Global Cities (Doctoral dissertation) (2016).

23. Lakha SF, Ballantyne P, Badr H, Agboatwala M, Mailis A, Pennefather P. Perspective of pain clinicians in three global cities on local barriers to providing care for chronic noncancer pain patients. *Pain Res Manag*. (2019) 2019:2,5. doi: 10.1155/2019/3091309

24. Mills T, Lawton R, Sheard L. Advancing complexity science in healthcare research: the logic of logic models. *BMC Med Res Methodol*. (2019) 19(1):55. doi: 10.1186/s12874-019-0701-4

25. Genes N, Violante S, Cetrangol C, Rogers L, Schadt EE, Chan YF. From smartphone to EHR: a case report on integrating patient-generated health data. *NPJ Digit Med*. (2018) 1(1):23. doi: 10.1038/s41746-018-0030-8

26. Yin RK. *Case study research design and methods*. Thousand Oaks, CA: Sage (2014. 282 pages. Canadian Journal of Program Evaluation. 2015; 30 (1).

27. Spinuzzi C. The methodology of participatory design. *Tech Commun*. (2005) 52 (2):163–74.

28. Clemensen J, Larsen SB, Kyng M, Kirkevold M. Participatory design in health sciences: using cooperative experimental methods in developing health services and computer technology. *Qual Health Res*. (2007) 17(1):122–30. doi: 10.1177/1049732306293664

29. Hill CE, Knox S, Thompson BJ, Williams EN, Hess SA, Ladany N. Consensual qualitative research: an update. *J Couns Psychol*. (2005) 52(2):196. doi: 10.1037/0022-0167.52.2.196

30. Malterud K. Qualitative research: standards, challenges, and guidelines. *Lancet*. (2001) 358(9280):483–8. doi: 10.1016/S0140-6736(01)05627-6

31. Mays N, Pope C. Assessing quality in qualitative research. *Br Med J*. (2000) 320 (7226):50–2. doi: 10.1136/bmj.320.7226.50

32. Wiech K, Tracey I. The influence of negative emotions on pain: behavioral effects and neural mechanisms. *Neuroimage*. (2009) 47(3):987–94. doi: 10.1016/j.neuroimage.2009.05.059

33. Janssen SA. Negative affect and sensitization to pain. *Scand J Psychol*. (2002) 43 (2):131–7. doi: 10.1111/1467-9450.00278

34. Öngür D, Price JL. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cerebral cortex*. (2000) 10(3):206–19. doi: 10.1093/cercor/10.3.206

35. Strigo IA, Simmons AN, Matthews SC, Arthur D, Paulus MP. Association of major depressive disorder with altered functional brain response during anticipation and processing of heat pain. *Arch Gen Psychiatry*. (2008) 65(11):1275–84. doi: 10.1001/archpsyc.65.11.1275

36. van Middendorp H, Lumley MA, Jacobs JW, Bijlsma JW, Geenen R. The effects of anger and sadness on clinical pain reports and experimentally-induced pain thresholds in women with and without fibromyalgia. *Arthritis Care Res (Hoboken)*. (2010) 62(10):1370–6. doi: 10.1002/acr.20230

37. Burns JW, Gerhart JI, Bruehl S, Peterson KM, Smith DA, Porter LS, et al. Anger arousal and behavioral anger regulation in everyday life among patients with chronic low back pain: relationships to patient pain and function. Health psychology: official journal of the division of health psychology. *Am Psychol Assoc*. (2015) 34(5):547–55. doi: 10.1037/hea0000091

38. Corbett M, Foster NE, Ong BN. Living with low back pain—stories of hope and despair. *Soc Sci Med*. (2007) 65(8):1584–94. doi: 10.1016/j.socscimed.2007.06.008

39. Schumacher KL, Koresawa S, West C, Dodd M, Paul SM, Tripathy D, Koo P, Miaskowski C. The usefulness of a daily pain management diary for outpatients with cancer-related pain. In: *Oncology nursing forum*. Oncology Nursing Society (2002). (Vol. 29, No. 9). p. 2–3. doi: 10.1188/02.ONF.1304-11313

40. Stinson JN, Lalloo C, Harris L, Isaac L, Campbell F, Brown S, et al. Icancope with pain^{TM}: user-centred design of a web-and mobile-based self-management program for youth with chronic pain based on identified health care needs. *Pain Res Manag*. (2014) 19(5):257–65. doi: 10.1155/2014/935278

41. Cano A, Goubert L. What's in a name? The case of emotional disclosure of pain-related distress. *J Pain*. (2017) 18(8):881–8. doi: 10.1016/j.jpain.2017.01.008

42. Fredrickson BL. The role of positive emotions in positive psychology: the broaden-and-build theory of positive emotions. *Am Psychol*. (2001) 56(3):218. doi: 10.1037/0003-066X.56.3.218

43. Blockley C, Alterio M. Patients' experiences of interpersonal relationships during first time acute hospitalisation. *Nurs Praxis N Z*. (2008) 24(2):23.

44. Peiris CL, Taylor NF, Shields N. Patients value patient-therapist interactions more than the amount or content of therapy during inpatient rehabilitation: a qualitative study. *J Physiother*. (2012) 58(4):261–8. doi: 10.1016/S1836-9553(12)70128-5

45. Oliver M. *Politics of disablement*. London: London MacMillan Press (1990).

46. Congress US. Americans with Disabilities Act of 1990, As Amended.

47. Barton B. Dreams deferred: disability definitions, data, models, and perspectives. *J Soc Soc Welfare*. (2009) 36:13. doi: 10.15453/0191-5096.3470

48. Grandpierre V, Milloy V, Sikora L, Fitzpatrick E, Thomas R, Potter B. Barriers and facilitators to cultural competence in rehabilitation services: a scoping review. *BMC Health Serv Res*. (2018) 18(1):23. doi: 10.1186/s12913-017-2811-1

49. Griful-Freixenet J, Struyven K, Verstichele M, Andries C. Higher education students with disabilities speaking out: perceived barriers and opportunities of the universal design for learning framework. *Disabil Soc*. (2017) 32(10):1627–49. doi: 10.1080/09687599.2017.1365695

50. Daly-Cano M, Vaccaro A, Newman B. College student narratives about learning and using self-advocacy skills. *J Postsecond Educ Disabil*. (2015) 28(2):213–27.

51. Agarwal N, Moya EM, Yasui NY, Seymour C. Participatory action research with college students with disabilities: photovoice for an inclusive campus. *J Postsecond Educ Disabil*. (2015) 28(2):243–50.

52. McCloy U, DeClou L. *Disability in Ontario: postsecondary education participation rates, student experience and labour market outcomes*. Toronto: Higher Education Quality Council of Ontario (2013).

53. Marquis E, Schormans AF, Jung B, Vietinghoff C, Wilton R, Baptiste S. Charting the landscape of accessible education for post-secondary students with disabilities. *Can J Disabil Stud*. (2016) 5(2):31–71. doi: 10.15353/cjds.v5i2.272

54. Rosenzweig K. Are today's general education teachers prepared to meet the needs of their inclusive students? (2009).

55. Degener T. 10 Years of convention on the rights of persons with disabilities. *Netherlands Quart Hum Rights*. (2017) 35(3):152–7. doi: 10.1177/0924051917722294

56. Niles AN, Haltom KE, Mulvenna CM, Lieberman MD, Stanton AL. Randomized controlled trial of expressive writing for psychological and physical health: the moderating role of emotional expressivity. *Anxiety Stress Coping*. (2014) 27(1):1–7. doi: 10.1080/10615806.2013.802308

frontiers | Frontiers in Digital Health

TYPE Original Research
PUBLISHED 29 January 2024
DOI 10.3389/fdgth.2024.1287340

OPEN ACCESS

EDITED BY
Kim Cornelia Marie-Louise Bul,
Coventry University, United Kingdom

REVIEWED BY
Joyce Poh Foong Lee,
Tunku Abdul Rahman University, Malaysia
Octavian Andronic,
Carol Davila University of Medicine and Pharmacy, Romania

*CORRESPONDENCE
Hayeon Song
✉ hayeon.song@gmail.com

RECEIVED 01 September 2023
ACCEPTED 10 January 2024
PUBLISHED 29 January 2024

CITATION
Jeong H, Yoo JH, Goh M and Song H (2024) Deep breathing in your hands: designing and assessing a DTx mobile app.
Front. Digit. Health 6:1287340.
doi: 10.3389/fdgth.2024.1287340

Deep breathing in your hands: designing and assessing a DTx mobile app

Harim Jeong[1,2], Joo Hun Yoo[2,3], Michelle Goh[1] and Hayeon Song[1]*

[1]Department of Interaction Science, SungKyunKwan University, Seoul, South Korea, [2]Department of Artificial Intelligence, SungKyunKwan University, Suwon, South Korea, [3]Hippo T&C, Suwon, South Korea

Digital Therapeutics (DTx) are experiencing rapid advancements within mobile and mental healthcare sectors, with their ubiquity and enhanced accessibility setting them apart as uniquely effective solutions. In this evolving context, our research focuses on deep breathing, a vital technique in mental health management, aiming to optimize its application in DTx mobile platforms. Based on well-founded theories, we introduced a gamified and affordance-driven design, facilitating intuitive breath control. To enhance user engagement, we deployed the Mel Frequency Cepstral Coefficient (MFCC)-driven personalized machine learning method for accurate biofeedback visualization. To assess our design, we enlisted 70 participants, segregating them into a control and an intervention group. We evaluated Heart Rate Variability (HRV) metrics and collated user experience feedback. A key finding of our research is the stabilization of the Standard Deviation of the NN Interval (SDNN) within Heart Rate Variability (HRV), which is critical for stress reduction and overall health improvement. Our intervention group observed a pronounced stabilization in SDNN, indicating significant stress alleviation compared to the control group. This finding underscores the practical impact of our DTx solution in managing stress and promoting mental health. Furthermore, in the assessment of our intervention cohort, we observed a significant increase in perceived enjoyment, with a notable 22% higher score and 10.69% increase in positive attitudes toward the application compared to the control group. These metrics underscore our DTx solution's effectiveness in improving user engagement and fostering a positive disposition toward digital therapeutic efficacy. Although current technology poses challenges in seamlessly incorporating machine learning into mobile platforms, our model demonstrated superior effectiveness and user experience compared to existing solutions. We believe this result demonstrates the potential of our user-centric machine learning techniques, such as gamified and affordance-based approaches with MFCC, which could contribute significantly to the field of mobile mental healthcare.

KEYWORDS

digital therapeutics (DTx), Human–Computer Interaction (HCI), mobile health interventions, machine learning feedback, gamification design, user engagement

1 Introduction

The digital healthcare and digital therapeutics (DTx) market has been experiencing a rapid expansion, a surge further catalyzed by the COVID-19 pandemic, which underscores the need for digital solutions to address mental health disorders (1). Current research indicates that of 18 identified DTx products, only six specifically target on treating

mental disorders with a particular emphasis on depression, anxiety disorders, and insomnia (2, 3). This pattern aligns with the U.S. Food and Drug Administration's (FDA) temporary policy to broaden patient access to DTx for mental healthcare amidst the pandemic (4).

In the realm of the mental health-centric DTx, the digital implementations of Cognitive Behavioral Therapy (CBT) are prevalent (3), often incorporating breathing exercises into these interventions (5). Notably, the integration of CBT and breathing exercises has been proposed to amplify therapeutic benefits (6). Building on this foundation, our study introduces an innovative approach to breathing exercises, aimed to enhance the effectiveness of an array of DTx targeting mental disorders. We aspire to increase their potency while ensuring sustained usage, a crucial aspect of successful DTx deployment. Our primary research goal is to foster the development of a more proficient DTx by evaluating the effectiveness of our proposed breathing exercise method.

Due to its potency in fostering engagement and elevating motivation, gamification is an active method in learning domains (7). It merges entertainment and tasks by integrating game elements into non-gaming contexts. This technique is especially prevalent in digital healthcare, essential for driving behavior change. In digital health services, where prolonged use is often necessary, gamification enhances engagement and supports sustained user participation. This aligns with the goals of these services. A systematic review demonstrates that gamification and serious games effectively encourage behavior change and heighten motivation (8). These elements contribute to the expectation of improved treatment outcomes within digital health interventions.

At present, DTx targeting mental health are mainly presented as mobile applications (3). Numerous mobile applications focus on deep breathing exercises. Unfortunately, many of these applications primarily offer passive animations and fail to actively engage users, a crucial component for effective DTx and healthcare. Research suggests employing machine learning techniques to offer personalized feedback in such applications (9). Yet, the limited computational prowess of current mobile devices introduces practical challenges. Accordingly, our study presents an efficient system designed to minimize the model weight for smooth operation in a mobile application environment, thus actualizing the methods proposed in prior research.

2 Literature review

2.1 Gamification effect

Following the broader discussion of gamification's role in digital healthcare in the introduction, we now focus on its specific impact and application in digital therapeutic interventions. While gamification has been acknowledged for its ability to merge entertainment with tasks, its application in digital health goes beyond mere engagement (10). It plays a significant role in facilitating sustained user interaction, particularly in applications requiring long-term commitment and adherence (11).

In the context of digital therapeutics, gamification is not just about adding game-like elements; it's about creating a more immersive and interactive experience that resonates with users (12). This approach is crucial in interventions where user motivation and continuous participation are key to successful outcomes. Moreover, gamification strategies in digital health have been shown to effectively drive behavior change, a central goal in many therapeutic interventions (13).

By integrating these elements into digital health interventions, we can transform the user experience from passive to active, thereby potentially improving adherence and treatment outcomes (14). This aligns with the overarching goal of digital therapeutics: to engage users in a meaningful way that promotes positive health behaviors and outcomes. The use of gamification in our study aims to leverage these benefits, creating an engaging and effective platform for deep breathing exercises.

Based on our review of gamification and its application in healthcare, the effects and flow of gamification have been summarized as follows. This summary illustrates how gamification can transition the user experience from a passive to an active role, underlining its significance in enhancing user engagement and motivation. This transformation is crucial in digital therapeutics, where active participation can lead to improved health outcomes. A visual representation of the effects and flow of gamification, as derived from our review, is depicted in Figure 1 below.

2.2 Gamification in deep breathing

As the digital therapeutic landscape continues to evolve, there's an increasing emphasis on enhancing user engagement to maximize therapeutic benefits. One innovative strategy that's gaining traction is the incorporation of gamification into deep breathing interventions. By gamifying deep breathing exercises, these interventions aim not only to harness the therapeutic advantages of controlled respiration but also to elevate user commitment, thereby improving the overall effectiveness of the intervention (15, 16).

Building upon the foundation set by previous studies, such as "Calm: Blow away your Stress" and "Breeze", our research proposes an optimal design for gamifying deep breathing interventions that target relaxation. For instance, the mobile application "Calm: Blow away your Stress" (15) invites users to dissipate twelve clouds using their breath through the microphone. The application's animated clouds and breath-based feedback stimulate user curiosity and engagement. However, the study did not distinctly address the explicit impact of these elements on engagement.

Another gamified deep breathing intervention is exemplified by "Breeze" (17). In this application, users control the sailing of a boat in a game-like setting by breathing directly into the microphone. The design incorporates a changing background as the boat sails

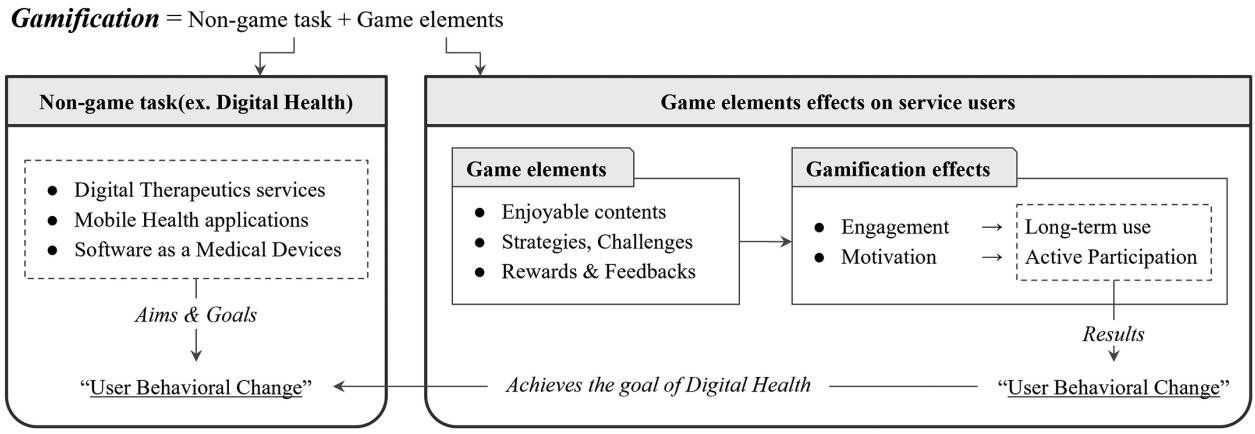


FIGURE 1
Effects of gamification and its flow.

further, encouraging continued user interaction. Feedback is provided based on the user's control over the boat's movement, with the Mel-Frequency Cepstral Coefficients (MFCC) feature extraction technique ensuring the feedback's accuracy by assessing the user's breathing state.

2.3 Affordance-based design

One of the main components of digital health is self-care. This feature signifies a shift from the traditional approach of relying on hospital visits and medical staff for diagnoses and treatments, empowering individuals to manage their own health needs as required (18). Because digital health or DTx requires self-management and system utilization by the user, Human–Computer Interaction (HCI) becomes an essential aspect. Specifically, appropriate guidance should be integrated within the system to ensure users can easily navigate digital health systems (19).

One design element to facilitate this intuitive guidance is affordance-based design. Affordances, defined as the inherent qualities of a design that dictate how an object should be interacted with, can help address a major challenge in digital health: providing comprehensive guidance with minimal user effort (20). Beyond the affordance design for simple self-management, it is posited that for digital health, the effectiveness of treatment can be maximized when the five objectives—social, cognitive, identity, emotional, and functional—are well-communicated through affordances (21).

The paper emphasizes the different affordances' importance in digital health, which includes social elements fostering a sense of belonging and support, cognitive aspects articulating the service's advantages, identity elements catering to users' ideal self-aspirations, emotional components generating positive sentiments, and functional attributes using technology to increase acceptance.

In summary, where self-care is paramount in digital health, an affordance-based design that helps users easily understand how to use the system is crucial. The effect can also be maximized when the five key affordances are adequately fulfilled.

2.4 Machine learning for deep breathing

Technological advancements in machine learning have significantly enhanced the domain of relaxation therapy, especially in the application of deep breathing exercises (22). The use of mobile devices' microphone modules for real-time assessment and feedback of breathing states represents a major leap forward (17). This development is crucial for providing users with immediate and accurate feedback on their breathing patterns, essential for effective relaxation therapy (23).

Machine learning techniques, such as Mel Frequency Cepstral Coefficient (MFCC), have been particularly instrumental in this regard. Originally used in speech recognition (24, 25), MFCC has now become a primary technique for analyzing breathing data. Its ability to extract significant features through spectral analysis has been crucial in developing more sophisticated and user-friendly relaxation therapy applications (26).

Building on this foundation, our study has implemented a deep breathing feature that provides accurate feedback using MFCC. Considering that most mental healthcare software is currently mobile-based, our focus has been on developing a system that is well-suited for the mobile environment. This approach ensures that our deep breathing exercises are not only effective but also accessible and practical for users on mobile platforms, aligning with the trend towards mobile health solutions in mental healthcare.

2.5 Heart rate variability (HRV) measurement

Heart Rate Variability (HRV) serves as a major indicator of the autonomic nervous system's functionality and has gained importance in the study of mental disorders (27). The analysis of

10.3389/fdgth.2024.1287340

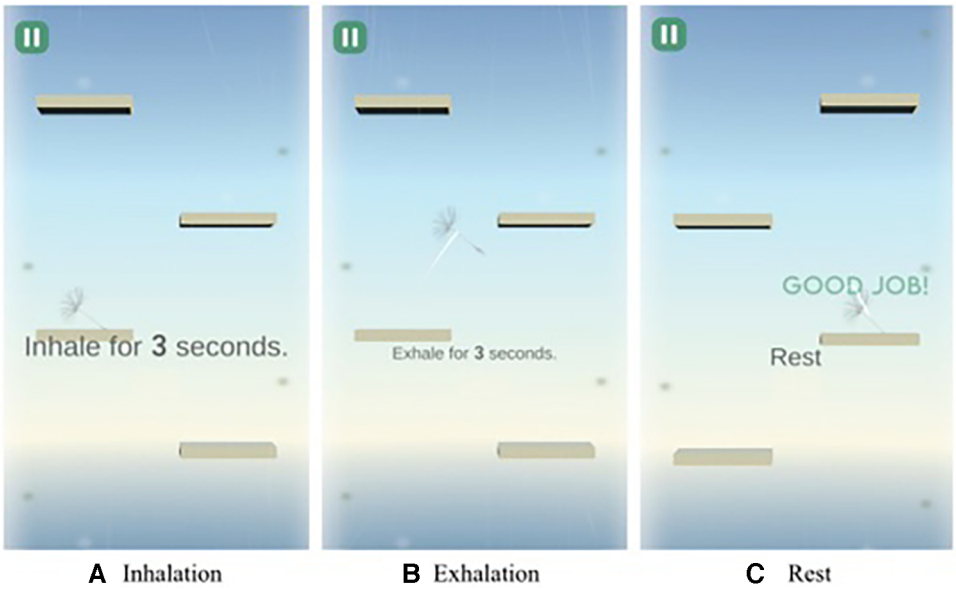


FIGURE 2
Gamified deep breathing app design.

HRV data, using methods such as frequency domain and time domain analysis, provides valuable insights into an individual's physiological responses under different conditions (28). These insights are pivotal for understanding the impact of various therapies, including deep breathing exercises, on mental well-being (29).

The measurement of HRV, therefore, becomes an integral part of our study, offering a window into the physiological impacts of our proposed deep breathing exercises. This rich source of data is instrumental in evaluating the efficacy of our interventions in a quantifiable and scientific manner. Considering the objective and scientific nature of HRV measurement, we have chosen this method as a key component to substantiate the effectiveness of our study. By proving the efficacy through HRV analysis, we aim to provide more objective and concrete evidence of the therapeutic benefits, further validating the practical implications of our research in mental health treatment.

3 Methodology

3.1 Designing a gamified deep breathing system

The present study introduces a gamified deep breathing system, infusing gamification principles into relaxation exercises. This novel approach aspires to alleviate the monotony of traditional deep breathing exercises, aiming for increased user immersion. To foster continuous engagement, our system offers real-time feedback, employing machine learning to provide accurate insights into the user's breathing dynamics.

Our design incorporates 'affordance-based design' principles, facilitating intuitive user perception and interactions within the application environment. The user interface employs straightforward and intuitive on-screen visuals corresponding with different breathing phases: inhalation, exhalation, and rest. During inhalation, a 'wind' effect descends on the screen, guiding users to breathe in. For exhalation, users blow into the microphone, causing the 'wind' effect to ascend, reflecting the exhale action. Finally, in the rest phase, the 'wind' effect disappears, signaling users to pause their breathing. The actual implementation of this breathing system can be seen in Figure 3.

In contrast, the application used by the control group, while also providing visual cues for breathing indication, lacked the gamification and affordance design elements of our system. This key difference lies in the absence of interactive elements such as the 'wind' effect and the dandelion seed metaphor, which are integral to our system's user engagement strategy. The control group's application primarily relied on simpler visual cues without the interactive and immersive components that characterize our gamified approach. An example screen of the application used by the control group can be found in Figure 4.

This design approach ensures a more engaging and interactive experience, aligning with our goal of enhancing user immersion in deep breathing exercises through gamification. By visually and interactively guiding users through each breathing phase, the system provides a unique and effective way to practice deep breathing, making the exercise both enjoyable and beneficial.

A pivotal design feature is the dandelion seed, a visual metaphor symbolizing the need for gentle, controlled breathing. Considering the key affordances of digital health outlined by Wong et al. (21), our design integrates elements of cognitive, emotional, identity, and functional affordances. The social aspect, though ideal, is outside this study's scope. For cognitive reinforcement, the application elucidates the merits of deep

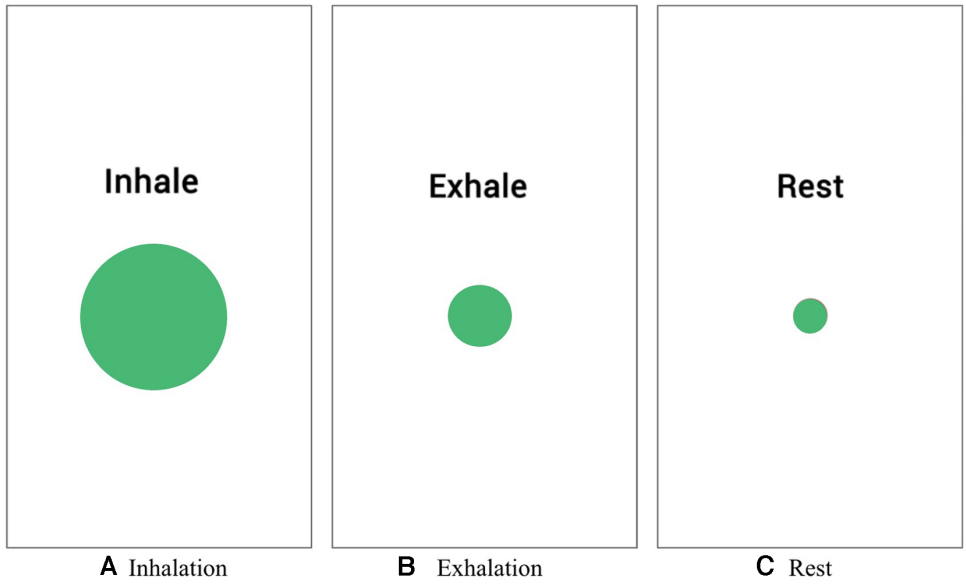


FIGURE 3
Circle relaxation app design.

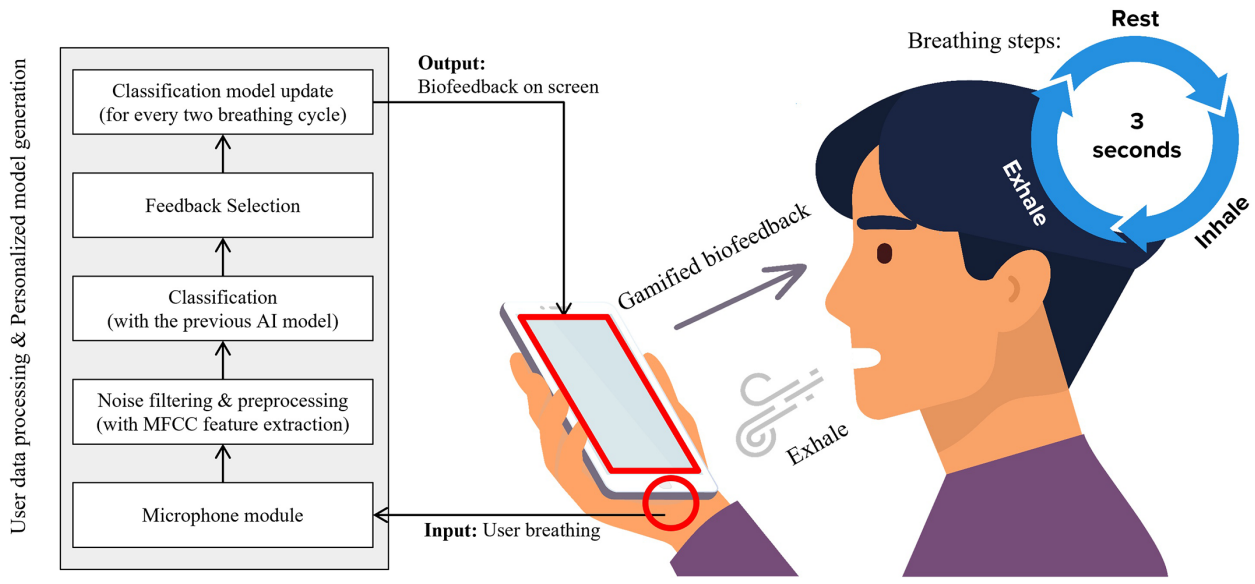


FIGURE 4
Overview of gamified deep breathing system.

breathing and its mental health impacts. Positive feedback messages uplift emotional affordance. The system's visualization of a dandelion ascending and ultimately blossoming reinforces identity. Clear insights into the system's workings (e.g., federated learning) amplify functional affordance.

For the social aspect, the system would ideally allow users to share their progress and engage with others, promoting a sense of belonging and support. However, given the scope of this study, this feature was not included in the current implementation. To enhance cognitive affordance, the application offers guidance on the benefits of deep breathing on relaxation and its overall influence on mental health before exercise. This approach strengthens users' understanding of the intervention's efficacy and mechanics, potentially boosting its overall effectiveness. Emotional affordance can be stimulated through positive reinforcement and fostering a positive self-perception.

The system achieves this by providing encouraging feedback, such as "Good Job" and "Almost there," following each game attempt. The incremental progress towards the goal and culminating in a blossoming flower upon a successful session establishes a rewarding cycle that promotes emotional satisfaction. Identity affordance is amplified through having a positive self-image through a positive transformation, enhancing both self-efficacy and motivation. Users are invited to embrace this positive identity by envisioning themselves as a dandelion gradually ascending the stairs and ultimately blooming into a flower in the end. This imagery embodies personal growth and positive transformation for the users. Lastly, functional affordance is elevated when users have a clear comprehension of the system's technical mechanisms, such as federated learning, promoting greater acceptance and trust in the system. We reinforced this by including messages like "synching your breath with the system" during moments of rest in the game.

To sustain user motivation, our system implements progressive disclosure, unveiling features progressively to retain user engagement (30). For instance, the dandelion seed's ascension up the stairs mirrors the user's progress.

The efficacy of these design features in fostering engagement, motivating users, and improving the intervention's effectiveness will be evaluated using appropriate evaluation metrics and methods, which will be elaborated upon in the following sections.

3.2 Implementation

Our implementation strategy involved leveraging mobile device technology to create an accessible and effective deep-breathing application. The microphone module of the device is used to detect the user's breathing, with ambient noise levels measured to ensure accuracy. The MFCC method is utilized for feature extraction, differentiating between correct and incorrect breathing patterns.

During the deep breathing exercises, the system guides users through a series of rest, inhale, and exhale phases, with the dandelion seed's movement on the screen corresponding to the user's breathing. We conducted a preliminary study with 40 college students to refine the deep breathing recognition system, creating a dataset of breathing recordings under diverse conditions. This data was used to train and refine our model, ensuring its accuracy and reliability.

The final system was developed using Python and TensorFlow, and the interface was built with Unity. Our approach emphasizes the personalization of the breathing exercises, tailoring the experience to each user's needs and progress.

The application's design leverages the built-in microphone of mobile devices to detect users' breathing patterns, thereby providing gamified biofeedback. As users engage with the system, their continuous breathing practice animates a dandelion seed in the game interface. However, challenges arise when ambient noise might erroneously influence the seed's movement. Addressing such potential inaccuracies was crucial, as inconsistencies in the feedback can erode users' trust in the application's effectiveness (31).

Upon the game's initiation, the microphone module automatically measures the surrounding noise level for a span of three seconds. This ambient noise reading serves as a baseline against which the user's breathing is compared with the help of a threshold algorithm. To refine the accuracy of this system, we employed the Mel-Frequency Cepstral Coefficient (MFCC) for feature extraction. Furthermore, convolutional neural network machine-learning techniques were used to discern genuine breathing patterns from both ambient noise and incorrect breathing phases. Figure 4 illustrates the steps of receiving user input, recognizing the breathing phase, collecting newly acquired data, and updating the application's classification model.

Guided by the application, users cycle through a series of rest, inhale, and exhale phases. During the exhale phase, the application assesses the sound it captures. The dandelion seed's movement on the screen corresponds to this classified exhalation. A strong exhalation might propel the seed considerably, while a weak one might not move the seed at all. Such visual feedback is instrumental, offering users insights into their breath strength. This dynamic aims not just to mirror the user's breathing pattern but also to discourage behaviors like hyperventilation and instead foster proper breathing techniques geared towards relaxation.

To refine the deep breathing recognition system, our team conducted a study involving 40 college students. This exercise produced a dataset of a total of 200 1-minute-long breathing recordings made under diverse environmental conditions. Within each recorded file, sections corresponding to exhale, inhale, and rest phase were further subdivided using a windowing method. These recordings were subsequently transformed from their initial time-domain waveform format to a more illustrative mel-spectrogram. The MFCC process was then used to extract distinct sound characteristics from each recording.

The recorded audio signal data, initially in the form of a time-domain waveform, was converted into a mel-spectrogram using a Fourier transform in the frequency spectrum. This conversion enabled us to represent the signal visually. Subsequently, the MFCC method was carried out to extract representative features of each recorded breathing phase (25). Firstly, windowing was applied to divide each crucial section, after which the section energy was computed using the Mel Filter Bank in the power spectrum for each divided signal data (32). By applying cepstral analysis to the Mel spectrum, which was analyzed through the Mel Filter Bank, we could obtain the MFCC. Each audio signal was thus converted into an MFCC format through these processing steps, which facilitated the extraction of unique characteristics of the corresponding sound.

Real-time classification of each breathing phase was made possible by harnessing a MobileNetV2-based model. The lightweight nature of MobileNetV2 made it an apt choice as the foundational architecture for our custom breathing classification model. When our initial model was tested against the MFCC feature data, it boasted an impressive 93.33% accuracy. Moreover, it achieved 93.65% average sensitivity and 96.58% average specificity across the three distinct breathing phases.

A notable aspect of the deep breathing application presented in this study is the personalized classification model update feature for each user. While previous studies also apply transfer learning and machine learning algorithms for user breathing measurement, they

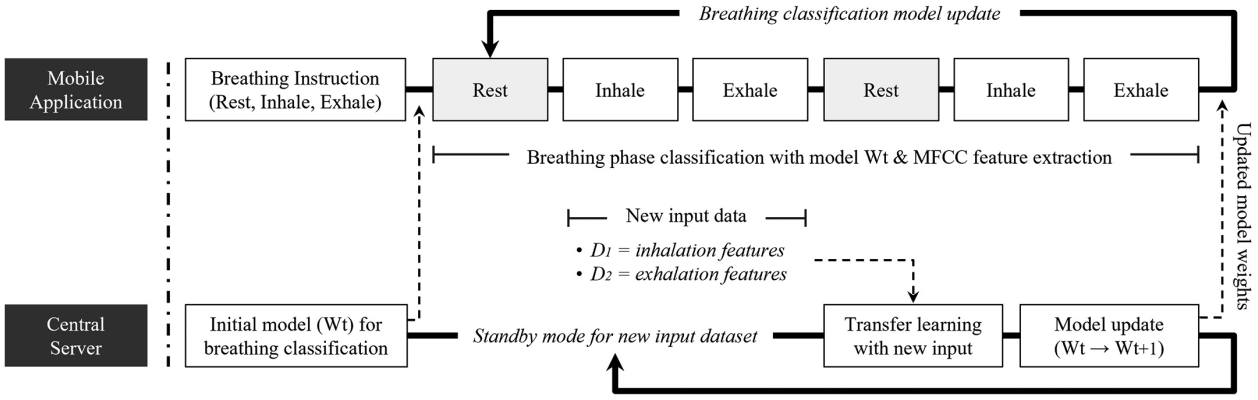


FIGURE 5
Classification model update process during deep breathing.

implemented a single, generalized weight set model into the system. Contrarily, we designed our proposed breathing application to maximize the effectiveness of digital treatments by enabling each user to cultivate their own optimized breathing model. Figure 5 describes the communication between mobile applications and central servers and how individual models are updated.

As previously mentioned, we adopted and utilized the model update process of the federated learning structure to create and update a model for each user. In a manner similar to federated learning, the initial model Wt is stored on a central server and provided to a new user before the deep breathing system commences. In the mobile application, the method repeats in the order of rest-inhale-exhale and carries out breathing classification and feature extraction using the current model Wt in each section. The newly inputted user breathing data during the inhale and exhale phases are processed, stored, and transmitted to the central server during the next rest cycle. The central server then applies the new input data to the transfer learning algorithm for model weight updates, which are then returned to the mobile application for system updates. A clustering algorithm was applied in the transfer learning process, similar to the technology mentioned in research (33), which helped set a threshold. This threshold allowed individual users to understand the characteristics of each label better. Through this personalized model weight update process, the more each user utilizes the app, the more tailored the model becomes, eventually leading to an optimized breathing threshold.

All these developments and refinements were achieved using Python 3.9.17 and Tensorflow 2.10.1. The final iteration of the system was built with Unity 2021.3.5f1 and optimized for a platform equipped with an Intel Core i7-10700 processor and NVIDIA GeForce RTX3070 graphics.

3.3 Experiment setting

To validate the effectiveness and usability of the proposed deep-breathing mobile application, we conducted an experiment with 70 participants. These participants were randomly assigned to either the intervention group, which used our developed system, or the control group, which used an existing deep-breathing mobile application that provides an image to guide the breathing rhythm. The experiment incorporated HRV (Heart Rate Variability) measurements, taken with the ubpulse T1 device developed by LAXTHA, to assess the relaxation effects of deep breathing. Initial baseline HRV measurements were recorded, followed by a one-minute deep breathing session using the respective applications. After the session, HRV measurements were taken again. Following the HRV measurement, participants completed a user experience survey to provide feedback.

Subsequently, our user experience survey was based on the Technology Acceptance Model (TAM). This model emphasizes perceived ease of use and perceived usefulness as key determinants of user attitudes and intentions towards adopting new technologies (34). The survey also incorporated measures of perceived enjoyment, especially considering the gamification elements integrated into our proposed system (35). The primary objective of this survey was to gauge user experiences and their intentions to continue using the deep breathing mobile application over time.

4 Results

In our experiment, we enrolled a total of 70 participants. Both the intervention group and the control group consisted of 35 individuals each. The participants had an average age of 33, including 39 females and 31 males. They were meticulously selected, ensuring none had specific heart conditions or were on medications that might influence the Heart Rate Variability (HRV) measurement.

4.1 Statistical analysis of HRV metrics

Our study revealed significant changes in the Heart Rate Variability (HRV) indices following the deep breathing exercises.

These indices, including normalized low frequency (normLF), normalized high frequency (normHF), Standard Deviation of NN Interval (SDNN), Total Power (TP), and HRV, are pivotal for monitoring the autonomic nervous system's functionality and assessing resilience to stress.

For the intervention group, the baseline for normLF was ($M = 53.58$, $SD = 4.44$). This value increased to ($M = 58.10$, $SD = 2.28$) during the deep breathing exercises. In tandem with this, the normHF values saw a decrease from its baseline of ($M = 46.42$, $SD = 4.43$) to ($M = 41.89$, $SD = 2.28$) during the exercises.

For the control group, the normLF values increased from a baseline of ($M = 52.45$, $SD = 5.56$) to ($M = 55.47$, $SD = 2.20$) during the exercises. Similarly, normHF values decreased from its baseline ($M = 47.54$, $SD = 5.56$) to ($M = 44.52$, $SD = 2.20$).

The t-test revealed significant p-values for changes in normLF and normHF for both groups, each being p 0.05. This indicates that during deep breathing, there's an increase in sympathetic nerve activity due to lung movement and a corresponding decrease in parasympathetic activity. Notably, the rise in sympathetic nerve activity before and after the deep breathing exercise in the intervention group was significantly pronounced, registering a difference of $+4.52$ ($t = -5.353$, $p < 0.001$) in comparison to the control group, which had a difference of $+3.02$ ($t = -2.986$, $p = 0.004$). This observation suggests that participants in the intervention group were more engaged, resulting in heightened sympathetic activity (36).

Following the deep breathing exercise, the LF and HF values were measured once again during the recovery phase. The intervention group displayed normLF values of ($M = 56.22$, $SD = 4.69$) and normHF values of ($M = 43.77$, $SD = 4.69$). On the other hand, the control group registered normLF values of ($M = 54.38$, $SD = 5.28$) and normHF values of ($M = 45.61$, $SD = 5.28$). When juxtaposed with the measurements taken during the deep breathing phase, it was evident that both groups exhibited a drop in normLF and a rise in normHF, signaling relaxation.

Delving deeper into the specifics, the intervention group witnessed a statistically significant decline in normLF by -1.88 ($t = 2.126$, $p = 0.038$) and a corresponding increase in normHF by $+1.88$ ($t = -2.126$, $p = 0.038$). Conversely, the control group noted a drop in normLF by -1.09 ($t = 1.13$, $p = 0.264$) and an increase in normHF by $+1.09$ ($t = -1.13$, $p = 0.264$), but neither of these changes was statistically significant. This indicates that the intervention group, who were more proactive during deep breathing, experienced pronounced activation of the sympathetic nerves and succeeded by activation of the parasympathetic nerves. However, the control group, which was relatively less engaged in deep breathing, did not exhibit a substantial relaxation response afterward.

Subsequent measurements encompassed additional HRV metrics, notably SDNN, TP, and HRV. To discern the significance of these fluctuations, t-test evaluations were executed. For the intervention group, the variations manifested as follows: SDNN had a difference of $+24.08$ ($t = -5.894$, $p < 0.001$), TP by $+1.07$ ($t = -5.118$, $p < 0.001$), and HRV by $+3.91$ ($t = -3.622$, $p < 0.001$), all being statistically significant. In the control group, SDNN changed by $+15.98$ ($t = -3.309$, $p = 0.001$), TP by $+0.88$ ($t = -3.702$, $p < 0.001$), and HRV by $+3.4$ ($t = -2.664$, $p = 0.009$). Though these changes were significant, the scope and amplitude of these modifications were more prominent in the intervention group.

These findings underscore significant shifts in markers of autonomic nervous system activity, notably including the Standard Deviation of the NN Interval (SDNN), TP, and HRV, during deep breathing exercises. SDNN, as a sensitive marker within HRV, plays a pivotal role in indicating an individual's stress response and overall autonomic nervous system balance (37). A lower SDNN is often associated with reduced stress resilience and poorer mental health outcomes. Thus, the observed reduction in SDNN in our intervention group is particularly meaningful, suggesting that our deep breathing system could be more effective than conventional techniques in enhancing mental health care. These markers not only trace changes in the autonomic nervous system but also reflect an individual's resilience to stress (38). The more pronounced alterations in the intervention group as opposed to the control group indicate that sustained use of our system might amplify relaxation and bolster stress resilience by positively modulating the autonomic nervous system. This offers valuable support in managing mental health and demonstrates the practical implications of our findings. A detailed breakdown of the HRV measurements is available in Table 1.

4.2 Usability analysis

Our analysis extended beyond evaluating the efficacy of the deep breathing system; we delved into its usability as well. A preliminary reliability analysis was performed using Cronbach's alpha for the gathered results.

The user acceptance of our technology was assessed based on the primary constructs of the Technology Acceptance Model (TAM), specifically: perceived ease of use (PEOU), perceived usefulness (PU), attitude towards using (ATT), and intention to use (ITU). Additionally, we considered perceived enjoyment (PENJ) given its relevance. The reliability metrics for each construct were: PEOU($\alpha = 0.86$), PU($\alpha = 0.73$), PENJ ($\alpha = 0.93$), ATT($\alpha = 0.83$), and ITU($\alpha = 0.92$). All the Cronbach's alpha values surpassed the 0.7 threshold, solidifying the reliability of our measures.

Breaking down the results for the intervention group: PEOU scored ($M = 6.10$, $SD = 0.99$), PU at ($M = 5.41$, $SD = 1.07$), PENJ reached ($M = 5.91$, $SD = 0.98$), ATT scored ($M = 5.80$, $SD = 0.89$), and ITU stood at ($M = 5.19$, $SD = 1.24$). In contrast, the control group results were: PEOU ($M = 6.24$, $SD = 0.97$), PU ($M = 5.28$, $SD = 0.99$), PENJ ($M = 5.00$, $SD = 1.37$), ATT ($M = 5.24$, $SD = 1.18$), and ITU ($M = 4.87$, $SD = 1.70$). Interestingly, aside from PEOU, all metrics indicated a preference for the intervention system.

In the subsequent t-test, only PENJ ($t = 3.175$, $p = 0.002$) and ATT ($t = 2.237$, $p = 0.028$) showcased significant disparities between the groups. However, when introducing a regression analysis, considering ITU as the dependent variable, both PENJ

TABLE 1 Statistical Analysis result of Intervention group and Control group.

	Frequency domain		Time domain		
	2*norm LF	2*norm HF	2*SDNN	2*TP	2*HRV
Intervention group - Mean (SD)					
Baseline	53.58 (4.44)	46.42 (4.43)	40.57 (18.20)	7.12 (0.92)	11.83 (4.27)
difference	+4.52	−4.53	+24.08	+1.07	+3.91
	($p < 0.001$)	**($p < 0.001$)**	**($p = 1.36e-07$)**	**($p < 0.001$)**	**($p < 0.001$)**
Relaxation	58.10 (2.28)	41.89 (2.28)	64.65 (15.90)	8.19 (0.82)	15.74 (4.75)
difference	−1.88	+1.88	−27.08	−1.25	−5.31
	($p = 0.038$)	**($p = 0.038$)**	**($p < 0.001$)**	**($p < 0.001$)**	**($p < 0.001$)**
Recovery	56.22 (4.69)	43.77 (4.69)	37.57 (17.46)	6.94 (0.98)	10.43 (3.44)
Control Group - Mean (SD)					
Baseline	52.45 (5.56)	47.54 (5.56)	48.82 (21.76)	7.43 (1.08)	12.90 (5.33)
difference	+3.02	−3.02	+15.98	+0.88	+3.40
	($p = 0.004$)	**($p = 0.004$)**	**($p = 0.001$)**	**($p < 0.001$)**	**($p = 0.009$)**
Relaxation	55.47 (2.20)	44.52 (2.20)	64.80 (18.47)	8.31 (0.91)	16.30 (5.33)
difference	−1.09	+1.09	−21.16	−1.06	−3.72
	($p = 0.264$)	($p = 0.264$)	**($p < 0.001$)**	**($p < 0.001$)**	**($p < 0.001$)**
Recovery	54.38 (5.28)	45.61 (5.28)	43.74 (15.74)	7.25 (0.77)	12.58 (3.95)

Bold indicates significance at p-value < 0.05.

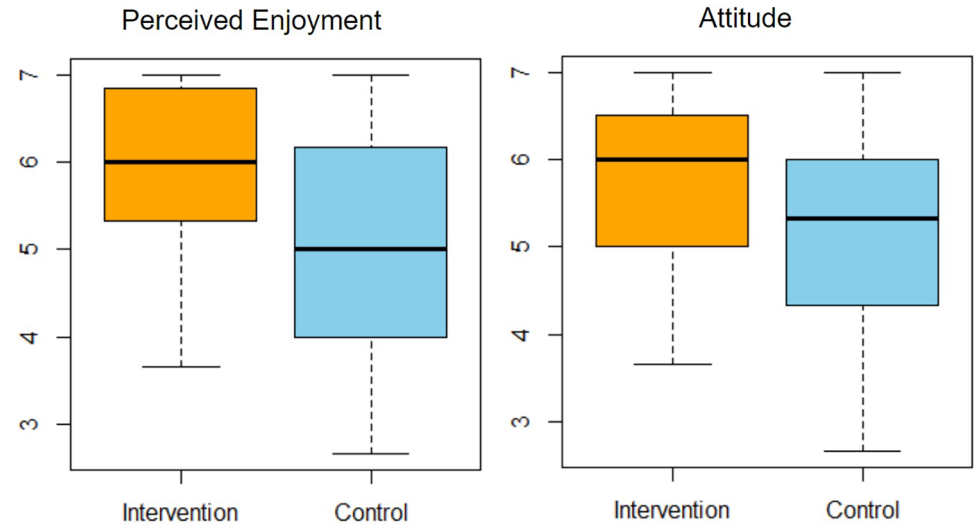


FIGURE 6
Boxplot of perceived enjoyment and attitude.

($t = 7.743$, $p < 0.001$) and ATT ($t = 7.516$, $p < 0.001$) emerged significantly. This suggests that PENJ and ATT might play pivotal roles in determining the long-term inclination towards our deep breathing system. These findings are graphically represented in Figure 6.

5 Discussion

In summarizing our findings regarding the efficacy and usability of our evolving deep breathing system in the realm of digital healthcare and Digital Therapeutics for mental well-being, our study illuminates several key aspects and future potentials.

The integration of gamification in our system, characterized by elements such as enjoyment and interactive feedback, has significantly improved participant commitment to deep breathing exercises. Furthermore, the application of machine learning, particularly the use of Mel Frequency Cepstral Coefficient (MFCC) for real-time assessment and feedback of breathing states, has shown promising results in enhancing the effectiveness of our deep breathing exercises. Our results showed his technological advancement is pivotal in creating a more personalized and responsive therapeutic experience.

Moreover, our study's utilization of Heart Rate Variability (HRV) as a dynamic indicator of the autonomic nervous system's functionality has provided invaluable insights. The analysis of HRV data, through frequency and time domain methods, has

helped us understand the physiological impact of our deep breathing exercises on mental well-being. This comprehensive measurement approach underpins the potential effectiveness of our interventions in managing mental health conditions like anxiety, stress, and mood disorders.

Looking towards the future, we envision our system's continued refinement leading to more widespread and effective use in DTx applications. Its ability to actively engage the autonomic nervous system suggests that it could be applied to a broader range of mental health conditions, potentially offering a non-invasive, user-friendly alternative to more traditional therapies (39). The observed patterns in Low Frequency (LF) and High Frequency (HF) values during exercises indicate a direct impact on the autonomic nervous system, providing a basis for exploring its application in managing conditions like anxiety, stress, and mood disorders (40).

In terms of usability, our system's positive reception compared to traditional deep breathing exercises positions it as a more engaging and effective tool for mental health management. This user-friendly approach, coupled with the potential for sustained use, underlines our system's capacity to deliver long-term therapeutic benefits (41). It also opens avenues for integration with other digital health applications, enhancing overall mental health care efficacy (42).

Overall, the implications of our research extend beyond immediate findings, suggesting a transformative potential for digital therapeutics in mental health management, both now and in the future.

5.1 Limitation & future works

The present study mainly pivoted towards optimizing for mobile platforms, mirroring the digital health industry's trajectory. Consequently, it might not have fully capitalized on the potential of existing computer technology. For instance, our study employed a model updating framework akin to federated learning for personalizing the classification model. However, inherent limitations in mobile-based self-learning prompted us to rely on a method that sends user data to a central server. This aspect highlights the necessity for subsequent research, especially as mobile technologies continue to evolve. Moreover, our system design didn't manage to incorporate every suggested design element derived from an affordance-based standpoint, primarily because our focus was on initial validation in the absence of any affordance-based designed breathing systems. We foresee more refined research addressing these gaps. Finally, while the gender and age distribution of our participants provided initial insights, the limited sample size of 70 may constrain the generalizability of our findings. Future studies with larger and more diverse samples are essential to validate and extend our findings.

6 Conclusion

In our research, we introduced and assessed a deep breathing system tailored for digital healthcare and Digital Therapeutics contexts. Our findings underscore the system's potential to substantially enhance mental health by fostering engaging and personalized user interactions. The integration of gamification and machine learning-driven personalized feedback has successfully motivated users to be more proactive during the deep breathing exercises. This heightened involvement manifests physiologically, evidenced by the rise in LF values during exercises, followed by an increase in HF, signifying exertion and subsequent relaxation effects. Moreover, the data suggest that our approach can invigorate the autonomic nervous system, potentially boosting stress resilience and further emphasizing the therapeutic promise of our solution.

From a usability perspective, participants expressed a marked preference for our system over traditional deep breathing alternatives, recounting a more delightful and affirmative experience. This enthusiasm is a robust testament to the system's potential for consistent use, a crucial aspect for realizing enduring therapeutic advantages.

Despite these positive strides, the study illuminated avenues for refinement and future exploration. Prioritizing the full spectrum of computer technology, embedding more affordance-driven design facets, and tapping into a wider participant demographic emerged as the study's limitations. These identified gaps are envisioned to steer forthcoming advancements and scholarly pursuits, paving the way for even more sophisticated digital therapeutic instruments dedicated to mental wellness.

In summation, our innovative deep breathing system represents a seminal advancement in digital healthcare. It not only acts as a potent stress alleviation tool but also cultivates a culture of active user participation, leading to enhanced mental well-being. The challenges and prospective research pathways spotlighted in our study offer exciting prospects for refinement and breakthroughs in this swiftly progressing sector.

Data availability statement

The datasets for this article are not publicly available due to concerns regarding participant anonymity. Requests to access the datasets should be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by Sungkyunkwan University Bioethics Committee. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

HJ: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review & editing; JHY: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Validation, Writing – original

draft, Writing – review & editing; MG: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing; HS: Conceptualization, Formal Analysis, Investigation, Methodology, Project administration, Validation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article.

This work was founded by the Technology Innovation Program (or Industrial Strategic Technology Development Program-Source Technology Development and Commercialization of Digital Therapeutics) (20014967, Development of Digital Therapeutics for Depression from COVID19) funded By the Ministry of Trade, Industry & Energy (MOTIE, Korea).

Conflict of interest

Author JHY is employed by Hippo T&C.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Torous J, Myrick KJ, Rauseo-Ricupero N, Firth J. Digital mental health, COVID-19: using technology today to accelerate the curve on access, quality tomorrow. *JMIR Ment Health*. (2020) 7:e18848. doi: 10.2196/18848

2. Cho C-H, Lee H-J. Could digital therapeutics be a game changer in psychiatry?. *Psychiatry Investig*. (2019) 16:97. doi: 10.30773/pi.2019.01.20

3. Hong JS, Wasden C, Han DH. Introduction of digital therapeutics. *Comput Methods Programs Biomed*. (2021) 209:106319. doi: 10.1016/j.cmpb.2021.106319

4. Kadakia K, Patel B, Shah A. Advancing digital health: FDA innovation during COVID-19. *Npj Digit Med*. (2020) 3:1–3. doi: 10.1038/s41746-020-00371-7

5. Zhu B, Hedman A, Feng S, Li H, Osika W. Designing, prototyping and evaluating digital mindfulness applications: a case study of mindful breathing for stress reduction. *J Med Internet Res*. (2017) 19:e197. doi: 10.2196/jmir.6955

6. Chien H-C, Chung Y-C, Yeh M-L, Lee J-F. Breathing exercise combined with cognitive behavioural intervention improves sleep quality and heart rate variability in major depression. *J Clin Nurs*. (2015) 24:3206–14. doi: 10.1111/jocn.12972

7. Alsawaier RS. The effect of gamification on motivation and engagement. *Int J Inf Learn Technol*. (2018) 35(1):56–79.

8. Sardi L, Idri A, Fernández-Alemán JL. A systematic review of gamification in e-Health. *J Biomed Inform*. (2017) 71:31–48. doi: 10.1016/j.jbi.2017.05.011

9. Oyebode O, Fowles J, Steeves D, Orji R. Machine learning techniques in adaptive, personalized systems for health, wellness. *Int J Hum Comput Interact*. (2023) 39 (9):1938–62.

10. Khaddage F, Lattemann C, Acosta-Díaz R. Mobile gamification in education engage, educate, entertain via gamified mobile apps. In: *Society for Information Technology & Teacher Education International Conference*. Association for the Advancement of Computing in Education (AACE) (2014). p. 1654–60.

11. Stepanovic S, Mettler T. Gamification applied for health promotion: does it really foster long-term engagement? a scoping review. In: *Proceedings of the 26th European Conference on Information Systems*. AIS (2018). p. 1–6.

12. Hammedi W, Leclerq T, Van Riel AC. The use of gamification mechanics to increase employee, user engagement in participative healthcare services: a study of two cases. *J Serv Manag*. (2017) 28:640–61. doi: 10.1108/JOSM-04-2016-0116

13. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. *JMIR Ser Games*. (2013) 1:e3139. doi: 10.2196/games.3139

14. Brown M, O'Neill N, van Woerden H, Eslambolchilar P, Jones M, John A. et al. Gamification, adherence to web-based mental health interventions: a systematic review. *JMIR Ment Health*. (2016) 3:e5710. doi: 10.2196/mental.5710

15. Agrawal V, Naik V, Duggirala M, Athavale S. Calm a mobile based deep breathing game with biofeedback. In: *Extended Abstracts of the 2020 Annual Symposium on Computer–Human Interaction in Play* (2020). p. 153–7.

16. Lukic YX, Teepe GW, Fleisch E, Kowatsch T. Breathing as an input modality in a gameful breathing training app (breeze 2): development, evaluation study. *JMIR Ser Games*. (2022) 10:e39186. doi: 10.2196/39186

17. Shih C-H, Tomita N, Lukic YX, Kowatsch T. Breeze: smartphone-based acoustic real-time detection of breathing phases for a gamified biofeedback breathing training. *Proc ACM Interact Mob Wearable Ubiquitous Technol*. (2019) 3:1–30. doi: 10.1145/3369835

18. Lupton D. The digitally engaged patient: self-monitoring and self-care in the digital health era. *Soc Theory Health*. (2013) 11:256–70. doi: 10.1057/sth.2013.10

19. Yoo JH, Jeong H, Chung T-M. Cutting-edge technologies for digital therapeutics: a review and architecture proposals for future directions. *Appl Sci*. (2023) 13:6929. doi: 10.3390/app13126929

20. Masoudi N, Fadel GM, Pagano CC, Elena MV. A review of affordances and affordance-based design to address usability. In: *Proceedings of the Design Society: International Conference on Engineering Design*. Cambridge University Press, vol. 1 (2019). p. 1353–62.

21. Wong CA, Madanay F, Ozer EM, Harris SK, Moore M, Master SO. et al. Digital health technology to enhance adolescent and young adult clinical preventive services: affordances and challenges. *J Adolesc Health*. (2020) 67:S24–S33. doi: 10.1016/j.jadohealth.2019.10.018

22. Cho Y, Bianchi-Berthouze N, Julier SJ. Deepbreath: deep learning of breathing patterns for automatic stress recognition using low-cost thermal imaging in unconstrained settings. In: *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE (2017). p. 456–63.

23. Cochrane KA, Loke L, Ahmadpour N, Schiphorst T, Campbell A, Núñez-Pacheco C. A comparison design study of feedback modalities to support deep breathing whilst performing work tasks. *Work*. (2021) 68:1187–202. doi: 10.3233/WOR-213448

24. Han W, Chan C-F, Choy C-S, Pun K-P. An efficient MFCC extraction method in speech recognition. In: *2006 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE (2006). p. 4.

25. Hossan MA, Memon S, Gregory MA. A novel approach for MFCC feature extraction. In: *2010 4th International Conference on Signal Processing and Communication Systems*. IEEE (2010). p. 1–5.

26. Sadi TM, Hassan R. Development of classification methods for wheeze and crackle using mel frequency cepstral coefficient (MFCC): A deep learning approach. *Int J Percept Cogn Comput*. (2020) 6:107–14. doi: 10.31436/ijpcc.v6i2.166

27. Kemp AH, Quintana DS. The relationship between mental and physical health: insights from the study of heart rate variability. *Int J Psychophysiol*. (2013) 89:288–96. doi: 10.1016/j.ijpsycho.2013.06.018

28. Appelhans BM, Luecken LJ. Heart rate variability as an index of regulated emotional responding. *Rev Gen Psychol*. (2006) 10:229–40. doi: 10.1037/1089-2680.10.3.229

29. Prinsloo GE, Derman WE, Lambert MI, Laurie Rauch H. The effect of a single session of short duration biofeedback-induced deep breathing on measures of heart rate variability during laboratory-induced cognitive stress: a pilot study. *Appl Psychophysiol Biofeedback*. (2013) 38:81–90. doi: 10.1007/s10484-013-9210-0

30. Springer A, Whittaker S. Progressive disclosure: empirically motivated approaches to designing effective transparency. In: *Proceedings of the 24th International Conference on Intelligent User Interfaces* (2019). p. 107–20.

31. Yu K, Berkovsky S, Taib R, Conway D, Zhou J, Chen F. User trust dynamics: an investigation driven by differences in system performance. In: *Proceedings of the 22nd International Conference on Intelligent User Interfaces* (2017). p. 307–17.

32. Madikeri SR, Murthy HA. Mel filter bank energy-based slope feature and its application to speaker recognition. In: *2011 National Conference on Communications (NCC)*. IEEE (2011). p. 1–4.

33. Yoo JH, Son HM, Jeong H, Jang E-H, Kim AY, Yu HY. et al. Personalized federated learning with clustering: non-iid heart rate variability data application. In: *2021 International Conference on Information and Communication Technology Convergence (ICTC)*. IEEE (2021). p. 1046–51.

34. Wallace LG, Sheetz SD. The adoption of software measures: A technology acceptance model (TAM) perspective. *Inf Manag*. (2014) 51:249–59. doi: 10.1016/j.im.2013.12.003

35. Hussain A, Mkpojiogu EO, Yusof MM. Perceived usefulness, perceived ease of use, and perceived enjoyment as drivers for the user acceptance of interactive mobile maps. In: *AIP Conference Proceedings*. AIP Publishing, vol. 1761 (2016).

36. Reyes del Paso GA, Langewitz W, Mulder LJ, Van Roon A, Duschek S. The utility of low frequency heart rate variability as an index of sympathetic cardiac tone: a review with emphasis on a reanalysis of previous studies. *Psychophysiology*. (2013) 50:477–87. doi: 10.1111/psyp.12027

37. Kim H-G, Cheon E-J, Bai D-S, Lee YH, Koo B-H. Stress and heart rate variability: a meta-analysis and review of the literature. *Psychiatry Investig*. (2018) 15:235. doi: 10.30773/pi.2017.08.17

38. Jarczok MN, Jarczok M, Mauss D, Koenig J, Li J, Herr RM. et al. Autonomic nervous system activity and workplace stressors—a systematic review. *Neurosci Biobehav Rev*. (2013) 37:1810–23. doi: 10.1016/j.neubiorev.2013.07.004

39. D'Ostilio K, Magis D. Invasive and non-invasive electrical pericranial nerve stimulation for the treatment of chronic primary headaches. *Curr Pain Headache Rep*. (2016) 20:1–9. doi: 10.1007/s11916-016-0589-1

40. Nakajima Y, Tanaka N, Mima T, Izumi S-I. Stress recovery effects of high-and low-frequency amplified music on heart rate variability. *Behav Neurol*. (2016) 2016:1–9. doi: 10.1155/2016/5965894

41. Bickmore TW, Picard RW. Establishing and maintaining long-term human-computer relationships. *ACM Trans Comput Hum Interact*. (2005) 12:293–327. doi: 10.1145/1067860.1067867

42. Dang A, Arora D, Rane P. Role of digital therapeutics and the changing future of healthcare. *J Family Med Prim Care*. (2020) 9:2207. doi: 10.4103/jfmpc.jfmpc_105_20

frontiers | Frontiers in Digital Health

TYPE Original Research
PUBLISHED 14 February 2024
DOI 10.3389/fdgth.2024.1297935

OPEN ACCESS

EDITED BY
Nikki Holliday,
Coventry University, United Kingdom

REVIEWED BY
Abid Hasan Khan,
Bangabandhu Sheikh Mujib Medical University (BSMMU), Bangladesh
Syeda Fatema Alam,
University of New South Wales, Australia

*CORRESPONDENCE
Stefan Rennick-Egglestone
✉ stefan.egglestone@nottingham.ac.uk

†These authors have contributed equally to this work

RECEIVED 20 September 2023
ACCEPTED 30 January 2024
PUBLISHED 14 February 2024

CITATION
Ali Y, Rennick-Egglestone S, Llewellyn-Beardsley J, Ng F, Yeo C, Franklin D, Perez Vallejos E, Ben-Zeev D, Kotera Y and Slade M (2024) Perception and appropriation of a web-based recovery narratives intervention: qualitative interview study. Front. Digit. Health 6:1297935. doi: 10.3389/fdgth.2024.1297935

Perception and appropriation of a web-based recovery narratives intervention: qualitative interview study

Yasmin Ali[1†], Stefan Rennick-Egglestone[1*†], Joy Llewellyn-Beardsley[1], Fiona Ng[1], Caroline Yeo[1,2], Donna Franklin[3,4], Elvira Perez Vallejos[5,6], Dror Ben-Zeev[7], Yasuhiro Kotera[1,8] and Mike Slade[1,9]

[1]School of Health Sciences, Institute of Mental Health, University of Nottingham, Nottingham, United Kingdom, [2]Department of Architecture and Built Environment, Faculty of Engineering, University of Nottingham, Nottingham, United Kingdom, [3]NEON Lived Experience Advisory Panel, Nottingham, United Kingdom, [4]School of Health and Related Research, University of Sheffield, Sheffield, United Kingdom, [5]School of Medicine, University of Nottingham, Nottingham, United Kingdom, [6]School of Computer Science, University of Nottingham, Nottingham, United Kingdom, [7]Department of Psychiatry and Behavioral Sciences, School of Medicine, University of Washington, Seattle, WA, United States, [8]Center for Infectious Disease Education and Research, Osaka University, Osaka, Japan, [9]Faculty of Nursing and Health Sciences, Health, and Community Participation Division, Nord University, Namsos, Norway

Introduction: Mental health recovery narratives are widely available to the public, and can benefit people affected by mental health problems. The NEON Intervention is a novel web-based digital health intervention providing access to the NEON Collection of recovery narratives. The NEON Intervention was found to be effective and cost-effective in the NEON-O Trial for people with nonpsychosis mental health problems (ISRCTN63197153), and has also been evaluated in the NEON Trial for people with psychosis experience (ISRCTN11152837). We aimed to document NEON Intervention experiences, through an integrated process evaluation.

Methods: Analysis of interviews with a purposive sample of intervention arm participants who had completed trial participation.

Results: We interviewed 34 NEON Trial and 20 NEON-O Trial participants (mean age 40.4 years). Some users accessed narratives through the NEON Intervention almost daily, whilst others used it infrequently or not at all. Motivations for trial participation included: exploring the NEON Intervention as an alternative or addition to existing mental health provision; searching for answers about mental health experiences; developing their practice as a mental health professional (for a subset who were mental health professionals); claiming payment vouchers. High users (10 + narrative accesses) described three forms of appropriation: distracting from difficult mental health experiences; providing an emotional boost; sustaining a sense of having a social support network. Most participants valued the scale of the NEON Collection (n = 659 narratives), but some found it overwhelming. Many felt they could describe the characteristics of a *desired narrative* that would benefit their mental health. Finding a narrative meeting their desires enhanced engagement, but not finding one reduced engagement. Narratives in the NEON Collection were perceived as authentic if they acknowledged the difficult reality of mental health experiences, appeared to describe real world experiences, and described mental health experiences similar to those of the participant.

10.3389/fdgth.2024.1297935

Discussion: We present recommendations for digital health interventions incorporating collections of digital narratives: (1) make the scale and diversity of the collection visible; (2) provide delivery mechanisms that afford appropriation; (3) enable contributors to produce authentic narratives; (4) enable learning by healthcare professionals; (5) consider use to address loneliness.

KEYWORDS

digital health intervention, online intervention, psychosis, recovery narrative, recovery story, lived experience narrative, autobiography, NEON Intervention

1 Introduction

Narratives describing recovery from mental health problems are readily available to the public. Whilst many are presented as written autobiographies (1), we have also encountered forms such as visual art, poetry, audio recordings, video recordings, and a mixture of modalities in one narrative (2). Llewellen-Beardsley (3) has argued that mental health recovery narratives are a distinctive genre, and Llewellyn-Beardsley and colleagues (4) have advocated for an inclusive recovery narrative definition, which encompasses narratives expressing the struggles and/or adversities that a narrator has experienced, alongside their strengths, successes, and survival. This is congruent with a longstanding orientation in which the term "recovery" is defined as "a way of living a satisfying, hopeful, and contributing life even within the limitations caused by illness" (5). Whilst our focus is mental health recovery narratives, we are aware that recovery narratives are regularly published to describe a range of health experiences, including for diabetes (6) and cancer recovery (7).

Recovery narratives have been used to create individual and societal benefits. A systematic review on the uses of mental health recovery narratives (8) identified 27 different uses, which were categorised as political (e.g., supporting policy change), societal (e.g., reducing mental health stigma), community (e.g., drawing attention to mental health concerns in a particular community), service level (e.g., improving mental health and social care services), and individual (e.g., as a therapeutic tool in an individual intervention). In mental health, some Digital Health Interventions (DHIs) have included recovery narratives as a proposed *active ingredient*, e.g., one of the proposed mechanisms by which the DHI improves outcome. For example, the Self-Management And Recovery Technology (SMART) study developed a web-based intervention presenting videos in which people with lived experiences of psychosis reflected on their recovery, alongside other supportive material such as health information (9). An interview study found that SMART access helped recipients who had experienced psychosis to gain a renewed belief in achieving recovery. Accessing SMART also helped participants to feel less alone, and more connected, hopeful, and inspired (9). Recovery narratives have also been used in interventions outside of mental health (10–12), including in a randomised controlled trial (RCT) which found that self-efficacy in achieving weight loss significantly increased in participants who had read indexed narratives compared to a control group (11).

As well as being available individually, recovery narratives can be grouped together into printed or online collections, often grouped by topic. For example, a published book of psychosis recovery narratives sought to create hope for the future, for people experiencing psychosis, and for those who offer care to them (13). National anti-stigma campaigns such as Time to Change in the UK (14) and Here to Help in Canada (15) have provided web-based access to large collections of recovery narratives with the intention of reducing mental health stigma (15). In the UK, it has become a common practice for units in the National Health Service (NHS) to share collections of service user recovery narratives, for example to support the understanding of mental health conditions and recovery by other service users (16). The people who produce narrative collections can be thought of as active participants in the selection, processing, and organisation of the collection (17), and hence can influence how mental health is seen and understood (2). In interview study with collection organisers has shown that narratives are often selected to create a desired influence (8).

1.1 The narrative experiences online intervention

The Narrative Experiences Online (NEON) study https://www.researchintorecovery.com/research/neon has developed the NEON Intervention, a web-based DHI providing access to a collection of 659 international mental health recovery narratives, donated by individuals and from existing collections (18). Each narrative in the NEON Collection is characterized using the 77-item Inventory of Characteristics of Recovery Stories (INCRESE) (19). The development process for the NEON Intervention began by establishing a verifiable theory (20) on the characteristics (4, 21) and recipient impact (22–25) of recovery narratives. A prototype web-application was co-produced, and then iteratively improved through a feasibility study with mental health service users, including through interviews collecting immediate responses to prototype features, and reflective responses after one month of use (20). Safety strategies were developed with academic and lived experience advice (20). They included the use of content warnings (26) before access to narratives containing content with the potential to cause distress, the ability to block (and unblock) individual narratives, an information page presenting both service signposting and self-management information, and the ability to rapidly exit the NEON Intervention (for example, if

participants feared stigma when using the NEON Intervention in a public place). As we developed the NEON Intervention, we were conscious that first contact with a healthcare technology can be challenging for people experiencing mental health problems (27). Hence to support new users to learn about the NEON Intervention, each participant was shown a narrative from a small set of narratives empirically identified as hope-inspiring for participants in a previous study (20), immediately after all online trial procedures and a personal profile had been completed. Participant format preferences recorded in the personal profile were respected e.g., a video first narrative was shown to participants not wanting text-based narratives. This was to aid inclusion for people experiencing difficulties processing specific narrative forms, e.g., dyslexia.

The central feature of the NEON Intervention is a homepage providing users with five narrative access mechanisms (Figure 1):

(1) Match me to a story: requests that a recommender system selects a narrative. Recommender systems are algorithms designed to match digital media material to users (28).
(2) Get me a random story: requests a randomly selected narrative.
(3) Browse stories: enables a user to browse NEON Collection narratives, by selecting from categories relating to characteristics of the narrator (Figure 2) or narrative content (Figure 3).
(4) My stories—bookmarked: enables a user to request a previously-saved narrative
(5) My Stories—hopeful: enables a user to request a previously viewed narrative which they rated as high on hopefulness.

NEON has conducted two pragmatic randomised controlled trials of the NEON Intervention with different populations, which have been designed to produce a definitive result (18, 29). The NEON-O Trial included participants who experienced mental health problems other than psychosis in the last 5 years ($N = 1{,}023$, https://www.isrctn.com/ISRCTN63197153). It found that NEON Intervention access was effective at increasing quality of life, increasing the presence of meaning in life, and was cost-effective from the perspective of the statutory health and social care system in England (30). The NEON Trial ($N = 739$, https://www.isrctn.com/ISRCTN11152837) included participants experiencing current mental health distress, and who had experienced psychosis in the last 5 years. An evaluation of effectiveness and cost-effectiveness will be reported elsewhere. DHIs can extend health service provision to people not currently using services, and hence for external validity, both trials

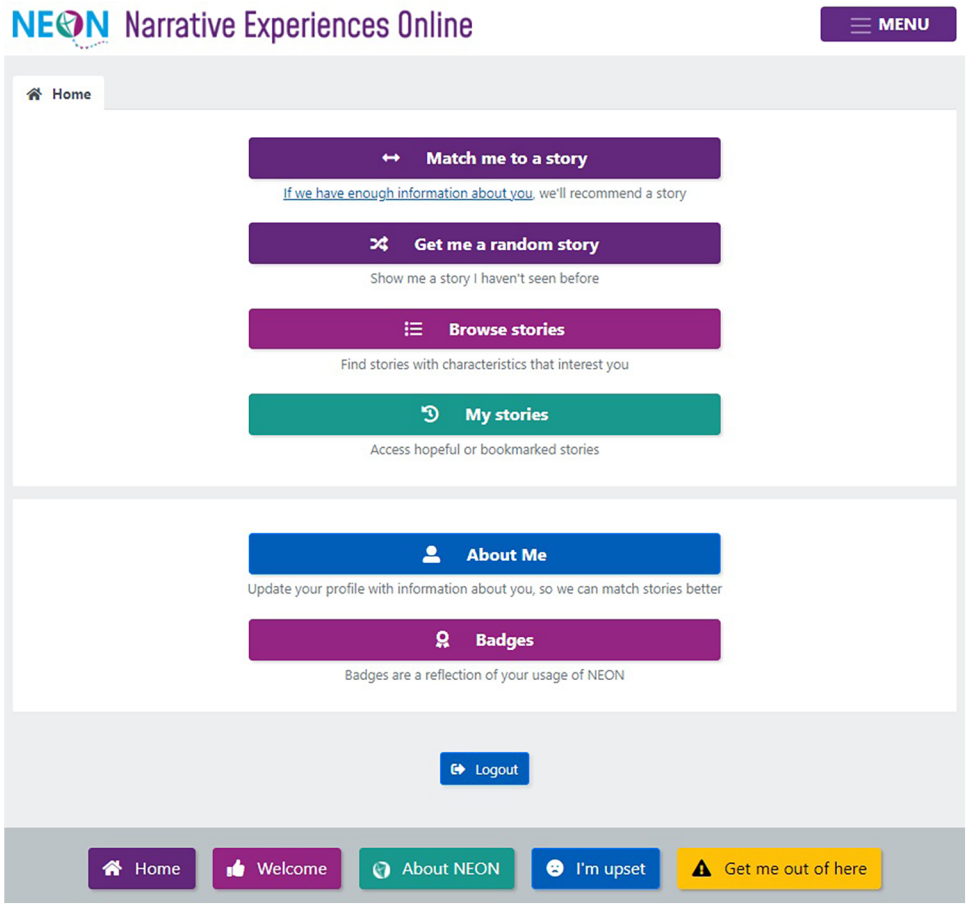


FIGURE 1
NEON Intervention home page.

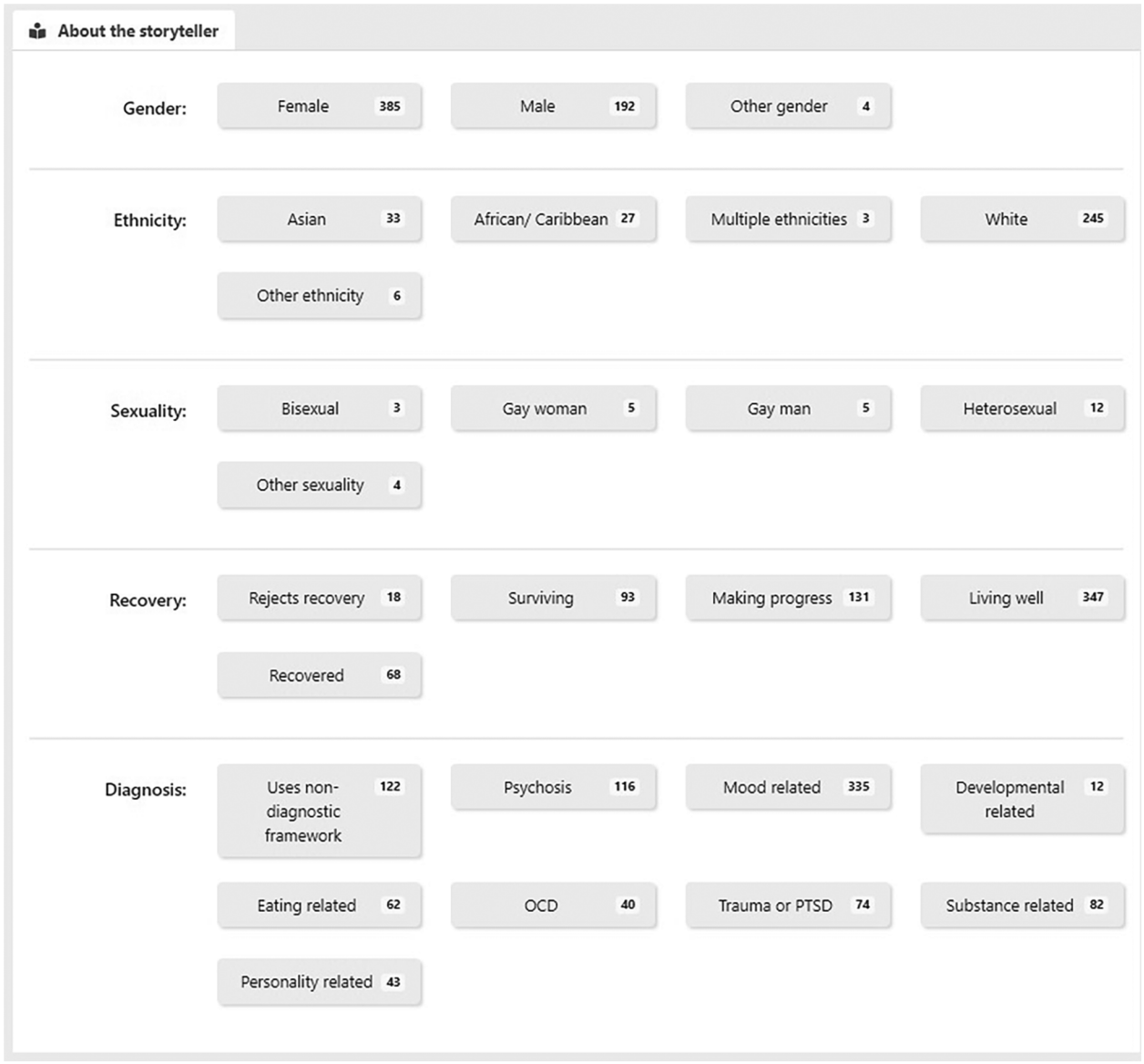


FIGURE 2
Characteristics of the narrator that can be used to filter for narratives.

recruited people who had used mental health services and those who had not (31).

Prior research in mental health has established the importance of documenting user experiences of DHIs (32), for example to generate knowledge to inform iterative development of the same and related DHIs. The aim of the current study was therefore to understand the experiences of NEON Trial and NEON-O Trial participants in registering for the NEON Trial and NEON-O Trial, and in using the NEON Intervention. The objectives for the analysis presented in this paper are:

1. to identify motivations for trial participation and intervention usage
2. to document how the NEON Intervention was appropriated
3. to document perceptions of the NEON Collection that influenced usage

Our selection of these foci was informed by prior user experience and DHI evaluation research. How to create technologies that motivate engagement has been identified as a critical topic for DHI developers with the potential to influence how DHIs are used (33). Appropriation refers to the processes by which newly acquired technologies become integrated into the everyday lives of their users (34), sometimes becoming a normal feature of those lives (35, 36), and hence effectively becoming "invisible" (37). A better understanding of how interventions are

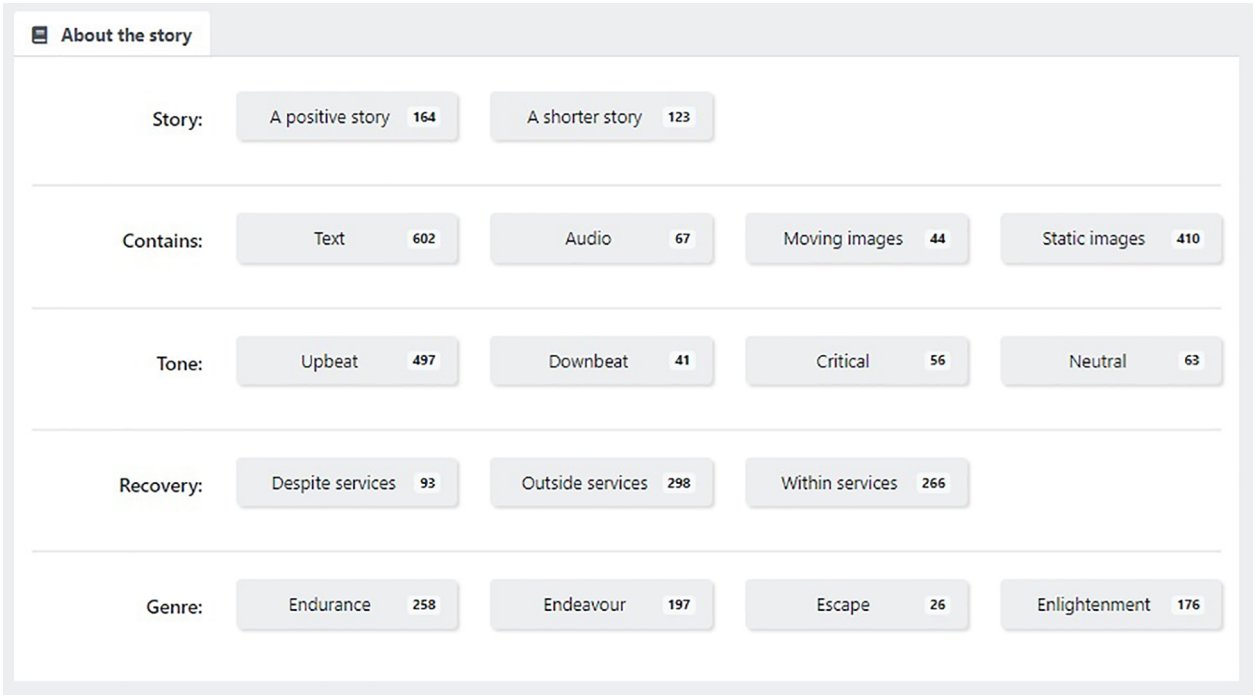


FIGURE 3
Characteristics of the narrative that can be used to filter for narratives.

appropriated can identify significant impacts and uncover unanticipated mechanisms underpinning success (38), providing developers with knowledge to guide iterative development work (39). Perceptions of DHIs can shape user experiences of those DHIs, and ultimately might influence their effectiveness (27).

2 Materials and methods

This paper reports an analysis of interviews collected during the process evaluation for the NEON Trial and NEON-O Trial. Ethical approval was granted in advance by Leicester Central Research Ethics Committee (REC reference: 19/EM/0326). Trial participants provided online consent when registering for the trials. All participants verbally re-confirmed their consent to take part in the interview. A description of resources allocated to trial recruitment work has been reported (31).

2.1 Sample

Shared trial inclusion criteria assessed at baseline were: resident in England; aged 18 or over at the time of registration; capable of accessing the Internet either on a personal computer, mobile, or device; able to understand written and spoken English; capable of providing informed consent; experience of mental health distress in the last 6 months, as assessed through three items from the Threshold Assessment Grid (40). NEON Trial participants had experience of psychosis in the last 5 years, and NEON-O Trial participants had experience of mental health problems other than psychosis in the last 5 years (both self-identified). Participants eligible for process evaluation interviews had been randomised to the intervention group; had provided all needed primary endpoint measures *or* were 32 or more days late in providing those measures (the lateness window was 31 days); and had provided consent to contact for interview through the informed consent form. All trial participants were recruited between 9th March 2020, and 26th March 2021. This coincided with a period in which the UK government imposed substantial restrictions on social interaction and physical mobility due to the COVID-19 pandemic, which were sustained throughout much of the trial follow-up period. Participants providing data late were included in case there was an association between low usage of the NEON Intervention and late provision of outcome measures, e.g., so as not to inadvertently exclude low users, who were seen as an important group to interview about their trial experiences.

An initial invitation to participate in the process evaluation was sent by email to an inception cohort of eligible trial participants, and respondents were interviewed. Work to invite participants for interview subsequently targeted a purposive sample with variation on three dimensions selected for their relevance to the user experience. These were (1) trial allocation; (2) intervention usage level; (3) mental health service usage (henceforth "service usage"). The latter was included because significant differences in baseline clinical and demographic characteristics were found between service users and non-service users in both trials, and

because service use history might influence how digital interventions are perceived and used (31).

For **Trial allocation**, participants recruited were either a part of the NEON Trial or the NEON-O Trial.

For **Intervention usage level**, a maximum variation sample was sought (41). Participants were categorised into disjoint groups, using logs of unique narratives requested between first access and the primary endpoint (52 weeks). Groups covered the entire range of usage:

Group 2.A: 0–1 narrative requests
Group 2.B: 2–9 narrative requests
Group 2.C: 10–30 narrative requests
Group 2.D: 31–79 narrative requests
Group 2.E: 80+ narrative requests

Since users are shown one narrative on gaining access intervention access, then membership of group 2.A indicates no intervention usage beyond the initial registration process. Low users were defined as members of 2.A or 2.B., and high users as members of 2.C, 2.D, or 2.E.

For **service usage**, a maximum variation sample was sought. Participants were categorised according to their use of statutory mental health services, as provided in England. Historical service use was collected at baseline on a web-based demographic form, and in-trial service use was collected at the primary endpoint through an abridged Client Service Receipt Inventory (CSRI). Forms are in Supplementary Appendix 1. Groups were:

Group 3.A (no mental health service treatment). Included participants who reported no mental health service usage of any kind before and during the trial.

Group 3.B (primary care mental health service treatment only). Included participants who had received treatment from primary care mental health services before the trial, and who did not initiate specialist mental health services before or during trial.

Group 3.C (specialist mental health treatment initiated during trial). Included participants who had no specialist mental health treatment before the trial but who initiated it during the trial.

Group 3.D (specialist mental health treatment before trial). Included participants who had received lifetime treatment from mental health specialist services before the trial.

2.2 Procedure

Semi-structured interviews were conducted with participants using Microsoft Teams or by telephone. Pilot interviews were conducted with members of the NEON study Lived Experience Advisory Panel ($n = 5$) to assess interview procedures, and refinements were made, e.g., giving participants the option to turn off their video if preferred. All interviews were audio-recorded using an encrypted recording device. Additionally, Microsoft Teams interviews were recorded using its recording function. Participants were offered £20 as compensation for their time and effort.

The topic guide requested participants to reflect on their year-long experience of using the NEON Intervention. A spreadsheet with summary information about their NEON Intervention usage up to the day before the interview was shown to each participant to aid reflection. During the interview, participants were asked to reflect on how relevant and inclusive the NEON Collection was. Participants were asked to describe how they had appropriated the NEON Intervention and to describe any impacts they felt had resulted from its use. Researcher prompts were used to support further questioning, and to aid participants explaining their responses in places. Regular meetings were held by research team members to revise the topic guide. A sample of the topic guide is presented in Supplementary Appendix 2.

Most interview recordings were transcribed by the research team, working from an initial Microsoft Teams auto-transcription where available. Some interviews with difficult intelligibility were transcribed professionally. All transcripts were pseudonymised and checked for accuracy by researchers before analysis. Participant names were replaced with fictional names, and named locations were redacted.

2.3 Analysis

An inductive thematic analysis was conducted using QSR International NVivo Version 12 Pro. The analyst team consisted of six members with various disciplinary backgrounds, some with personal experience of mental health problems. A preliminary coding framework was established through parallel coding of three transcripts. YA integrated codes produced by individual analysts into a coherent framework. The coding framework was then refined through the analysis of 32 further transcripts by YA, with qualitative analyst meetings held at critical points in this process as selected by YA, to discuss the analysis of transcripts, compare findings, and enrich developing codes and themes. Due to approaching saturation, the remaining 19 transcripts were examined for discrepant content only (e.g., content not already accounted for in existing codes), and the coding framework refined where this was identified. Nodes with the greatest relevance for understanding trial behaviours and supporting future intervention development work were selected and described. Nodes describing the impact of the NEON Intervention will be reported elsewhere. Selected descriptive characteristics of participants who were interviewed were tabulated. Ethnicity responses were grouped into two disjoint categories (White British, racialised ethnicity) to avoid risk of self-identification due to small numbers of participants in most categories other than White British, following UK Data Service guidance (42). Gender categories were Female; Male; Other.

3 Results

Fifty-four trial participants were interviewed. Most interviews were conducted using Microsoft Teams ($n = 49$), but five participants chose a telephone interview ($n = 5$). Interviews lasted between 45 and 116 min. Participant characteristics are presented in Table 1.

TABLE 1 Participant characteristics and group allocation.

Characteristic	*N*
Total Participants	*N* = 54
Gender	Female: 30
	Male: 21
	Other: 3
Age	
Mean (SD)	40.4 (12.2)
Median (min, max)	40 (21,72)
Ethnicity	White British: 46
	Minoritized ethnicity: 8
Trial Allocation (Dimension 1)	NEON Trial: 34
	NEON-O Trial: 20
Intervention Usage level (Dimension 2)	Low user: 23
	2.A (0–1 narratives): 5
	2.B (2–9 narratives): 18
	High User: 31
	2.C (10–30 narratives): 12
	2.D (31–79 narratives): 13
	2.E (80 + narratives): 6
Health service usage (Dimension 3)	3.A (No mental health service treatment): 0
	3.B (Primary care mental health service treatment only): 2
	3.C (Specialist mental health treatment initiated during trial): 0
	3.D (Specialist mental health treatment before trial): 52

Ninety nine trial participants met the criteria for category 3.A (31), and yet none were successfully recruited for interview, despite a disproportionately large effort to recruit people into this important purposive sampling category. We conclude that participants who had received no mental health service treatment were harder to engage than participants in other service use categories. Only two participants initiated specialist mental health treatment during the trial, and hence category 3.C had low potential for interview recruitment.

3.1 Objective 1: motivations for trial participation and NEON intervention usage

3.1.1 Extrinsic motivations

Receiving a payment voucher was a primary motivation for trial participation for some low users in the NEON Trial. Patrick stated:

> *"The reason why I used it was for the rewards you get, the voucher. I am being honest you know. Because I just need the voucher to buy things with, yeah. And also because of the issues I have got as well helps obviously." (Patrick, NEON, Low user)*

Some low users described having no further interaction with the NEON web-application once they had claimed their voucher, including never using the NEON Intervention.

The NEON Intervention was used by some only in preparation for their process evaluation interview, motivated by knowledge that they would be questioned about their NEON Intervention usage. When reflecting on the spreadsheet describing her usage, Rose expressed:

> *"Yeah, I didn't use it that much, that [increase in use] was because I knew this interview was coming up." (Rose, NEON, Low user)*

Usage after the primary endpoint was not accounted for in the definition of groups used in the purposive sampling strategy, but was reflected in the usage summary spreadsheet shown to participants, and hence has not influenced the usage category participants were placed into.

3.1.2 Intrinsic motivations

High users were typically intrinsically motivated, i.e., they wanted to use the NEON Intervention for their own sake, with the intention of reaching their personal goals. We identified three specific intrinsic motivations described by participants.

3.1.2.1 To explore the potential of an alternative or additional recovery approach

Some engaged with the NEON Intervention to explore its potential as an **alternative** to statutory National Health Service (NHS) Mental health services in England, which often have long waiting lists and capacity limitations. Some engaged to explore its potential as alternative to pharmaceutical treatment, especially for those with no desire to use medication long-term. For example:

> *"I had I felt, I felt like I tried everything available to me and I was sort of waiting on NHS services, which obviously take forever and I'm kind of willing to try anything at that stage, I think like most people, I don't want to have to take medication for the rest of my life and I'm very keen on non-pharmaceutical options and this seemed quite interesting." (Serena, NEON, High user)*

Some participants used the NEON Intervention as an alternative to mental health services that they had no access to, despite actively pursuing them:

> *"Well, basically I can't get any mental health support in my area because I am diagnosed with a personality disorder … because I have got a formal diagnosis, I am refused any help in my area because … they don't have anybody qualified to deal with my complexity of issues. I have tried alternative routes to try and get help. In [location redacted] I had a very good support system, and they were very helpful and luckily, I did some DBT therapy but when I moved to [location redacted] it is a completely different system, they left me with no support whatsoever. So that's what motivated me really to use NEON, to give me the help I can't get elsewhere right now." (Zendaya, NEON, High user)*

This alternative status meant that NEON Intervention usage could be influenced by changing patterns of health service

availability. For example, usage could be lower if alternative support became available, and greater as it became less available:

> *"The only thing with NEON is when I have got my care co-ordinator, I have got my therapist, I have got my employment support and I've got all my other things that I get from the EIS (Early Intervention Service), I used NEON less then because I've got so much support from them but I probably will use it more when I don't have any of their services in December and I'll use it again next year because I won't have any support then." (Hattie, NEON, High user)*

Some participants described a motivation to explore the NEON Intervention as an **addition** to a pre-existing set of mental health strategies:

> *"I wanted to see if this tool would add anything different to what I already do for my mental health, but I've already got a good routine going, so maybe this tool would've been more years ago, right at the start of it all." (Dorian, NEON-O, High user)*

3.1.2.2 To search for answers about mental health

The NEON Intervention was described as a "safe place" by some participants who stated that they felt it was a secure platform to search for answers regarding their mental health experiences. This sense of security enabled participants to safely meet their other intrinsic motivations such as interest in recovery narratives and learning different forms of recovery. Users also appreciated that NEON Intervention narrators were open about their struggles and the reality of mental health.

> *"Obviously I have a history of mental illness so it was always a case of anything I can do to help and yes it was a really interesting, especially reading stories, other people's recovery stories and yes in a way as well I wasn't expecting it to make that much of a difference but when it did oh there's other people like me and I was really interested to read and learn from that point of view and NEON was a safe place to do that as well." (Paul, NEON-O, High user)*

3.1.2.3 Due to working as a mental health professional

An unanticipated finding was that some trial participants were professionals who worked in the field of mental health (including healthcare workers and researchers), whose primary stated motivation to register for the NEON trials and access recovery narratives was to gain knowledge that aided their work around mental health conditions, rather than for the personal mental health benefits of narrative access. Erin described how her experiences with the NEON Intervention influenced her working practices as a peer support worker:

> *"… I did do a searching using psychosis and then found appropriate stories just to kind of support the work that I was doing, yes, I did use it for that and it did work in that respect quite well." (Erin, NEON, High user)*

[The NEON Intervention does not allow keyword searches and hence in this example we infer that Erin was inspired by NEON Intervention usage to search for published recovery narratives].

Florence described looking for insights arising due to narratives being under the control of narrators:

> *"So, I come under the job titles like service user researcher, so everyone in our team has their own personal experience and mental health problems, I listen to, and I talked to people who have mental health problems in my work all the time. And, you know, that's more kind of conversation. So, I thought this would be a really good opportunity to see all the other ways that people kind of…because there's a difference between, you know, having a conversation, and then someone writing something that sort of might be anonymous. You know, the people that are kind of, I suppose it's, maybe they've written them in a bit more of a safe way. So, people could maybe be more honest, was what I was anticipating or maybe hoping for some good insight or for them to be helpful." (Florence, NEON-O, High user)*

Whilst for these and other participants, motivation to participate was grounded in their profession, we have verified through inspection of transcripts that all participants who were motivated in this way were also legitimate trial participants, in that they consistently met inclusion criteria, including having personal experience of mental health problems.

3.2 Objective 2: appropriation of the NEON intervention

Some participants described how they consciously integrated the NEON Intervention into their everyday lives, developing their own approach to working with it which provided benefits for them. For some high users, NEON Intervention usage was an almost daily occurrence; this is reflected in usage logs demonstrating that some users accessed hundreds of unique narratives through the NEON Intervention. Descriptions of successful appropriation through thoughtful use were particularly likely to be provided by high users. The conscious nature of these processes is evident in emergent participant usage of the term "tool" to describe the NEON Intervention, usage of which demonstrates that some participants were actively thinking about the properties of the NEON Intervention and how these might fit into their lives. Descriptions of appropriation were present across a broad range of interviews with both high and low users, e.g.,

> *"I really believe in therapies other than medication, I mean I also do take medication, but I believe in talking therapies and all that sort of thing, so I'd heard about these recovery stories and I thought it maybe the tool for me and I was really*

> *poorly at the time, so I really wanted to do it for myself, so that's why I took part." (Trish, NEON, High user)*

Participants also described a broad range of outcomes of NEON Intervention use which were enabled by successful appropriation. These were mostly congruent with outcomes identified in the NEON Impact Model (20), a validation of which will be reported elsewhere. Below, we describe the three forms of appropriation which were most strongly evident in interviews.

3.2.1 Tool for distraction

The NEON Intervention was appropriated as a source of distraction by some, for example as a distraction from negative internal thoughts, through being immersed in the narration of others.

> *"Even though I kind of like I throw myself into the world of stories, and that does distract me, usually from me. I never use it as a kind of coach, but I think their words do kind of distract me even on a subconscious level which is brilliant. And so that source of distraction is important whatever you do." (Kieran, NEON-O, High user)*

> *"I feel like mental health services have a role like in my life have a massive role. But like there's stuff outside of services that also impacts on like how you are. So, I was like, reading people's stories or looking at artwork that they've made which was helpful for me. So, these stories became something that I could add to my daily arsenal of like, Oh, if it's a rubbish day, or like, if I need distraction …" (Tia, NEON, High user)*

Grace indicated that for her, the happier narratives were the ones accessed more for distraction:

> *"There's sometimes when you can get really, really to a point when you're at your lowest. When I'm there I just keep myself in that low mindset, which obviously, it could be quite counterproductive. It's just, it's trying to crack into that. And get yourself to do something. The stories helped distract me, but I need to say that it was more of the happy ones. I think I used it a lot for happy distractions." (Grace, NEON, High user)*

3.2.2 Tool for emotional boost

Some participants reported turning to the NEON Intervention when they needed to boost their mood.

> *"These stories were a kind of quick pick me up almost that I could use daily." (Tia, NEON, High user)*

> *"So, I kind of thought okay this is what obviously maybe give me a bit of inspiration to uplift me when I needed it. You know if people have recovered and have gone to extra learning and be able to do full time jobs and it was just kind of give me a little glimmer of hope that there's something beyond this point and when you're at your lowest point, you need something to bring you out of that, give you hope and you think you're not alone. I'd go on NEON when I needed that uplift." (Hattie, NEON, High user)*

3.2.3 Tool for sustaining the perception of a support network

Some participants described perceiving NEON Collection narratives as a support network that enabled an indirect sense of connection with others, and which made them feel a part of a supportive community, despite having no direct contact with NEON Intervention narrators:

> *"It's a bit like a bit of inspiration and a bit of anonymous support in a round-about way, you can't help but be inspired, I mean there's some stories on there you think how did they get out of it or how do you get that low down in recovery (…) it's a support network. As I say, it doesn't matter which story you're reading, even the ones that I couldn't relate to, they've all been inspiring as a support network." (Matty, NEON, High user)*

This was congruent with the outcome of "connectedness" described in the NEON impact model, and elucidated in other work (23), and was a particularly helpful form of impact for individuals that felt isolated from friends and family.

In common with other participants, and as well as using NEON as a tool for emotional boost, Matty described turning to the NEON Intervention due to his perception of it as a support network, and due to help that came from enhancing this perception:

> *"… it gets to the point that it book marked on my phone, the website, then you start thinking I'll just go back and read a story here, see how someone else dealt with it and it ends up a sort of staple and you start using it as a resource for recovery, so when I've been getting more and more down I've found myself instead of listening to a podcast go back and just reading through some NEON stories, there were various other resources I used as well but as my mood has gone down I've used NEON more as a resource to get myself back up again and as a little support network, even though you don't know the people who've written the stories, a little support network and you can relate to it. NEON has grown into this fabulous resource, I mean my own personal circumstances have changed this year regarding my mental health but it's nice to have a new resource, almost like a peer support, although you don't know who these people are it's like peer support, it's just been a fantastic tool to have and I feel I've just been lucky that I've stumbled across NEON at the same time as I needed them" (Matty, NEON, High user)*

3.3 Objective 3: perceptions of the NEON collection

Participants discussed their perceptions of the NEON Collection in substantial detail. We identified the following

findings as being of particular salience to DHIs integrating narrative collections.

3.3.1 Disengagement after early contact

The first few narrative(s) that participants were recommended or selected themselves influenced their engagement with the NEON Intervention for some. Some low users stopped using the NEON Intervention when the initial narratives they encountered were perceived as not relevant:

> *"I didn't read more than three, maybe four. To be honest they just weren't relevant to me, kind of a waste. I just stopped using it after that. There was no point, do you know what I mean?" (Yara, NEON-O, Low user)*

Not finding the right content could lead to a negative evaluation of the NEON Collection as a whole (even if that content was available, but the participant had not succeeded in locating it):

> *"It's kind of only really reflecting on it now, that I realised that that might have been why I didn't really use it much. Because I didn't find the content that I needed to begin with, which maybe made me feel like it wasn't helpful." (Alice, NEON, Low user)*

These experiences were not useful; some high users described persisting with the NEON Collection, even if they did not immediately find helpful narratives.

3.3.2 The desired narrative

Many participants felt they could articulate characteristics of recovery narratives that might most benefit them, and desired to find narratives with these characteristics in the NEON Collection. They frequently felt more engaged with the NEON Intervention if they found narratives that matched their ideal, and less engaged if they failed or found it difficult to find such a narrative. Hence being able to find a narrative with their desired characteristics influenced NEON Intervention engagement for those participants who felt they could articulate those characteristics. In Table 2, we synthesise the characteristics of desired narratives described by participants.

Some forms of desired narrative content presented in Table 2 were identified from transcript fragments where participants described content that they perceived as missing from the NEON Collection, but would ideally like to have found. For example, the NEON trials were open during a period that roughly coincided with the COVID-19 pandemic, and some participants desired to find narratives describing recovery from mental health experiences relating to events occurring in the pandemic, or due to the sustained period of isolation encountered during the pandemic (which were not present in the NEON Collection). Some participants desired narratives describing recovery experiences from a broader range of mental health diagnoses than were explicitly available—many NEON Collection narratives do not explicitly use diagnostic labels when describing mental health experiences, and hence participants looking for narratives discussing less-common diagnoses may not have found them. Most generally, we can conclude that there is no narrative that would be perceived as ideal for all participants, as e.g., advice-based narratives and experience-focused narratives were coded as mutually exclusive categories due to how they were described.

TABLE 2 Characteristics of narratives considered ideal by study participants.

Narratives with the potential for impact	Narratives with the potential to shift the participant's understanding of their mental health experiences
	Helpful narratives [nb. Participants frequently indicated that they wanted to find narratives with the potential to help them, without being specific about what this mean]
Narrative form	Narratives that avoid formulaic structures
	Positive narratives
	Hopeful narrative
	Reflective narratives
	Brave narratives
	Authentic narratives
Narrative content	Narratives about isolation
	Physical health narratives
	Religious narratives
	Music-based narratives
	Up to date narratives
	Narratives describing a broad range of mental health diagnoses
	Advice-based narratives
	Experience-focused narratives [that avoided giving advice]

3.3.3 Characteristics of narratives perceived as authentic

The perceived authenticity of narrator and narrative was reported by some as a central factor on which they decided whether to continue interacting with a narrative.

> *"But as long as it's authentic, the worst thing would be to force all the stories into saying that narrative of hope. It would then start to feel contrived and not real. So, like I say, although maybe a little bit disappointing at the time the video story didn't really take me anywhere, it still felt it gave me trust, that these were genuine stories, and not sort of fake stories told for a purpose, which I wouldn't enjoy at all." (Rob, NEON, Low user)*

For some users, an anonymous narrator aided perceptions of authenticity, since being anonymous in a narrative could be perceived as a human response to mental health stigma or perhaps a mechanism enabling honesty that may not otherwise have been possible:

> *"I really liked the stories where the author had no name, you didn't know who they are, they were just that bit more believable." (Angela, NEON, High user)*

We examined all discussions of authenticity in transcripts, and identified characteristics of narratives that were more likely to lead to perceptions of authenticity. These were:

- Acknowledgement of difficulty: the narrator acknowledges the difficult reality of mental health experiences
- Realism: The narrative appears to describe real world experiences
- Shared experience: The mental health experiences of the narrator are similar to those of the participant [enabling the participant to relate to those experiences]

The final point suggests that perceptions of authenticity (or not) can arise as an interplay between characteristics of the narrative and of the recipient, and hence that perceived authenticity is not simply a feature of a narrative alone.

3.3.4 Scale and diversity of the NEON collection

For most participants, the substantial scale and diversity of narratives in the NEON Collection supported people to remain interested and engaged in using the NEON Intervention.

> *"I found it very interesting, I found that the stories were very, some of them they were powerful and coming from a lot of insight of the lived experience and very so it was quite a big variety of some experiences, so it's just very, very every sort of diagnosis and so yes I found it very well done, kept me interested, and it was very well think of and very easy to follow and yes and powerful as well yes." (Katlyn, NEON, High user)*

Some participants who were high users greatly appreciated the large scale of the NEON Collection. For example, always being able to access new narratives kept a high user [178 narratives accessed] engaged despite experiencing ADHD:

> *"… I had a really late diagnosis of ADHD which explained a lot of things and now it's about really making the most of what I'm good at rather than kicking myself because I can't do Excel spreadsheets because it's not that I don't try hard enough, I just don't quite have the skills or the neuron connections to do it like other people. So yes, NEON's good, I love the fact that despite all efforts I haven't been sent a story twice, that's pretty cool and kept me intrigued to read and explore more." (Stefanie, NEON, High user)*

High users also reported that they appreciated there being sufficient narratives they could access in the NEON Intervention regularly.

The scale of the NEON Collection was found to be overwhelming for some due to the magnitude of narratives available, and this affected engagement:

> *"There were so many stories that I sometimes switched off." (Angela, NEON, High user)*

Erin found that NEON Collection use led her to access narratives outside of the NEON Intervention, and that the scale of what was available outside of the NEON Intervention could be overwhelming.

> *"I think there's other stories on there that might have a similar impact on me that I've not found yet, not discovered that collection of stories so it's kind of good that I would still have ongoing access so that I can find out a little bit more about things that I might have missed but also I suppose I can see why it's overwhelming because I've set up the thing where I carry on researching even more and look into other stuff so I suppose it kind of opens up even more avenues but those are ones that I've chosen to go down myself, rather than it all being on NEON." (Erin, NEON, High user)*

Finally, having a variety of narrative formats was perceived as unanimously beneficial, for example as a mechanism for accommodating opposing preferences by different participants.

> *"Just to know that there's a story here for you, it's kind of a thing where there is a good variety of stories. I think format is important for people. You know, I don't like listening to audiobooks, whereas people I know absolutely love it. And that's how they get their books. We've got images, videos, so, it's sort of a case of what you want." (Rob, NEON, Low user)*

4 Discussion

4.1 Principle findings

Our findings have demonstrated that the NEON Intervention was successfully appropriated by some participants, and suggest that appropriation is a factor in high use. Some engagement (or the lack of it) was explained by participants' perceptions of the NEON Collection that underpins the NEON Intervention. Engagement was typically enhanced when people found narratives that they perceived as authentic, or narratives with characteristics that matched those that they desired to find. For some participants, failing to find relevant narratives early on contributed to low engagement, and could contribute to a perception that the NEON Collection as a whole lacked relevance for them, even though this might not have been correct. The diversity of narrative formats contained in the NEON Collection was universally perceived as positive, and the large scale and diversity of content was perceived as positive by most. A minority of participants found the scale of the NEON Collection overwhelming, which contributed to disengagement for some. Engaging with narratives in the NEON Intervention could lead to participants searching for recovery narratives outside of the NEON Intervention, but in turn the scale of what is publicly available could feel overwhelming. Some participants found that accessing the NEON Collection provided a perception of having a distributed social support network despite there being no mechanism to directly engage with narrators or other users. Narrative collections such as the NEON Collection may be a mechanism for creating a beneficial perception of *common humanity*, ie. a recognition that struggle is a common human experience across all societies (43).

4.2 Relationship to prior work

Some trial participants perceived anonymous narratives as being authentic, due to a belief that a narrator would choose to be anonymous because of anticipated stigma about mental health problems. However, our own work suggests that anonymity can be enforced on narrators by curators of narrative collections who fear narrator harm through narrator identifiability (2). Since The NEON Collection integrates narratives from more than thirty existing collections with varying curatorial processes, then it is possible that this anonymity, and hence these perceptions of authenticity, were due to curator rather than narrator choice.

In parallel, Winstone et al. have examined how people who self-harm evaluate lived experience narratives. They found that perceiving a narrative as authentic can contribute to help-seeking by a recipient, and that narrative authenticity can be promoted by factors such as presenting realistic representations of recovery as a non-linear process (include setbacks or relapses), and honesty about self-harm as a coping mechanism. They recognise that presenting more "extreme" accounts of crisis might contribute to perceptions of authenticity whilst raising a danger of triggering self-harm in participants (44). This is in keeping with our own systematic review on the impact on recipients of recovery narratives, which identified that access to recovery narratives describing eating disorder experiences might encourage disordered eating in susceptible participants (22). Collectively, these insights point to narrative authenticity as both an opportunity and a responsibility challenge for narrative intervention developers, who may wish to minimise potential harms, while maximising both narrative impact and user safety.

Work to understand perceptions of authenticity may be informed by a wider body of related work on the credibility of health information (45) which demonstrates that sources perceived as having higher credibility are also perceived as more useful to individuals (45). The collation of narratives for narrative interventions may be supported by artefacts such as production guides, which can accumulate knowledge on what makes a narrative feel authentic. A video recovery narrative production guide has been co-produced and published by the KLIFAD study, which has examined the impact of alcohol misuse recovery narratives on people misusing alcohol. The KLIFAD production guide contains content enabling videographers to support participants in expressing narratives likely to be perceived as authentic, derived from an existing theory base (46). This guide was used to produce a collection of video recovery narratives for use in a feasibility trial (47). Since narrative diversity was mostly perceived as beneficial, then people curating collections may benefit from guidance on how to assess and plan for diversity. Kotera et al. have proposed diversity and inclusivity metrics for recovery narratives. These are broadly applicable beyond mental health (48), and are also well-aligned with responsible innovation practices (49) which are emerging as critically-important issues in healthcare research and development.

4.3 Strengths and limitations

Strengths include the use of a purposive sampling strategy that enabled access to a broad range of NEON Intervention experiences, from those that involved regular use, to those that involved no use. Whilst we sought for variation on service use history, we were limited by being unable to recruit participants who had never used any form of statutory mental health service, who were hard to engage. Future studies may consider new tactics to reach people in this category, such as enhanced payments, or prospective invitations to interview at the point of trial registration. Whilst we made a deliberate choice to only interview participants after the primary endpoint, so as not to influence quantitative outcome assessment, our interviews were reliant on participant memory. Our findings may have been influenced by the impact of the COVID-19 pandemic, including government-imposed restrictions on socialisation and mobility, which coincided with much of the trial recruitment and follow-up period.

4.4 Recommendations

Drawing on our findings and discussion, we propose the following recommendations to promote responsible innovation practices aimed at intervention developers and implementers.

(1) DHIs using digital media collections should make the scale, diversity, and inclusivity of collections clearly visible to users at first contact and beyond, and should provide access mechanisms that avoid users feeling overwhelmed.

If users can be rapidly deterred by perceiving a digital media collection as limited, then attempts to communicate the scale, diversity, and inclusivity of a media collection may enhance engagement. To support users at risk of feeling overwhelmed, more limited views of a collection should be provided. One tactic would be to provide smaller sub-collections of digital media items, for example a recommended set of ten narratives describing recovery after depression.

(2) Intervention developers should provide delivery mechanisms that afford appropriation

For example, since we found that some users accessed narratives on an almost daily basis, then a digital service providing a recommendation per day might afford appropriation by people who find benefits in that particular pattern of use.

(3) Curators of narrative collections should enable their contributors to produce authentic narratives.

This may involve creating resources such as production guides setting a clear expectation that authentic narratives are acceptable, and that narrators do not need to hide their experiences (whilst setting boundaries on experience representations that may cause harm to others, such as graphic descriptions of self-harm).

(4) Intervention developers should enable learning by legitimate users who are also healthcare professionals

This may provide an alternative route to impact from the intervention, for example by healthcare professionals using knowledge gained from recovery narratives to improve their own practice, or enabling healthcare professionals to recommend narratives from a collection to their clients.

(5) Web-based delivery of recovery narratives might be used to tackle loneliness across different health conditions due to their capacity to create a perception of having a social network

Implementation work might target who could benefit from feeling socially connected, which in turn may improve mental health (50–52).

5 Conclusions

Through our process evaluation of the NEON Trial and NEON-O Trial, we have developed knowledge on motivations for trial registration and NEON Intervention usage, forms of appropriation, and perceptions of our collection of mental health recovery narratives, and used to our findings to identify specific recommendations for developers of interventions making use of digital narratives. In considering how our intervention was appropriated, we have described three specific reasons for appropriation that were presented by our participants. Given that most use of digital healthcare technologies will need to take place as part of normal daily routines, and will have to sit well with those routines, then we propose that a research focus on understanding appropriation is important (though of course technologies can still be appropriated by people who have been separated from daily routines, such as people who have been hospitalised). In keeping with its study outside of healthcare, the study of appropriative processes for digital healthcare technologies may require socially oriented methods such as ethnomethodology, to provide direct observation of appropriative processes at work.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Leicester Central Research ethics committee. The studies were conducted in accordance with the local legislation and institutional requirements. Trial participants provided online consent when registering for the trials. All participants verbally re-confirmed their consent to take part in the interview.

Author contributions

YA: Formal Analysis, Methodology, Resources, Writing – original draft. SR: Conceptualization, Data curation, Formal Analysis, Methodology, Supervision, Writing – original draft, Writing – review & editing. JL: Formal Analysis, Writing – review & editing. FN: Formal Analysis, Writing – review & editing. CY: Formal Analysis, Writing – review & editing. DF: Formal Analysis, Writing – review & editing. EP: Writing – review & editing. DB: Writing – review & editing. YK: Formal Analysis, Writing – review & editing. MS: Conceptualization, Funding acquisition, Investigation, Methodology, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article.

This study is funded by the NIHR [Personal experience as a recovery resource in psychosis: Narrative Experiences Online (NEON) Programme (RP-PG-0615-20016)]. Mike Slade and Stefan Rennick-Egglestone acknowledge the support of the NIHR Nottingham Biomedical Research Centre (NIHR203310). Fiona Ng acknowledges support from an NIHR Advanced Fellowship (NIHR302218) and an Anne McLaren Fellowship. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2024.1297935/full#supplementary-material

References

1. Smith R. *Stop Paddling, Start Sailing: A Journey and Some Ideas*. Baltimore: PublishAmerica (2004).

2. McGranahan R, Rennick-Egglestone S, Ramsay A, Llewellyn-Beardsley J, Bradstreet S, Callard F, et al. The curation of mental health recovery narrative

collections: systematic review and qualitative synthesis. *JMIR Ment Health*. (2019) 6 (10):e14233. doi: 10.2196/14233

3. Llewellyn-Beardsley J. *"Not the Story You Want,I'm Sure": Mental Health Recovery and the Narratives of People from Marginalised Communities*. Nottingham: University of Nottingham (2023).

4. Llewellyn-Beardsley J, Rennick-Egglestone S, Callard F, Crawford P, Farkas M, Hui A, et al. Characteristics of mental health recovery narratives: systematic review and narrative synthesis. *PloS One*. (2019) 14(3):e0214678. doi: 10.1371/journal.pone.0214678

5. Slade M, Longden E. Empirical evidence about recovery and mental health. *BMC Psychiatry*. (2015) 15(1):1–14. doi: 10.1186/s12888-015-0678-4

6. Barton SS. Using narrative inquiry to elicit diabetes self-care experience in an aboriginal population. *Can J Nurs Res Arch*. (2008) 40(3):16–37.

7. Chou W-YS, Hunt Y, Folkers A, Augustson E. Cancer survivorship in the age of YouTube and social media: a narrative analysis. *J Med Internet Res*. (2011) 13(1):e7. doi: 10.2196/jmir.1569

8. Yeo C, Rennick-Egglestone S, Armstrong V, Borg M, Franklin D, Klevan T, et al. Uses and misuses of recorded mental health lived experience narratives in healthcare and community settings: systematic review. *Schizophr Bull*. (2021) 48(1): sbab097. doi: 10.1093/schbul/sbab097

9. Williams A, Fossey E, Farhall J, Foley F, Thomas N. Recovery after psychosis: qualitative study of service user experiences of lived experience videos on a recovery-oriented website. *JMIR Ment Health*. (2018) 5(2):e9934. doi: 10.2196/mental.9934

10. Bartolucci M, Batini F. The effect of a narrative intervention program for people living with dementia. *Psychol Neurosci*. (2019) 12(2):307. doi: 10.1037/pne0000141

11. Manuvinakurike R, Velicer WF, Bickmore TW. Automated indexing of internet stories for health behavior change: weight loss attitude pilot study. *J Med Internet Res*. (2014) 16(12):e3702. doi: 10.2196/jmir.3702

12. Li PW, Yu DS. A modeling-based narrative intervention to promote timely care-seeking in patients with acute myocardial infarction: a pilot randomized controlled trial and feasibility analysis. *Eur J Cardiovasc Nurs*. (2019) 18(3):215–23. doi: 10.1177/1474515118810126

13. Cordle H, Carson J, Richards P, Fradgley J. *Psychosis: stories of recovery and hope*. Salisbury:: Quay Books (2010).

14. Time to Change. Time to Change Website. Available online at: https://www.time-to-change.org.uk/ (accessed August 27, 2018).

15. Here to Help Canada. Here to Help: Personal Stories. (2022). Available online at: https://www.heretohelp.bc.ca/personal-stories (accessed August 27, 2018).

16. South London and Maudsley NHS Foundation Trust. Patient Stories: Stories of Recovery (2017). Available online at: https://www.slam.nhs.uk/patients-and-carers/patient-information/patient-stories (accessed June 12, 2019).

17. Yeo C, Hare-Duke L, Rennick-Egglestone S, Slade M. The VOICES typology of curatorial decisions in narrative collections of the lived experiences of mental health service use, recovery or madness: a qualitative study of interviews. *JMIR Ment Health*. (2020) 7(9). doi: 10.2196/16290

18. Rennick-Egglestone S, Elliott R, Newby C, Robinson C, Slade M. Impact of receiving recorded mental health recovery narratives on quality of life in people experiencing non-psychosis mental health problems (NEON-O Trial): updated randomised controlled trial protocol. *Trials*. (2022) 23:90. doi: 10.1186/s13063-022-06027-z

19. Llewellyn-Beardsley J, Barbic S, Rennick-Egglestone S, Ng F, Roe J, Hui A, et al. INCRESE: development of an inventory to characterize recorded mental health recovery narratives. *J Recov Ment Health*. (2020) 3(2):25–44.

20. Slade M, Rennick-Egglestone S, Llewellyn-Beardsley J, Yeo C, Roe J, Bailey S, et al. Recorded mental health recovery narratives as a resource for people affected by mental health problems: development of the narrative experiences online (NEON) intervention. *JMIR Form Res*. (2021) 5(5):e24417. doi: 10.2196/24417

21. Llewellyn-Beardsley J, Rennick-Egglestone S, Bradstreet S, Davidson L, Franklin D, Hui A, et al. Not the story you want? Assessing the fit of a conceptual framework characterising mental health recovery narratives. *Soc Psychiatry Psychiatr Epidemiol*. (2020) 55(3):295–308. doi: 10.1007/s00127-019-01791-x

22. Rennick-Egglestone S, Morgan K, Llewellyn-Beardsley J, Ramsay A, McGranahan R, Gillard S, et al. Mental health recovery narratives and their impact on recipients: systematic review and narrative synthesis. *Can J Psychiatry*. (2019) 64(10). doi: 10.1177/0706743719846108

23. Rennick-Egglestone S, Ramsay A, McGranahan R, Llewellyn-Beardsley J, Hui A, Pollock K, et al. The impact of mental health recovery narratives on recipients experiencing mental health problems: qualitative analysis and change model. *PloS One*. (2019) 14(12):e0226201. doi: 10.1371/journal.pone.0226201

24. Ng F, Charles A, Pollock K, Rennick-Egglestone S, Cuijpers P, Gillard S, et al. The mechanisms and processes of connection: developing a causal chain model capturing impacts of receiving recorded mental health recovery narratives. *BMC Psychiatry*. (2019) 19(1):1–15. doi: 10.1186/s12888-018-1996-0

25. Ng F, Newby C, Robinson C, Llewellyn-Beardsley J, Yeo C, Roe J, et al. How do recorded mental health recovery narratives create connection and improve hopefulness? *J Ment Health*. (2022) 31(2):273–80. doi: 10.1080/09638237.2021.2022627

26. Charles A, Hare-Duke L, Nudds H, Franklin D, Llewellyn-Beardsley J, Rennick-Egglestone S, et al. Typology of content warnings and trigger warnings: systematic review. *PloS One*. (2022) 17(5):e0266722. doi: 10.1371/journal.pone.0266722

27. Rennick-Egglestone S, Knowles S, Toms G, Bee P, Lovell K, Bower P. *Health technologies' in the wild': experiences of engagement with computerised CBT. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. (2016). pp. 2124–35

28. Resnick P, Varian HR. Recommender systems. *Commun ACM*. (1997) 40 (3):56–8. doi: 10.1145/245108.245121

29. Rennick-Egglestone S, Elliott R, Smuk M, Robinson C, Bailey S, Smith R, et al. Impact of receiving recorded mental health recovery narratives on quality of life in people experiencing psychosis, people experiencing other mental health problems and for informal carers: narrative experiences online (NEON) study protocol for three randomised controlled trials. *Trials*. (2020) 21(1):1–34. doi: 10.1186/s13063-020-04428-6

30. Slade M, Rennick-Egglestone S, Elliott RA, Newby C, Robinson C, Gavan SP, et al. Effectiveness and cost-effectiveness of online recorded recovery narratives in improving quality of life for people with non-psychotic mental health problems: a pragmatic randomized controlled trial. *World Psychiatry*. (2024) 23(1):101–12. doi: 10.1002/wps.21176

31. Rennick-Egglestone S, Newby C, Robinson C, Yeo C, Ng F, Elliott R, et al. Differences between online trial participants who have used statutory mental health services and those who have not: analysis of baseline data from 2 pragmatic trials of a digital health intervention. *J Med Internet Res*. (2023) 25. doi: 10.2196/44687

32. Knowles SE, Toms G, Sanders C, Bee P, Lovell K, Rennick-Egglestone S, et al. Qualitative meta-synthesis of user experience of computerised therapy for depression and anxiety. *PLoS One*. (2014) 9(1):e84323. doi: 10.1371/journal.pone.0084323

33. Balaam M, Rennick Egglestone S, Fitzpatrick G, Rodden T, Hughes A-M, Wilkinson A, et al. *Motivating mobility: designing for lived motivation in stroke rehabilitation. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 2011*. (New York: ACM; 2011).

34. Proudfoot JG, Parker GB, Pavlovic DH, Manicavasagar V, Adler E, Whitton AE. Community attitudes to the appropriation of mobile phones for monitoring and managing depression, anxiety, and stress. *J Med Internet Res*. (2010) 12(5):e1475. doi: 10.2196/jmir.1475

35. Hutchinson H, Mackay W, Westerlund B, Bederson BB, Druin A, Plaisant C, et al. *Technology probes: inspiring design for and with families. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 2003*. (New York: ACM; 2003).

36. Marcelo Álvarez S. The teacher as a decision maker in the appropriation of digital technologies for didactic intervention: the benefits of technological sovereignty based on the free paradigm. *Revista Innovaciones Educativas*. (2021) 23(SPE1):110–9.

37. Tolmie P, Pycock J, Diggins T, MacLean A, Karsenty A. *Unremarkable computing. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*; (2002).

38. Ryan K, Linehan C, Dockray S. *Appropriation of digital tracking tools in an online weight loss community: individual and shared experiences. Designing Interactive Systems Conference 2021*; (2021).

39. Rennick-Egglestone S, Brundell P, Koleva B, Benford S, Roussou M, Chaffardon C. Families and mobile devices in museums: designing for integrated experiences. *J Comput Cult Herit (JOCCH)*. (2016) 9(2):1–13. doi: 10.1145/2891416

40. Slade M, Powell R, Rosen A, Strathdee G. Threshold assessment grid (TAG): the development of a valid and brief scale to assess the severity of mental illness. *Soc Psychiatry Psychiatr Epidemiol*. (2000) 35(2):78–85. doi: 10.1007/s001270050011

41. Teddlie C, Yu F. Mixed methods sampling: a typology with examples. *J Mix Methods Res*. (2007) 1(1):77–100. doi: 10.1177/1558689806292430

42. Hundepool A, Domingo-Ferrer J, Franconi L, Giessing S, Lenz R, Longhurst J, et al. *Handbook on Statistical Disclosure Control*. London: ESSnet on Statistical Disclosure Control (2010).

43. Neff K. Self-compassion: an alternative conceptualization of a healthy attitude toward oneself. *Self Identity*. (2003) 2(2):85–101. doi: 10.1080/15298860309032

44. Winstone L, Mars B, Ferrar J, Moran P, Penton-Voak I, Grace L, et al. Investigating how people who self-harm evaluate web-based lived experience stories: focus group study. *JMIR Ment Health*. (2023) 10:e43840. doi: 10.2196/43840

45. Sbaffi L, Rowley J. Trust and credibility in web-based health information: a review and agenda for future research. *J Med Internet Res*. (2017) 19(6):e218. doi: 10.2196/jmir.7579

46. Rennick-Egglestone S, Subhani M, Knight H, Jones K, Hutton C, Jackson T, et al. Transient elastography and video recovery narrative access to support recovery from alcohol misuse: development of a novel intervention for use in community alcohol treatment services. *JMIR Form Res*. (2023) 7. doi: 10.2196/47109

47. Subhani M, Enki DG, Knight H, Jones KA, Sprange K, Rennick-Egglestone S, et al. Does knowledge of liver fibrosis affect high-risk drinking behaviour

(KLIFAD): an open-label pragmatic feasibility randomised controlled trial. *eClinicalMedicine*. (2023) 61. doi: 10.1016/j.eclinm.2023.102069

48. Kotera Y, Rennick-Egglestone S, Ng F, Llewellyn-Beardsley J, Ali Y, Newby C, et al. Assessing diversity and inclusivity is the next frontier in mental health recovery narrative research and practice. *JMIR Ment Health*. (2023) 10(1):e44601. doi: 10.2196/44601

49. Owen R, Pansera M. *Responsible Innovation and Responsible Research and Innovation*. Cheltenham: Edward Elgar Publishing (2019).

50. Stickley T. *"Ah, Look at All the Lonely People…": Loneliness and Solitude in Public Mental Health*. London, England: SAGE Publications Sage UK (2021). p. 186–7.

51. Ridout B, Campbell A. The use of social networking sites in mental health interventions for young people: systematic review. *J Med Internet Res*. (2018) 20 (12):e12244. doi: 10.2196/12244

52. Anderson K, Laxhman N, Priebe S. Can mental health interventions change social networks? A systematic review. *BMC Psychiatry*. (2015) 15(1):1–8. doi: 10.1186/s12888-015-0684-6

frontiers | Frontiers in Digital Health

TYPE Original Research
PUBLISHED 12 April 2024
DOI 10.3389/fdgth.2024.1334840

OPEN ACCESS

EDITED BY
Kim Cornelia Marie-Louise Bul,
Coventry University, United Kingdom

REVIEWED BY
Candice A. Alfano,
University of Houston, United States
Jakub Swacha,
University of Szczecin, Poland

*CORRESPONDENCE
Christine Seaver
christine.seaver@ucf.edu

RECEIVED 07 November 2023
ACCEPTED 22 March 2024
PUBLISHED 12 April 2024

CITATION
Seaver C, Bowers C, Beidel D, Holt L and Ramakrishnan S (2024) A game-based learning approach to sleep hygiene education: a pilot investigation.
Front. Digit. Health 6:1334840.
doi: 10.3389/fdgth.2024.1334840

A game-based learning approach to sleep hygiene education: a pilot investigation

Christine Seaver[1]*, Clint Bowers[1], Deborah Beidel[1], Lisa Holt[2] and Sridhar Ramakrishnan[2]

[1]UCF RESTORES, University of Central Florida, Orlando, FL, United States, [2]BlueHalo, Rockville, MD, United States

Introduction: Sleep hygiene education (SHE) consists of environmental and behavioral practices primarily intended to reduce sleep problems. Currently considered ineffective as a stand-alone treatment, the manner in which the education is typically delivered may be ineffective for the acquisition of new knowledge. The purpose of this study was to determine if a more engaging teaching medium may improve the efficacy of sleep hygiene education. This study examined the use of game-based learning to teach SHE to individuals with sleep problems.
Methods: 35 participants played the SHE games for 30 days. Differences in pre- and post-state anxiety and sleep quality measures were examined.
Results: Participants had significant improvements in sleep quality and state anxiety after using the app for 30 days, although scores for the majority of patients remained elevated.
Discussion: This pilot investigation provides initial evidence for the efficacy of a game-based approach to SHE.

KEYWORDS

sleep hygiene, game-based learning, serious games, games for health, sleep hygiene education (SHE)

Introduction

Quality sleep is imperative for health and overall quality of life. In fact, dimensions of sleep quality have been linked to numerous poor health outcomes such as coronary heart disease and hyptertension in numerous cross-sectional (1, 3) and cohort (4) studies. These studies generally demonstrate a moderate effect of sleep quality on health outcomes. Cross-sectional (5) and experimental (6) studies have also demonstrated relationships between sleep problems and increased levels of anxiety. More specifically, one study (7) found that individuals reported significanly lower sleep quality on days when they experienced higher levels of stress. Similarly, another study (8) found that stressful life events had a direct effect on sleep quality. These findings are not surprising given that poor sleep quality leads to a physiological neuroendocrine stress response in which sympathetic tone and cortisol levels are increased (9). Finally, research has shown that reduced sleep quality can bave impacts beyond physical and mental health. Indeed, a cross-sectional study indicated that poor sleep leads to occupational impairments, such as reduced productivity and efficiency (10).

Two primary indicators of good sleep quality are (1) a sleep latency of ≤15 min and (2) sleeping through the night (or only having few awakenings of ≤5 min) (11). However, a recent National Health Interview (12) found that over a 30 day period, 14.5% of adults in

the United States had difficulty falling asleep and 17.8% of adults staying asleep. These problems were more prevalent depending on factors such as race, socioeconomic status, urbanization level, age, and biological sex (12). In particular, the report found that individuals with lower levels of education and income had more trouble falling asleep. Additionally, non-hispanic white adults had more trouble staying asleep compared to non-hispanic black, Hispanic, and non-hispanic Asian adults. Furthemore, those in more rural areas had more trouble falling and staying asleep. Given this host of negative outcomes, a significant percentage of the population could benefit from evidence-based interventions to improve sleep.

Primary care providers (PCP) are often the first point of contact for individuals presenting with difficulties falling or staying asleep (13). A retrospective chart review of the patients who discussed sleep problems with their PCP found that 51.5% of the sample were offered a sleep medication. Intensive Cognitive Behavior Therapy was recommended to 5% of the patients, whereas sleep hygiene education (SHE) was provided to 31.5% of the patients (14). Sleep hygiene consists of recommendations that optimize sleep (15, 16), and a review (13) included caffeine intake, tobacco use, alcohol intake, napping, wake/sleep time, and exercise as the most covered topics. These recommendations are typically delivered verbally or printed on a handout (17). One survey (18) revealed that although 97% of general practitioners utilized SHE as a management tool for their patients, they often perceived SHE alone to be ineffective. Prior to concluding that SHE alone is ineffective, examining whether an alternative delivery method that would optimize acquisition and implementation of appropriate sleep behaviors merits consideration.

Delivering sleep hygiene education via gaming

SHE requires the acquisition and practice of certain behaviors, but typically is delivered in a single meeting with a practitioner. Such formats may not encourage engagement or provide opportunities to review and practice the behaviors. An alternative format is game-based learning, where learners interact with challenging interactive activities with a clear set of goals, constraints, and rules (19). Game-based learning formats offer more engaging learning mediums, and opportunities to practice newly acquired skills within a realistic context. Additionally, gaming environments afford the learner more control and feedback over traditional digital training, which, in turn, can increase motivation (19, 21). Finally, the "play" factor of games may increase motivation to learn (22–24), enhance educational achievement in comparison to a traditional learning environment (25, 26). As such, play can engage individuals in the process of learning, and thereby lead to better outcomes. Given its numerous advantages, it is not surprising that game-based learning has been on the rise over the last several decades. Games are used for learning in several industries, including mental health (27), healthcare (28), and training (29). In particular, serious game interventions have been shown to improve sleep quality in children. A study by Wilson, Miller, Bonuck, Lumeng, and Chervin (2014) (30) the "Sweet Dreamzzz Early Childhood Sleep Education Program™" "Dreamzzz Early Childhood Sleep Education Program™". One component of this program utilized game-based learning to promote sleep hygiene. Although the study did not parse out the efficacy of each individual component, the intervention led to improvements in sleep duration. Another study by Almondes and Leonardo (2019) (31) evaluated the effects of a serious game called "Perfect Bedroom: learn to sleep well" in children. The study found small but positive changes in the sleep routines of those in the experimental group. In perhaps the only study that addressed clinical insomnia, adults who engaged in cognitive training games (e.g., Sudoku) over 15-days had improved cognitive performance, mood, and changes in sleep quality (32). These data suggest that games are efficacious in promoting knowledge and skill acquisition, providing support for the testing of this format to teach SHE.

In summary, sleep problems are common in the general population and most people initially seek help from their PCP. Pharmacological agents are most commonly offered but the most common non-pharmacological intervention is SHE (14), which is considered only minimally effective, with effect sizes no larger than those found for psychological placebo (33). Although superior psychological treatments for sleep problems (e.g., CBT-I) exist, it is probable that many individuals never get referred, or are unable to pay, for this type of specialized treatment (14). Thus, many patients may be limited to receiving SHE as typically delivered in one office encounter with no follow-up. This manner of delivery is inconsistent with how acquisition of new knowledge and skills are acquired and maintained. One question is therefore whether delivery of SHE in a different format, such as game-based learning, may improve learning and thereby, its efficacy at relieving sleep difficulties. As such, this pilot study examines the use of a game-based sleep hygiene application in improving sleep quality.

Hypotheses

1. Sleep problems as assessed by scores on the Pittsburgh Sleep Quality Index (PSQI) will decrease (indicating better quality sleep) from pre-assessment to post-assessment.
2. Self-reported state anxiety as assessed by the State-Trait Anxiety-State (STAI-S) scores will decrease (indicating lower levels of state anxiety from pre-assessment to post-assessment).

Methods

Inclusion and exclusion criteria

Participants were at least 18 years old and reported trouble sleeping as assessed by scores ≥5 on the Pittsburgh Sleep Quality Index (PSQI) (34) at pre-assessment. Scores of ≥5 on the PSQI have been previously validated as indicative of poor sleep and is a commonly used cutoff score (34–36). Participants were included regardless of whether they were taking sleep medications.

Recruitment

This study was approved by the University of Central Florida's Institutional Review Board (IRB). Flyers were posted on community boards throughout Orange County, Florida. A link to the digital version of this flyer was included in a weekly email sent to the University of Central Florida community. Interested individuals contacted the researcher to receive an informed consent document. The signed consent was returned prior to beginning the study.

Demographics

Of the 35 participants, 5 declined to answer the demographic questionnaire. Demographics of the remaining 30 participants can be found in Table 1.

TABLE 1 Participant demographics.

Age	35.31 (SD = 16.25)
Female	63.4%
Male	36.6%
White/Caucasian	60%
Black	25%
Asian	5%
Pacific Islander	0%
Other	10%

Materials and procedure

Intervention description

Restful Journey is a mobile application designed to teach SHE and apply that knowledge in several play scenarios. Developed by the University of Central Florida (UCF) and Intelligent Automation, a Blue Halo Company, Restful Journey contains a series of mini-games that teach and reinforce good sleep hygiene by rewarding users for achieving sleep goals. The sleep hygiene concepts covered throughout the game were based on the American Academy of Sleep Medicine's (AASM) guidance (37). See Table 2 for a list of sleep hygiene behaviors taught in Restful Journey.

Restful Journey encompasses a sailing exploration theme. Players begin by learning sleep hygiene in preparation for their "journey". Over time as they progress throughout the game (i.e., achieve sleep hygiene goals) they are able to explore different islands and obtaining jewels from each.

"Healthy Habits" is the first mini-game. This is a quiz-based game comprised of three levels. Players learn SHE through playing the game and receiving feedback regarding their responses. Players must accurately answer 90% of questions to proceed to the next level. Each level increases in difficulty by decreasing the time allotted to answer the questions. In keeping with the theme of Restful Journey, the overarching goal of this game is to teach the "sailor" (i.e., the user) on how to stay well-rested in preparation of their journey.

TABLE 2 Sleep hygiene concepts taught in restful journey.

Napping for 30 min or less (or not at all).
Getting at least 30 min of physical activity.
Keeping the bedroom temperature between 63 and 66 degrees.
Waking up at the same time as the previous day.
Going to bed at the same time as the previous day.
Engaging in an appropriate activity before bedtime (e.g., journaling, yoga, reading, sex).
Having last caffeine intake 4–6 h before bed (or not at all).
Having last alcohol intake at least 4–6 h before bed (or not at all).
Having last meal or snack at least 3 h before bed.
Falling asleep within 20 min.
If not falling asleep within 20 min, doing an appropriate activity.
If getting up due to not falling asleep- waiting until tired to go back to bed.
Keeping the lights dim or off prior to going to bed.
Only engaging in sleep and/or sex in bed.
Getting 7–9 h of sleep.

"Sleep Scrutiny" is the second mini-game in which players progress to "Sleep Captain". Players must inspect their crewmate's quarters and identify objects that are conducive or detrimental to sleep. Again, there are three levels and rooms must be inspected with 100% accuracy to move forward. The time given to identify objects and answer questions decreases with each level. In keeping with the story of Restful Journey, players must help their crewmates stay well-rested to ensure that the crew is ready to set sail.

Participants were asked to dedicate 10–15 min per day to playing these first two games (i.e., "Healthy Habits" and "Sleep Scrutiny"), with the goal of completing them with a 90% correct score by the end of the first week of the study.

The final game, "Vigilant Voyage" translates game-based learning into real-life. Each day, players open the game and indicate which sleep hygiene behaviors they used the previous day. Coins are awarded for each behavior. Once the check-in is completed, the game generates a random sleep goal. If the player achieved that goal, they received bonus coins. Once enough coins are earned, the player can explore various islands and obtain jewels. Players must cash-in their coins to explore an island, requiring them to continue earning coins each day by meeting sleep hygiene goals to continue exploring other islands. Players "win" the game by collecting all of the jewels.

The game is structured to encourage all participants to meet the same sleep hygiene goals. Additionally, the game-based learning aspect provides motivation to work on any unmet goals.

See Figures 1–4 for visual depictions of Restful Journey.

Assessment measures

Assessment measures were completed at pre and posttreatment. The assessment included:

- The Pittsburgh Sleep Quality Index (PSQI) (34) is a self-report measure of sleep quality. The PSQI distinguishes "good" and "poor" sleepers. This assessment measures seven dimensions of

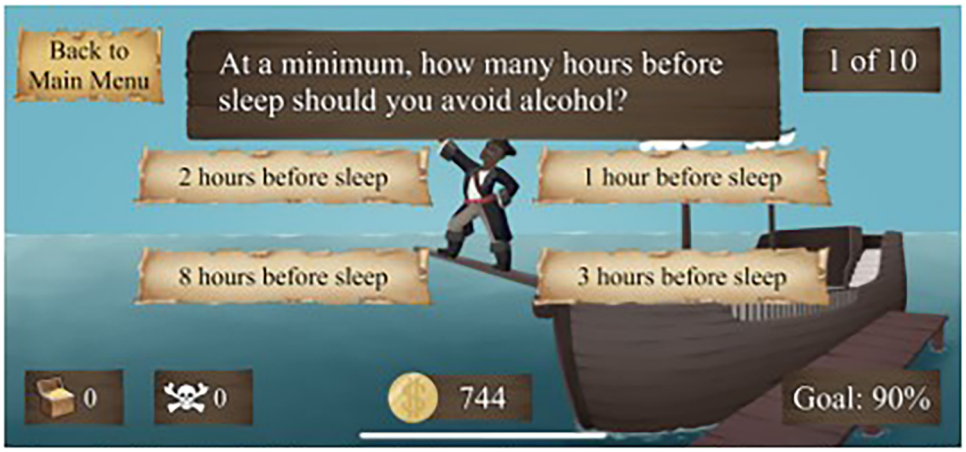


FIGURE 1
Healthy Habits. A Quiz-based game to teach players sleep hygiene principles.

FIGURE 2
Sleep Scrutiny. Players prepare virtual bedroom for sleep.

FIGURE 3
Vigilant Voyage. Players progress along the map as they earn points for meeting sleep goals.

FIGURE 4
Players log in daily to claim which sleep goals they met the previous day.

sleep quality: (1) subjective sleep quality, (2) sleep latency, (3) sleep duration, (4) sleep efficiency, (5) sleep disturbance, (6) use of sleeping medication, and (7) daytime dysfunction. These components can be looked at individually or added together to achieve a "global score of sleep quality." Global PSQI score above a "5" indicates "poor sleep quality." It should be noted that a score above a "5" does not necessarily indicate a particular clinical sleep disorder. This assessment was chosen to measure sleep quality due to its robust psychometric properties (38–40). The PSQI is the most frequently used generic measure of sleep quality among clinicians and researchers (40). As such, many studies have used the PSQI to measure improvements in sleep quality following an intervention (41–44).

- The State-Trait Anxiety Inventory (STAI-S) (45) is a widely used measure of state anxiety (i.e., current feelings of anxiety). Participants indicate the extent of their agreement with several items on a 4-point scale. The scale is internally consistent ($\alpha = .90$). Higher means of the sum of scores indicate higher degrees of anxiety. THE STAI has been used to evaluate the treatment of medical and anxiety disorders in thousands of studies (46) and is considered to be an adequate measure of anxiety in research and clinical settings.

After completing the screening and pre-assessment, participants received instructions to download Restful Journey on their smartphone. It was downloadable on Android or iPhone devices.

Participants were emailed a daily check-in questionnaire each day, which tracked the sleep hygiene goals achieved the previous night, which game(s) they played, and the time spent playing the game. The purpose of the daily check-in was twofold; (1) assess which sleep hygiene behaviors changed over the course of the study and (2) to ensure that participants were interacting with the game daily throughout the study duration. All participants completed at least 90% of the daily check-ins. Upon completion of the post-assessment (which was identical to the pre-assessment) participants received a digital $25 Amazon gift card.

Power analysis

An *a priori* power analysis was conducted using G*Power version 3.1 (47) for sample size estimation. The primary outcome measure was the PSQI (36). According to the power analysis, which was conducted for a paired samples *t*-test using an effect size of .5, and a significance criterion of $\alpha = .05$, the sample size needed to achieve a power of .85 is 31. An effect size of .5 was chosen as it is generally considered to be a moderate effect size (48) and is therefore sufficient to demonstrate practical significance of the results. Thus, the obtained sample size of $n = 35$ is more than adequate to test the study hypotheses.

Results

App usage

Based on the Daily Check-In data, participants self-reported spending an average of 9.16 min per day (SD = 7.32) playing Restful Journey. The average number of days played was 28.4. All participants reported that they completed all mini-games and the game in its entirety.

Sleep quality and state anxiety

All data analyses were conducted using paired sample *t*-tests. Due to a substantial number of tests ($n = 9$), we applied a false discovery rate correction to calculate adjusted *p* value thresholds for each test (see Table 3). The pre and post-scores for each of the dependent variables are displayed in Table 3. The results indicated that Restful Journey resulted in statistically significant improvements on several aspects of sleep quality assessed by the PSQI, including: global sleep quality, subjective sleep quality,

TABLE 3 Assessment scores and Benjamin and Hochberg false discovery rate correction.

Test no. (ranked in order by size of *p* value)	Variable	Mean pre-test score	Mean post-test score	Effect size (*d*)	Estimated *p* value[a]	Corrected threshold[b]	Outcome
1	PSQI- Global score	10.54 (*SD* = 3.25)	7.03 (*SD* = 3.34)	1.06	<.001	.005	Significant
2	PSQI- Subjective sleep quality	1.69 (*SD* = .63)	1.03 (*SD* = .66)	.91	<.001	.011	Significant
3	PSQI- Sleep latency	2.26 (*SD* = .82)	1.66 (SD = .87)	.71	<.001	.016	Significant
4	PSQI- Sleep duration	1.34 (SD = .76)	.77 (SD = .73)	.94	<.001	.022	Significant
5	PSQI- Daytime dysfunction	1.51 (SD = .78)	.91 (SD = .70)	.74	<.001	.027	Significant
6	PSQI- Sleep efficiency	1.49 (SD = 1.17)	.94 (SD = 1.03)	.55	.002	.033	Significant
7	STAI- State anxiety	38.46 (SD = 11.66)	35.14 (SD = 10.83)	.39	.026	.038	Significant
8	PSQI- Sleep disturbance	1.46 (SD = .70)	1.20 (SD = .58)	.39	.027	.044	Significant
9	PSQI- Medication usage	.71 (SD = 1.10)	.68 (SD = 1.18)	.36	.831	.05	Not significant

[a]Estimated *p* value refers to *p* value produced by each of the individual tests performed.
[b]Corrected threshold refers to the multiple comparison corrected α level, the metric to which *p* values are compared.

sleep latency, slepe duration, daytime dysfunction, sleep efficiency, and sleep disturbance. The average score on the PSQI at post-assessment was 7.03. After playing Restful Journey for 30 days, 22.8% of participants had PSQI global scores of <5 at the post-assessment, meaning that these individuals no longer met the criteria for being "poor sleepers". The results also indicated that Restful Journey resulted in statistically significant improvement of state anxiety as measured by the STAI.

Discussion

The overarching aim of this paper was to examine the utility of game-based learning to deliver SHE and examine its impact on numerous subcomponents of sleep quality, as well as state anxiety. This pilot investigation provides modest initial evidence for the use of *Restful Journey* to improve multiple facets of sleep quality as defined by the PSQI. In particular, these improvements were reported on the subcomponents of subjective sleep quality, sleep latency, and sleep duration. Moreover, participants had a significant improvement in sleep efficiency, another subcomponent of the PSQI. There was also a significant improvement in sleep disturbance. Finally, there was a significant decrease in daytime dysfunction.

Although participants experienced a statistically significant improvement in overall sleep quality as well as six of its seven components, the average score on the PSQI at post-assessment was above "5", indicating that participants were still having some problems with sleep. Additionally. participants did not significantly decrease their use of sleep medication, as measured by a subcomponent of the PSQI. The daily check-in data indicated that the participants were using the app almost every day, so adherence to playing the game was not an issue. In other words, the participants played the game daily, thus simply providing more opportunities for knowledge acquisition does not explain SHE's limited efficacy.

When placing these results in context within the larger literature, there are relatively few studies that have investigated the efficacy of sleep hygiene interventions in nonclinical samples. Overall, this small body of work provides some preliminary support for the use of sleep hygiene education in nonclinical populations, but the findings are inconsistent. Sleep hygiene education did not significantly improve sleep quality in adolescents (49, 50) or a sample of IT employees (51). On the other hand, sleep hygiene education was successful in improving sleep quality in athletes (52, 53), female workers (54), and women with HIV or AIDS (55). Restful Journey produced effects more in line with the latter studies demonstrating modest improvements in sleep quality.

Although outside the scope of the present study, differences in SHE efficacy among these different populations may suggest an effect of individual differences (56). In the present study, nearly one quarter of participants had PSQI global scores of <5 at the post-assessment, indicating a positive response to playing the game in these individuals. Although the personal characteristics that might predict who would positively respond to game-based learning are not known at this time, this percentage indicates that a subset of the general population might respond to a game-based application, a concept sometimes known as "person-centered care". In fact, a unique strength of serious games is the ability to implement machine learning (57), thus allowing the game to tailor SHE to each individual's unique needs. Future work may examine whether individual differences are indeed predictive of SHE outcomes, and whether tailoring goals and learning outcomes through the implementation of machine learning algorithms into game-based SHE interventions may yield better results. Physicians generally agree that sleep quality decreases with increased age (58). In fact, an epidemiological survey of over 9,000 senior individuals found that more than 80% had at least one complaint related to their sleep quality (59). Thus, future work should also seek to identify whether age may moderate the relationship between SHE and sleep quality. Finally, although its efficacy in terms of improving sleep quality appears to be comparable to traditional SHE approaches, perhaps it boils down to personal preference; the game format of Restful Journey may appeal to certain individuals who may not be motivated to engage with traditional SHE formats. The "play" factor behind Restful Journey may make it more appealing and therefore potentially increase the uptake of SHE, even if the outcomes are ultimately similar to those of traditional SHE.

Finally, state anxiety scores also significantly decreased from pre to posttreatment after playing *Restful Journey*. Previous links

between sleep difficulties and anxiety have established that these conditions are often comorbid and have a bidirectional predictive relationship (60). Both problems have mental health implications. Therefore, it was hypothesized that state anxiety would improve after playing Restful Journey, which was supported, as participants had a significant decrease in state anxiety from pre- to post- assessment. This study is consistent with previous research that has demonstrated that improved sleep quality can lessen state anxiety.

Limitations

As a pilot investigation, this study has several limitations. The most obvious limitation is that it was not a randomized controlled trial and therefore did not have a control group. Comparing Restful Journey to the traditional form of SHE and a waitlist control group would help provide further evidence for the notion of a game-based SHE format. Another limitation of the present study is that there was no assessment for underlying sleep disorders among participants. In the general population, sleep disorders are largely undiagnosed and the disorder is often complex. There are a number of reasons for this underdiagnosis; (1) limited training in the recognition of sleep disorders, (2) an uncertainty of how to treat sleep disorders, and (3) the patient and/or provider fails to mention sleep issues (61). It is therefore possible that some study participants could have more serious sleep disorders, such as sleep apnea, which may limit the extent to which SHE alone can improve their sleep quality. Additionally, it is unknown whether participants may have sought other sleep interventions during the study. Therefore, we cannot be certain if the effects could be attributed to other interventions or treatments. Finally, this study was limited to thirty days of app usage. It is not known (1) if longer use of the app may have resulted in continued improvement in sleep or (2) whether the significant improvements were stable several months later. Future investigations should address these issues.

Contributions and future directions

Despite these limitations, the present study found initial evidence for a game-based SHE format. The nature of the game-based interventions circumvents a number of barriers that individuals face when seeking relief for sleep problems; including ease of access, ease of use, minimal financial cost, and extensive opportunities to learn and practice these skills/behaviors in an engaging environment. The present study also demonstrated that state anxiety significantly decreased. Given that trouble sleeping can exacerbate anxiety (62), this decrease in anxiety is likely due to the improvement in sleep quality. As nearly one quarter of participants no longer met the criteria for being "poor sleepers" after using Restful Journey for 30 days, we propose that Restful Journey may be a tool for some individuals with non-clinical sleep problems to improve their sleep quality. Those that do not experience a significant improvement in sleep should seek more intensive treatment. Likewise, the addition of this application to more intensive clinical treatments may yield optimal results.

Future work may examine if the game may also reduce other mental health issues by improving overall sleep quality. Furthermore, future studies may seek to compare Restful Journey to the traditional method of SHE, which is typically a simple informational sheet. Finally, future work may examine the use of Restful Journey in combination with more intensive treatments (e.g., CBT-I) to improve treatment efficacy. Doing so would also provide insight as to whether game-based SHE is just as effective as the present gold-standard approach to alleviating sleep problems.

Data availability statement

The datasets presented in this article are not readily available because due to the funding source, there are restrictions that apply to the sharing of data. Reasonable requests will be considered on an individual basis. Please contact the corresponding author, CS, Ph.D. at christine.seaver@ucf.edu. Requests to access the datasets should be directed to christine.seaver@ucf.edu.

Ethics statement

The study involving humans was approved by University of Central Florida's Institutional Review Board. The study was conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

CS: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Writing – original draft. CB: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing. DB: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing. LH: Conceptualization, Funding acquisition, Writing – review & editing. SR: Conceptualization, Funding acquisition, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article.

This work is supported by the US Army Medical Research and Materiel Command under Contract Number W81XWH-17-C-0191. The views opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation. In the conduct of

research where humans are the subjects, the investigator(s) adhered to the policies regarding the protection of human subjects as prescribed by Code of Federal regulations (CFR) Title 45, Volume 1, Part 46; Title 32, Chapter, 1 Part 212; and Title 21, Chapter 1, Part 50 (Protection of Human Subjects). Article processing charges were provided in part by the UCF College of Graduate Studies Open Access Publishing Fund.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

References

1. Liao D, Liu J, Vgontzas AN, Rodriguez-Colon S, Calhoun S, Li X, et al. Cardiac autonomic modulation and sleep-disordered breathing in children. *Sleep Med Clin*. (2009) 4(1):27–36. doi: 10.1016/j.jsmc.2008.12.001

2. Fiorentini A, Valente R, Perciaccante A, Tubani LJ. Sleep's quality disorders in patients with hypertension and type 2 diabetes mellitus. *Int J Cardiol*. (2007) 114 (2):E50–2. doi: 10.1016/j.ijcard.2006.07.213

3. Stranges S, Dorn JM, Cappuccio FP, Donahue RP, Rafalson LB, Hovey KM, et al. A population-based study of reduced sleep duration and hypertension: the strongest association may be in premenopausal women. *J Hypertens*. (2010) 28(5):896–902. doi: 10.1097/HJH.0b013e328335d076

4. Rod NH, Vahtera J, Westerlund H, Kivimaki M, Zins M, Goldberg M, et al. Sleep disturbances and cause-specific mortality: results from the GAZEL cohort study. *Am J Epidemiol*. (2011) 173(3):300–9. doi: 10.1093/aje/kwq371

5. Narmandakh A, Roest AM, de Jonge P, Oldehinkel AJ. The bidirectional association between sleep problems and anxiety symptoms in adolescents: a TRAILS report. *Sleep Med*. (2020) 67:39–46. doi: 10.1016/j.sleep.2019.10.018

6. McGowan SK, Behar E, Luhmann M. Examining the relationship between worry and sleep: a daily process approach. *Behav Ther*. (2016) 47(4):460–73. doi: 10.1016/j.beth.2015.12.003

7. Marcusson-Clavertz D, Sliwinski MJ, Buxton OM, Kim J, Almeida DM, Smyth JM. Relationships between daily stress responses in everyday life and nightly sleep. *J Behav Med*. (2022) 45(4):518–32. doi: 10.1007/s10865-021-00281-3

8. Li Y, Gu S, Wang Z, Li H, Xu X, Wang F, et al. Relationship between stressful life events and sleep quality: rumination as a mediator and resilience as a moderator. *Front Psychiatry*. (2019) 10:449915.

9. McEwen BS, Karatsoreos IN. Sleep deprivation and circadian disruption: stress, allostasis, and allostatic load. *Sleep Med Clin*. (2015) 10(1):1–0. doi: 10.1016/j.jsmc.2014.11.007

10. Rosekind MR, Gregory KB, Mallis MM, Brandt SL, Seal B, Lerner D. The cost of poor sleep: workplace productivity loss and associated costs. *J Occup Environ Med*. (2010) 52(1):91–8. doi: 10.1097/JOM.0b013e3181c78c30

11. Ohayon MM. Epidemiology of insomnia: what we know and what we still need to learn. *Sleep Med Rev*. (2002) 6(2):97–111. doi: 10.1053/smrv.2002.0186

12. Centers for Disease Control and Prevention. Sleep Difficulties in Adults: Unites States, (2020). (cited January 17, 2024). Available online at: https://www.cdc.gov/nchs/products/databriefs/db436.htm

13. Irish LA, Kline CE, Gunn HE, Buysse DJ, Hall MH. The role of sleep hygiene in promoting public health: a review of empirical evidence. *Sleep Med Rev*. (2015) 22:23–36. doi: 10.1016/j.smrv.2014.10.001

14. Sun J, McPhillips MV, Chen KC, Zang Y, Li J, Oehlke J, et al. Primary care provider evaluation and management of insomnia. *J Clin Sleep Med*. (2021) 17 (5):1083–91. doi: 10.5664/jcsm.9154

15. Murawski B, Wade L, Plotnikoff RC, Lubans DR, Duncan MJ. A systematic review and meta-analysis of cognitive and behavioral interventions to improve sleep health in adults without sleep disorders. *Sleep Med Rev*. (2018) 40:160–9. doi: 10.1016/j.smrv.2017.12.003

16. Hauri P. *Current Concepts: The Sleep Disorders*. Naperville, IL: a Scope publication: Upjohn (1977).

17. Hauri PJ. Sleep hygiene, relaxation therapy, and cognitive interventions. In: Hauri PJ, editors. *Case Studies in Insomnia*. Boston, MA: Springer US (1991). p. 65–84. Available online at: http://link.springer.com/10.1007/978-1-4757-9586-8_5

18. Everitt H, McDermott L, Leydon G, Yules H, Baldwin D, Little P. GPs' management strategies for patients with insomnia: a survey and qualitative interview study. *Br J Gen Pract*. (2014) 64(619):e112–9. doi: 10.3399/bjgp14X677176

19. Tekinbaş KS, Zimmerman E. *Rules of Play: Games Design Fundamentals*. Cambridge, MA: The MIT Press (2004).

20. Nastasi BK, Clements DH. Motivational and social outcomes of cooperative computer education environments. *J Comput Childhood Educ*. (1993) 4(1):15–43.

21. Roussos M, Johnson A, Moher T, Leigh J, Vasilakis C, Barnes C. Learning and building together in an immersive virtual world. *Presence (Camb)*. (1999) 8(3):247–63. doi: 10.1162/105474699566215

22. Rieber LP. Seriously considering play: designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educ Technol Res Dev*. (1996) 44(2):43–58. doi: 10.1007/BF02300540

23. Lardinois C. Simulation, gaming and training in a competitive, multimodal, multicompany, intercity passenger-transportation environment. *J Oper Res Soc*. (1989) 40(10):849–61. doi: 10.1057/jors.1989.153

24. Romme AGL, Putzel R. Designing management education: practice what you teach. *Simul Gaming*. (2003) 34(4):512–30. doi: 10.1177/1046878103258202

25. Baltra A. Language learning through computer adventure games. *Simul Gaming*. (1990) 21(4):445–52. doi: 10.1177/104687819002100408

26. Pange J. Teaching probabilities and statistics to preschool children. *Inf Technol Childhood Educ Annual*. (2003) 15:163–72.

27. Bowers C, Procci K, Joyce R, Verduin M, LaRowe S, Mynck H, et al. Serious games for therapy: a training perspective. *J Cyber Ther Rehabil*. (2011) 4 (4):447–53.

28. Sipiyaruk K, Gallagher JE, Hatzipanagos S, Reynolds PA. A rapid review of serious games: from healthcare education to dental education. *Eur J Dent Educ*. (2018) 22(4):243–57. doi: 10.1111/eje.12338

29. Freitas S, Jarvis S. *Towards a Development Approach to Serious Games. Games-Based Learning Advancements for Multi-Sensory Human Computer Interfaces: Techniques and Effective Practices*. Hershey, PA: IGI Global (2009). p. 215–31.

30. Wilson KE, Miller AL, Bonuck K, Lumeng JC, Chervin RD. Evaluation of a sleep education program for low-income preschool children and their families. *Sleep*. (2014) 37(6):1117–25. doi: 10.5665/sleep.3774

31. Almondes KM, Leonardo ME. Games as education tools. *Sleep Sci*. (2019) 12 (2):100. doi: 10.5935/1984-0063.20190069

32. Tapia JL, Puertas FJ, Duñabeitia JA. Digital therapeutics for insomnia: assessing the effectiveness of a computerized home-based cognitive stimulation program. *J Integr Neurosci*. (2023) 22(2):34. doi: 10.31083/j.jin2202034

33. Chung KF, Lee CT, Yeung WF, Chan MS, Chung EWY, Lin WL. Sleep hygiene education as a treatment of insomnia: a systematic review and meta-analysis. *Fam Pract*. (2018) 35(4):365–75. doi: 10.1093/fampra/cmx122

34. Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Res*. (1989) 28(2):193–213. doi: 10.1016/0165-1781(89)90047-4

35. Buysse DJ, Hall ML, Strollo PJ, Kamarck TW, Owens J, Lee L, et al. Relationships between the Pittsburgh sleep quality index (PSQI), epworth sleepiness scale (ESS), and clinical/polysomnographic measures in a community sample. *J Clin Sleep Med*. (2008) 4(6):563–71. doi: 10.5664/jcsm.27351

36. Zheng S, Feng J, Lin R, Yan Y, Zhang R, Huang H, et al. The impact of negative mood state on sleep-related attentional bias in insomnia. *J Sleep Res*. (2019) 28(2): e12748. doi: 10.1111/jsr.12748

37. American Academy of Sleep Medicine. Healthy Sleep Habits. (Place unknown: American Academy of Sleep Medicine). (2016 Oct 1) (updated 2020 Aug; cited Jan 18, 2024). Available online at: https://sleepeducation.org/healthy-sleep/healthy-sleep-habits/

38. Carpenter JS, Andrykowski MA. Psychometric evaluation of the Pittsburgh sleep quality index. *J Psychosom Res*. (1998) 45(1):5–13. doi: 10.1016/S0022-3999(97)00298-5

39. Hinz A, Glaesmer H, Brähler E, Löffler M, Engel C, Enzenbach C, et al. Sleep quality in the general population: psychometric properties of the Pittsburgh sleep quality index, derived from a German community sample of 9284 people. *Sleep Med*. (2017) 30:57–63. doi: 10.1016/j.sleep.2016.03.008

40. Mollayeva T, Thurairajah P, Burton K, Mollayeva S, Shapiro CM, Colantonio A. The Pittsburgh sleep quality index as a screening tool for sleep dysfunction in clinical and non-clinical samples: a systematic review and meta-analysis. *Sleep Med Rev*. (2016) 25:52–73. doi: 10.1016/j.smrv.2015.01.009

41. Horsch CH, Lancee J, Griffioen-Both F, Spruit S, Fitrianie S, Neerincx MA, et al. Mobile phone-delivered cognitive behavioral therapy for insomnia: a randomized waitlist controlled trial. *J Med Internet Res*. (2017) 19(4):e70. doi: 10.2196/jmir.6524

42. Martin JL, Song Y, Hughes J, Jouldjian S, Dzierzewski JM, Fung CH, et al. A four-session sleep intervention program improves sleep for older adult day health care participants: results of a randomized controlled trial. *Sleep*. (2017) 40(8):z s x 079. doi: 10.1093/sleep/zsx079

43. Murawski B, Plotnikoff RC, Rayward AT, Oldmeadow C, Vandelanotte C, Brown WJ, et al. Efficacy of an m-health physical activity and sleep health intervention for adults: a randomized waitlist-controlled trial. *Am J Prev Med*. (2019) 57(4):503–14. doi: 10.1016/j.amepre.2019.05.009

44. Schlarb AA, Friedrich A, Claßen M. Sleep problems in university students–an intervention. *Neuropsychiatr Dis Treat*. (2017) 13:1989–2001. doi: 10.2147/NDT.S142067

45. Spielberger CD. *State-trait Anxiety Inventory for Adults*. Washington, DC: APA PsycTests (1983). doi: 10.1037/t06496-000.

46. Spielberger CD, Reheiser EC, Owen AE, Sydeman SJ. Curiosity in treatment planning and outcomes assessment. In: Maruish ME, editor. *The Use of Psychological Testing for Treatment Planning and Outcomes Assessment: Volume 3: Instruments for Adults*. New York, NY: Routledge (2004). p. 421.

47. Faul F, Erdfelder E, Lang AG, Buchner A. G*power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. (2007) 39(2):175–91. doi: 10.3758/bf03193146

48. Sullivan GM, Feinn R. Using effect size—or why the *P* value is not enough. *J Grad Med Educ*. (2012) 4(3):279–82. doi: 10.4300/JGME-D-12-00156.1

49. De Sousa IC, Araújo JF, De Azevedo CV. The effect of a sleep hygiene education program on the sleep-wake cycle of Brazilian adolescent students. *Sleep Biol Rhythms*. (2007) 5:251–8. doi: 10.1111/j.1479-8425.2007.00318.x

50. Otsuka Y, Kaneita Y, Spira AP, Mojtabai R, Itani O, Jike M, et al. Trends in sleep problems and patterns among Japanese adolescents: 2004–2017. *Lancet Reg Health–Western Pacific*. (2021) 9:100107. doi: 10.1016/j.lanwpc.2021.100107

51. Kakinuma M, Takahashi M, Kato N, Aratake Y, Watanabe M, Ishikawa Y, et al. Effect of brief sleep hygiene education for workers of an information technology company. *Ind Health*. (2010) 48(6):758–65. doi: 10.2486/indhealth.MS1083

52. Van Ryswyk E, Weeks R, Bandick L, O'Keefe M, Vakulin A, Catcheside P, et al. A novel sleep optimisation programme to improve athletes' well-being and performance. *Eur J Sport Sci*. (2017) 17(2):144–51. doi: 10.1080/17461391.2016.1221470

53. O'Donnell SH, Driller MW. Sleep-hygiene education improves sleep indices in elite female athletes. *Int J Exerc Sci*. (2017) 10(4):522.

54. Chen PH, Kuo HY, Chueh KH. Sleep hygiene education: efficacy on sleep quality in working women. *J Nurs Res*. (2010) 18(4):283–9. doi: 10.1097/JNR.0b013e3181fbe3fd

55. Hudson AL, Portillo CJ, Lee KA. Sleep disturbances in women with HIV or AIDS: efficacy of a tailored sleep promotion intervention. *Nurs Res*. (2008) 57 (5):360–6. doi: 10.1097/01.NNR.0000313501.84604.2c

56. McAlpine T, Mullan B, Clarke PJ. Assessing the daily association of sleep hygiene behaviours with sleep: a between and within persons approach. *J Behav Med*. (2024) 47(2):255–70. doi: 10.1007/s10865-023-00448-0

57. Gombolay MC, Jensen RE, Son SH. Machine learning techniques for analyzing training behavior in serious gaming. *IEEE Trans Games*. (2017) 11(2):109–20. doi: 10.1109/TCIAIG.2017.2754375

58. Espiritu JR. Aging-related sleep changes. *Clin Geriatr Med*. (2008) 24(1):1–4. doi: 10.1016/j.cger.2007.08.007

59. Foley DJ, Monjan AA, Brown SL, Simonsick EM, Wallace RB, Blazer DG. Sleep complaints among elderly persons: an epidemiologic study of three communities. *Sleep*. (1995) 18(6):425–32. doi: 10.1093/sleep/18.6.425

60. Webb E, Ashton C, Kelly P, Kamali F. Alcohol and drug use in UK university students. *Lancet*. (1996) 348(9032):922–5. doi: 10.1016/s0140-6736(96)03410-1

61. Roth T, Zammit G, Kushida C, Doghramji K, Mathias SD, Wong JM, et al. A new questionnaire to detect sleep disorders. *Sleep Med*. (2002) 3(2):99–108. doi: 10.1016/s1389-9457(01)00131-9

62. Pires GN, Bezerra AG, Tufik S, Andersen ML. Effects of acute sleep deprivation on state anxiety levels: a systematic review and meta-analysis. *Sleep Med*. (2016) 24:109–18. doi: 10.1016/j.sleep.2016.07.019

frontiers | Frontiers in Digital Health

TYPE Curriculum, Instruction, and Pedagogy
PUBLISHED 18 April 2024
DOI 10.3389/fdgth.2024.1307817

OPEN ACCESS

EDITED BY
Edith Talina Luhanga,
Carnegie Mellon University Africa, Rwanda

REVIEWED BY
Jeanne Markram-Coetzer,
Central University of Technology, South Africa
Vera Colombo,
National Research Council (CNR), Italy

*CORRESPONDENCE
Michael Loizou
michael.loizou@plymouth.ac.uk

RECEIVED 05 October 2023
ACCEPTED 03 April 2024
PUBLISHED 18 April 2024

CITATION
Loizou M, Arnab S, Lameras P, Hartley T, Loizides F, Kumar P and Sumilo D (2024) Designing, implementing and testing an intervention of affective intelligent agents in nursing virtual reality teaching simulations—a qualitative study.
Front. Digit. Health 6:1307817.
doi: 10.3389/fdgth.2024.1307817

Designing, implementing and testing an intervention of affective intelligent agents in nursing virtual reality teaching simulations—a qualitative study

Michael Loizou[1]*, Sylvester Arnab[2], Petros Lameras[2], Thomas Hartley[3], Fernando Loizides[4], Praveen Kumar[5] and Dana Sumilo[6]

[1]Centre for Health Technology, University of Plymouth, Plymouth, United Kingdom, [2]Centre for Postdigital Cultures, Coventry University, Coventry, United Kingdom, [3]Department of Computer Science, University of Wolverhampton, Wolverhampton, United Kingdom, [4]School of Computer Science and Informatics, Cardiff University, Cardiff, United Kingdom, [5]School of Health and Social Wellbeing, University of the West of England, Bristol, United Kingdom, [6]Warwick Medical School, University of Warwick, Warwick, United Kingdom

Emotions play an important role in human-computer interaction, but there is limited research on affective and emotional virtual agent design in the area of teaching simulations for healthcare provision. The purpose of this work is twofold: firstly, to describe the process for designing affective intelligent agents that are engaged in automated communications such as person to computer conversations, and secondly to test a bespoke prototype digital intervention which implements such agents. The presented study tests two distinct virtual learning environments, one of which was enhanced with affective virtual patients, with nine 3rd year nursing students specialising in mental health, during their professional practice stage. All (100%) of the participants reported that, when using the enhanced scenario, they experienced a more realistic representation of carer/patient interaction; better recognition of the patients' feelings; recognition and assessment of emotions; a better realisation of how feelings can affect patients' emotional state and how they could better empathise with the patients.

KEYWORDS

affective, virtual, reality, VR, nursing, intelligent, teaching, simulation

1 Introduction

Simulation is an educational tool that is becoming increasingly prevalent in nursing education (1) provides students with realistic clinical situations allowing them to practice and learn in a safe environment (2). Human patient simulators provide experiences that are realistic and offer students an opportunity to assess, intervene, and evaluate patient outcomes (3). Findings from a meta-analysis suggests that simulation education improved nursing students' performance in clinical reasoning skills (4).

A considerable amount of literature focuses on the use of digitally-mediated tools in healthcare. For example, a systematic review of 12 studies investigated the effectiveness of dementia carer-oriented interventions delivered through the internet and showed improvements in carer wellbeing (5). The review highlighted that, multicomponent individually tailored programmes that combine information, caregiving strategies, and contact with other carers can increase their confidence and self-efficacy, reducing stress

and depression. The evolution of nursing education teaching technologies has witnessed a shift towards interactive and immersive learning experiences. Building upon the observation by Baysan et al. (6) regarding the prevalent use of videos, recent studies have explored the integration of virtual reality (VR) and augmented reality (AR) simulations in nursing education. For instance, a study by Nakazawa et al. (7) shows that training with AR is effective in enhancing caregivers' physical skills and fostering greater empathy towards their patients. This approach not only enhances learner engagement but also facilitates active learning and skill retention.

However, despite the potential benefits of digital health interventions, addressing the emotional challenges and stressors faced by participants remains a critical concern. Research has highlighted the impact of environmental factors and basic emotions on user experience and outcomes in health teaching simulations (8, 9). Another study found undergraduate students felt unready, were anxious about having their mistakes exposed, worried about damaging teamwork and were afraid of evaluation (10).

According to Rippon (9), anger in healthcare environments is a very common emotion, leading to aggression and violence, and healthcare professionals are exposed to it daily with an increasing number of them suffering from signs of post-traumatic stress disorder. Although these problems have been identified and the potential of technologies has been widely recognised, most Human-Computer Interaction applications tend to overlook these emotions when responding to user input (11).

Thus, there is an increasing need for computer systems to endow affective intelligent agents with emotional capabilities and socially intelligent behaviour, and the aim of having an affective interaction between virtual and human users and an effect on the affective state of the user (12–14). This article proposes the design, implementation, and evaluation of a model for incorporating emotional enhancements (concentrating on negative emotions such as stress, fear, and anxiety) into virtual agents applied to virtual teaching applications for healthcare provision. For this purpose, we have created the following research objectives:

- RO1—To what extent does a virtual learning scenario incorporating emotional virtual agents provide a realistic experience.
- RO2—To what extent does the incorporation of emotions into virtual agents stimulate a better set of responses from the human user.
- RO3—To what extent do emotional virtual agents improve the learning experience.
- RO4—To what extent do emotional virtual agents allow the learners recognise their emotions and understand how these affect the virtual agents and, subsequently, allow them to feel more empathy towards them.

2 Related work underlying the educational activity

Although it has been increasingly accepted that emotions are an important factor in improving human-machine interaction, many times digital teaching systems do not consider the emotional dimension that human users expect to encounter in an interaction, and this can lead to frustration (11). To be able to provide affective interaction between intelligent agents and human users, the computer system needs to provide the virtual agents with capabilities for emotional and socially intelligent behaviour which should, in real time, have a measurable effect on the user (15).

Emotion recognition has been widely explored in human-robot interaction (HRI) and affective computing (16). Recent works aim to design algorithms for classifying emotional states from different input modalities, such as facial expression, body language, voice, and physiological signals (17, 18). Ability to recognise human emotional states can encourage natural bonding between humans and robotic artifacts (19).

Andotra (20) found that by integrating emotional intelligence with technological capabilities, chatbots or virtual agents could enhance user engagement and well-being through personalised and empathetic interactions. To enhance conversational skills integrating emotional capabilities in chatbots is essential. AI-driven chatbots can detect user sentiments in a conversation, thus triggering the chatbot to comprehend the user's emotional state and generate an appropriate response (21).

There are several models for agent decision-making in interactive virtual environments, such as interactive healthcare teaching applications, for example:

- Belief, Desire, and Intention (BDI) (22), explained in more detail below
- Psi-theory (23), OCC (24) which models the human mind as an information processing agent, controlled by a set of basic physiological, social and cognitive drives
- and the Five factor Model (FFM), a grouping of five unique characteristics used to study personality: conscientiousness, agreeableness, neuroticism, openness to experience, extraversion.

Models are needed because agents are part of an environment that has limited resources, such as memory and computational power. This means that agents cannot take an infinite amount of time to process their next move(s). One of the most prominent models is the BDI model that uses a set of rules designed to reduce the number of possible actions the agent can perform, thus reducing the time needed for the agent to decide its next action. The environment within which the agents act changes continuously so the agent must decide upon its next move and act upon it quickly enough so that it ensures that the environment does not change in such a way that it may invalidate the selected action. BDI constrains the reasoning needed by the agent and, in consequence, the time needed for making a decision. Within BDI beliefs represent beliefs about oneself, other agents, and the world. Desires represent what the agent would like to achieve, such as situations it would like to accomplish. For example, an agent could desire a patient in a simulation to be helped with their meal. Intentions represent actions/plans the agent has committed to do.

Research on agent design architectures and emotions for virtual teaching of personnel for stressful situations focuses on emotion

integration, the design of affective intelligent agents that interact with other agents (virtual or real) based on social norms and on how to improve agent realism and believability. Towards this goal some of the research combines and/or extends other theories and models.

For example, the extended BDI (EBDI) model (25), adds the influence of primary (first, more instinctive, emotional response to an event e.g., fear, sadness, anger) and secondary emotions (arising from higher cognitive processes and based on what we perceive that the result of a certain situation will be e.g., joy and relief). This makes agents more engaging and believable so that they can better play a human-like role in affective intelligent agent scenarios.

In addition to the user's emotions, virtual agents in our model (described in the next section—Architecture Overview) also adapt their actions and emotions based on the user's personality traits. The cognitive-behavioural architecture (26) incorporates the notion of personality as a determining factor for the agent's future emotions and coping preferences. This approach enables improved character diversity and personality coherence across different situations.

3 Design of the online learning environment

We present an architecture for creating a virtual tutor with the ability to alter both the tutor's and the virtual patient's behaviour and mood based on the responses, emotions, and personality of human participants. Responses to human participant's actions are selected from a database of possible rules that are based on previous research and input from psychologists and experts in the area of health training. The testing scenario used in this research focuses on teaching nurses caring for dementia patients.

The proposed architecture makes use of the BDI model and extends it by incorporating the emotions and personality of the trainees in the decision-making process. BDI is one of the most popular models for developing rational agents based on how humans act based on the information derived from an environment (27). BDI was chosen because it is one of the most prominent approaches to building agent and multiagent systems, it is based on a sound psychological theory, Michael Bratman's theory of human practical reasoning (28) and it has been successfully used for the design of realistic affective intelligent agents for many years (29–31).

The architecture proposed in this work extends BDI by incorporating two new elements. Firstly, it incorporates the current emotional state of the participant into its decision-making process. Systems that have explored emotion integration, such as the BDI tutor model (29) and Model Social Game Agent (MSGA) (30) have focused on just incorporating the emotions of trainees and do not take other factors into account, such as a human participant's personality traits. Our proposed model focuses on individual users and on the emotions of both the human participants and of the virtual agents. The emotional state of the human participant is not approximated by the model but is the dependant variable and reported by the participants in real time in order to convey the emotional state to the simulation for it to adapt. This introduces realism and personalises the process for the human participant as it focuses on the individual's emotional state and how this changes throughout the teaching session. With advances in artificial intelligence, virtual agents are starting to play an active role in various fields such as information presentation, sales, training, education, and healthcare (32, 33). To help induce a positive emotion and thereby enhance social bonding in human–virtual agent interaction, affective communication between people and virtual agents has become important (34).

Secondly, our architecture incorporates the personality of the participant. For the purposes of this research the Five Factor Model (also known as the "Big 5") personality test (35) is used for measuring the participant's level of neuroticism. This has been suggested to be related to the level that negative emotions affect an individual during training (36–38). The feedback provided to the participant as well as the virtual patient's emotional state is affected by both these elements.

Regarding the reasoning rule-based model, the Jboss DROOLS Expert rules engine is used for providing feedback to the user and uses the adapted BDI model for providing emotional intelligence feedback. The front end of the learning environment has been designed using the UNITY3D game development system and programmed using C# to communicate with the server back end. The back end was programmed using JAVA and runs under Tomcat so that several front-end sessions can communicate concurrently with and between modules on the server.

3.1 Architecture overview

A high-level overview of the emotional virtual agent's architecture can be seen in Figure 1. The architecture comprises 4 sections. The "User" section illustrates the real world and the participant taking part in the virtual teaching simulation. The "Agent Interface" section illustrates the communication between the participant and the virtual agent. The "Cloud Communication" section illustrates the data communication between the virtual teaching environment and the emotional agent architecture on the server. Finally, the "Evaluation" section illustrates the decision-making process based on the BDI system, available plans, human participant emotions and personality. Once inputted, the events, emotions and personality characteristics of the participant are sent from the virtual learning environment to the server, as shown in the "Cloud Communication" section.

The interaction between the participant and the agent interface, shown in the "User" and "Agent Interface" sections of Figure 1, is performed using traditional input devices, such as the keyboard and mouse. The participants take part in an interactive 3D virtual teaching scenario. The communication between the participant and the virtual tutor/patient is achieved using dialog boxes. The virtual tutor has no embodiment within the virtual environment. It presents messages and feedback to the participant via text boxes.

FIGURE 1
High-level overview of the proposed emotional virtual agent's architecture.

The virtual patient is represented within the virtual environment by a 3D human model. They are controlled using scripted behaviours; however, the virtual tutor can alter the selection of behaviour and the mood based of the virtual patient based on the responses, emotions, and personality of human participant.

The participants have a first-person view of the virtual teaching environment and select their choice of action using selection boxes, as shown in Figure 2. Participants input their personality type and their current emotional status by selecting the type and intensity of their emotions, using selection grids. This method of emotional input is used as it is a quick and reliable way for collecting information on the emotional status of a trainee in a real time environment.

3.2 Adapting the BDI model

The BDI model is used as the basis for interactive virtual environment decision-making in the emotional virtual agent's architecture. The virtual tutor's beliefs are created and updated based on its desires and current scenario events. The tutor uses dialogue boxes for teaching the participants and prompting them to perform different actions. The participant responds by performing the required actions and replies to the tutor using multiple choice and true/false responses.

As illustrated in Figure 1, the decision-making process consists of the following modules:

- The agent's Beliefs module accepts as input events from the current scenario and the agent's desires. Based on these inputs, the module outputs a list of the agent's beliefs to the agent's desires and intentions modules. For example, a scenario input could be a virtual patient that needs help finishing his meal and the tutor's desire to provide the needed help to the patient. The outputs of the module would include the intention to ask a participant to help the patient and the desire to follow through by confirming at a later time that the participant fulfilled her task.
- The agent's Desires module accepts as input a list of the agent's beliefs and outputs a list of the agent's desires to both the agent's beliefs and agent's intentions modules. Based on the example illustrated above, the module could receive the desire of the agent to confirm that the participant helped the virtual patient finish his meal. The module would then output the intention of asking the participant at a later time if she fulfilled the task of helping the virtual patient and if the patient did, indeed, manage to finish his meal.
- The agent's Intentions module accepts as input a list of possible plans, the agent's desires and the agent's beliefs. It outputs a list of possible actions to the reasoning module. For example, inputs could include the plan of the tutor to ask the participant to help the virtual patient finish his meal by helping him sit at the table and cut the meat in manageable pieces, the desire of the tutor to help the patient and the belief that the patient needs the participant's help for finishing his meal.
- The agents Reasoning Module accepts as inputs a list of possible participant responses to the tutor's question and the participant's emotions and personality traits. It outputs the chosen actions to be communicated to the participant and a set of gestures and facial expressions to be used by the virtual patient's 3D avatar. For example, if a participant's inputs indicate that they are tired and not very empathetic then the tutor would ask her in a firmer way to help the virtual patient. The tutor would be less firm when communicating with a rested participant that is more likely to feel empathy for the patient that has difficulty finishing his meal.

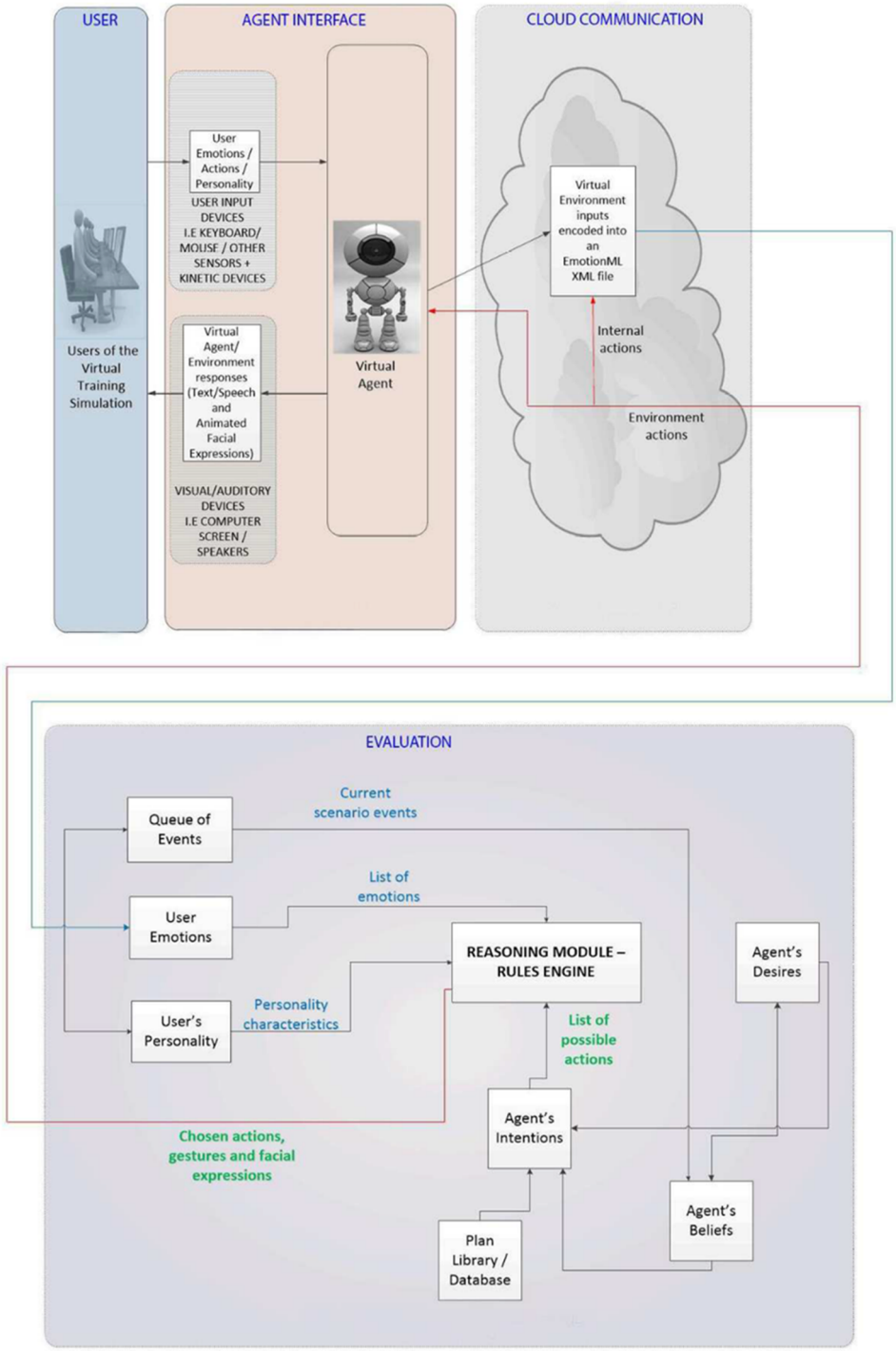


FIGURE 2
A screenshot of the virtual teaching environment and an interaction with the virtual agent. Please enter a scale for the emotion Stress and then reply to the question below—Current Value 0; Please enter a scale for the emotion Agitation and then reply to the question below—Current Value 0; Please select Emotions; (Q1 of 5) Should I be here in this place?:- I know you are worried about being here in hospital. Did you need to use the toilet?; You are in hospital and you have to stay in for quite some time; Please stop talking to me and interrupting me while I am trying to do the medicine round.

3.3 Incorporating emotions and personality into belief–desire–intention model

Participant emotions and personality in the proposed architecture are inputted by the participant using the keyboard/mouse within the virtual learning environment. These are then transferred over a network to the server/back-end to be interpreted by the inference engine. The participant's input their emotions and personality by replying to multiple choice questions, selecting one or more emotions and their intensity

(using a Likert scale). As discussed above, for this research the Five Factor Model (also known as the "Big 5") personality test (35) is used for measuring the participant's level of neuroticism.

The participant's personality and emotions and the virtual agent's desires, intentions and beliefs are the inputs to the emotion creation and revision module. Emotions affect the best possible way an agent can respond to his requests. Personality affects the intensity and duration of these emotions. For example, using the Myers Briggs Personality Types, a Facilitator Caretaker incorporating the following Myers-Briggs Personality Preferences: Extraversion Sensing Feeling Judging, is more likely to feel compassion and sadness for a patient than a Conceptualizer Director that incorporates the following preferences: Introversion, iNtuition, Thinking, Judging.

The virtual agents in our model predict how to respond to the indicators of a participant's emotions depending on an inference of the participant's personality. The participant's neuroticism level is used then to adjust the scale grade of the reported emotion. As people with higher levels of neuroticism tend to over-report their negative emotions, the rule is to decrease the level of emotion reported on an 1–10 scale from participants with higher than average levels of neuroticism and lower the reported level for participants with lower than average levels of neuroticism.

3.4 The reasoning module

A set of rules adapts the responses of the virtual tutor based on the trainee's emotions, emotion strength, and personality traits.

Our system outputs the feedback to be given to the participant, and information that identifies how the participant's response affected the emotional state of the virtual patient. This information can then be used to adjust the virtual patient's behaviour and animations. For example, an animation can be played, and an emoticon can be displayed that show the patient's resulting emotional state. This approach allows the system to decide which actions will be performed by the virtual agent and what types of gestures, facial expressions and emoticons will be used for interacting with the participant within the virtual teaching environment. In a similar manner, depending on the virtual patient feedback, which is directly related to the emotion felt by the patient, an emoticon is displayed on the screen that shows clearly how the patient feels. The possible emotions portrayed are happiness, sadness, disgust, fear and anger. Figure 3 illustrates an example of sadness for the virtual patient.

4 Results

The chosen scenario for the data collection was designed by a team of nursing tutors that specialise in mental health and, specifically, in treating patients with dementia. The participants for the case studies consisted of nine nursing students with an interest in mental health. The data collection consisted of:

1. A pre-study interview that gathered background information about the participant;
2. An online personality test (Five Factor Model). This was used to determine the level of neuroticism of the participants;
3. Two teaching sessions. During the sessions, the participants interacted with a virtual patient with dementia and responded to common questions asked by such patients. During the interaction, interruptions occurred that had to be dealt with (these were random in occurrence and duration, in the same way these would occur in the real world), such as another patient asking for water or the fire alarm going off. These interruptions were chosen based on input from experienced nurses. The training comprised two different scenarios:
   a. Scenario 1—A teaching scenario where the feedback given to the participants did not depend on their emotions and personality and the patient did not demonstrate any visual or auditory change in their mood. This session was delivered as a point of reference to the participant on how systems that do not cater for different personalities and emotions work;
   b. Scenario 2—an updated system based on the model designed for this research; the second scenario provided visual (emoticons and animations) and auditory feedback on the patient's emotional state. These were based on the trainee's emotions, personality, and responses to the patient's questions.
4. An interview was conducted for gathering data on how the teaching session affected the learning process, emotional states and general satisfaction of the participants. Interview questions (Appendix A) were semi-structured and the participants were prompted to discuss and expand their replies.

A set of qualitative data was collected using observation and semi-structured interviews. The data was used to draw out patterns of how visual and auditory representations of a virtual patient's mood may affect the learning process, emotional states, and general satisfaction of the trainees.

- **Data analysis**

This section presents and interprets the qualitative results from the testing of the developed emotional virtual agent's architecture.

The participants were selected using convenience sampling from a 3-year nursing course at a university. Students were in their early 20s, 7 were female and 2 male, a representative ratio of nursing students according to statistics (39). Inclusion criteria included that the participants were taking part in their hospital placements in the mental health ward and were working with patients with dementia. The study was approved by the university's ethical committee and the participants signed a consent statement before participating. The sample size was deemed to be large enough for a small-scale qualitative test of the technology to evaluate feasibility and acceptability. A future quantitative study with a larger number of participants will be used to evaluate statistical significance.

A summative content analysis was first used to identify themes related to the research objectives. In the table (Appendix B) the number of occurrences that support each research theme for

FIGURE 3
User of animations, sounds and emoticons for portraying the emotion of sadness for the virtual patient. "A few minutes for someone with dementia is meaningless. Mrs Smith may be unable to work out how long that might be. The response may make her feel agitated that she has to wait. She may be left feeling that she is not being taken seriously. This may lead to her getting up out of her chair very quickly and asking the same question. PLEASE TRY AGAIN."

each of the nine participants are listed. The research objectives (RO) are as follows:

- RO1—To what extent does a virtual learning scenario incorporating emotional virtual agents provide a realistic experience.
- RO2—To what extent does the incorporation of emotions into virtual agents stimulate a better set of responses from the human user.
- RO3—To what extent do emotional virtual agents improve the learning experience.
- RO4—To what extent do emotional virtual agents allow the learners recognise their emotions and understand how these affect the virtual agents and, subsequently, allow them to feel more empathy towards them.

Below is a summary of the occurrences for each RO from interviewing the participants

Table 1 illustrates how the scenario that provided text responses that were affected by the participant's personality traits and emotions, in addition to using visual and auditory cues, was more successful in providing a more realistic experience, better responses, improved learning and increased empathy.

In more detail, Increased Empathy was the research objective that had the highest number of occurrences. This is a very important part of what we tried to achieve with our VR scenario educational design as we aimed to have a patient-centered

TABLE 1 Summary of the occurrences for each RO from interviewing the participants.

Theme Description	Realistic experience	Better responses	Improved Learning Experience	Increased Empathy
Interview occurrences Participant 1	4	4	4	6
Interview occurrences Participant 2	4	3	3	5
Interview occurrences Participant 3	2	3	3	3
Interview occurrences Participant 4	2	2	2	2
Interview occurrences Participant 5	1	1	2	2
Interview occurrences Participant 6	2	3	1	1
Interview occurrences Participant 7	7	2	1	2
Interview occurrences Participant 8	3	2	3	4
Interview occurrences Participant 9	1	1	4	3

educational design, something extremely important in general, but even more in the mental health sector, The next RO with 26 occurrences was Realistic Experience, something that a VR experience can certainly provide if designed correctly. Improved learning experience and Better Responses to patient questions followed closely with a still high 23 and 21 occurrences respectively, showing that all of the ROs were sufficiently achieved.

In the remainder of this section, we will provide a detailed analysis of the qualitative data collected for this study. The analysis is organised according to the research objectives outlined at the start of this section.

- **To what extent does a virtual learning scenario, incorporating emotional virtual agents provide a realistic experience**

All participants ($n = 9$) reported that the scenario questions were realistic, based on their experience of roleplaying teaching and actual first-hand experience of working with people with dementia during hospital placement, and made them feel that they were interacting with a real patient, especially in the case that the patients reacted in a way that made their emotions clear. They also felt that the realism of the scenario could help them improve their nursing skills, especially with the integrated distractions which corresponded to what they would expect to have to deal with in a real-world scenario. They found that the realism of the system could provide a good alternative to training with real people either at university or during their placements. The realistic sounds from the patients (laughter, screams etc.) were also found to be very useful and increased the realism of the experience.

One participant stated in the initial interview that took place before the teaching session that he believed that online teaching would never replace practicing with real people.

> "I'm a firm believer that simulated and online experiences will never replace the total quality of experiencing that in practice." (Participant 9).

After the session the same participant reported that he found the online teaching to be much more realistic than expected:

> "I think that's exactly the kind of challenges that a healthcare worker will have. They will go into a bay, they'll be approached by one patient but there will be others also demanding on their time, there'll be other important tasks and it's important to make sure that the patient is cared for and you also meet the needs of others. So, the scenarios felt very real to me." (Participant 9).

- **To what extent does the incorporation of emotions into virtual agents stimulate a better set of responses from the human user**

All the participants reported that the emotionally enhanced teaching scenario made them realise how their actions can affect the patient's mood and subsequent reactions and how by being calm and not letting their emotions guide their actions they can help the patients much more. They also agreed that seeing how happy the patient was because of their care also improved their mood. The affective virtual agents also made the participants reflect on their work, when the results were positive but, mostly, when they did not succeed in keeping the patient happy. For example, one participant stated:

> "… in one of the questions the patient was shouting, he wanted to get out of there. If I was shouting back it would have escalated, it would have made the patient more angry, whereas if I just calmly say we can help then at least the patient's hear that there is some help for them." (Participant 1).

They continued to say that reporting these emotions helped him try to control them:

> "… so I've got to keep it in and just find an easier, calmer way to say, 'Let's sit down and we'll sort it out.'" (Participant 1).

The participant also said that asking him to report these emotions helped him think about how he felt and prompted him to try and choose the answer that would make it less likely to transfer his negative emotions onto the patient.

> "Yeah, because there were some answers that I was reading, and I thought if I answer with that it will make the scenario escalate and make the patient more agitated and angry if I tell them, "Oh just sit down, I'll deal with you later." They don't feel like they're being cared for, so they'll just want to leave, they won't want to sit down and talk to you but if you say, 'Sit down, I'll get you a drink and we'll talk about how worried you are.'" (Participant 1).

The online teaching environment also helped them respond better as they felt more secure not having to deal with real patients while still at the initial stages of learning as the embarrassment of getting one question wrong in the real-world would sometimes make them get more wrong answers.

- **To what extent do emotional virtual agents improve the learning experience**

Again, all the participants stated that the feedback based on their responses made them understand clearly how their answers affected the patient and it allowed them to better understand how important it is to know how the patient feels at every stage of their interaction. They reported that the experience seemed realistic and close to what they would expect to do in real life, and this helped them learn better. It was also something they found more enjoyable and less stressful as real-life teaching scenarios sometimes made them feel uncomfortable and as they were judged by their tutor. They felt that these teaching sessions within a "safe" environment would allow them to practice and improve their skills before attempting to work with a real patient. For example:

> "So, like when we did it face to face with a member of staff in a way you could feel that you were being judged by that member of staff. I suppose to test out your skills first online then you're sort of in a safer environment than when you're face to face with somebody. Learning this online is more of a safe environment to get it wrong and where you can get your skills up to scratch really to be able to do it face to face with somebody." (Participant 3).

The teaching sessions also made them realise that this kind of teaching can also be applied to other patients with different conditions where it would also help improve the learning experience. Responses included:

> "I think that's the thing that's missing in practice is that nobody tells you, you are so driven within the moment of stress that you don't realise you are and you do need to take a step back. I think that's perhaps the difference here between the virtual environment and the real environment, there is no button in the real environment to press, to gauge my anxiety or my stress, to help me realise it, to help me control it." (Participant 9).

- **To what extent do emotional virtual agents allow the learners recognise their emotions and understand how these affect the virtual agents and, subsequently, allow them to feel more empathy towards them**

The extent to which affective virtual agents allow carers to increase their empathy was investigated. Both the replies to the interview questions and the observation from the researcher provided data that supported the research objective. More specifically the participants reported that they realised how patients can feel threatened when carers let their negative emotions show and, on the other hand, how by seeing a patient being calm and happy they know that they succeeded in giving them the right care and they did not let their stress, anger or other emotions affect their work. For example, one participant commented on how having to report how stressed and agitated he was before every reply to a question helped him understand better how those emotions could affect the patient.

> "Well, it made me really think of the patient. That it could escalate the patient's condition and make him feel threatened if the trainee let those emotions show." (Participant 1).

In the same way he noted that reporting those emotions help him think about them more, and subsequently made him try to control them.

> "Yeah, if I'm shouting, even though I might feel angry and just want to finish, if I show that it might make the patient feel threatened." (Participant 1).

He also stated that asking him to report those emotions helped think about how he felt and try to choose the right answer that wouldn't affect the patient negatively.

> "Yes, you may not have the time to and you may need to do other stuff but that way at least you've calmed the patient down a little." (Participant 1).

The reactions of the patient in the second scenario made them realise clearly how the patient felt and made them try even more to be compassionate and calm and not allow their actions to have a negative impact on the patient. All the participants agreed that a happy patient made them feel happy and contented too whereas a distressed patient affected them negatively. Responses included:

> "Yes, because if you feel you're affected emotionally by the situation you do think about what it is you've got to say next and what you're going to do next because you know that you don't want your emotions to impact on the patient because it's not about you, it's about the patient and that side of care." (Participant 4).

The results of the comparison between the two different systems are outlined below. The system that provided the visual and auditory representations of the virtual patient's emotions based on the participant's emotions, personality and responses was reported by the nice participants to have the following advantages:

1. Provided a more realistic representation of the carer/patient interaction;
2. Performed better in helping the carers:
    a. recognise how the patients feel;
    b. recognise and evaluate their own emotions;
    c. realise how their actions can affect the patient's emotional state;
    d. realise how their emotions can affect the patient's emotional state;
    e. empathise with the patients.

To summarise, all (100%) of the participants reported, when using the enhanced scenario, a more realistic representation of carer/patient interaction; better recognition of the patients' feelings; recognition and assessment of emotions; a better realisation of how feelings can affect patients' emotional state and how they could better empathise with the patients.

5 Discussion

Our aim was to explore how intelligent virtual agents in healthcare provision teaching simulations can improve the human participant's learning experience by incorporating visual and auditory representations of their emotional states. These emotional states and the virtual agent's responses were adapted based on an indication of the participant's emotions and personality.

We found that visual and auditory representations of the patient's emotional state based on our adapted BDI architecture positively affected the learning process, emotional states, and general satisfaction of trainee nurses. Two scenarios were used; one that did not include a virtual patient with different mood states and an updated system that provided visual and auditory feedback based on the patient's emotional state.

There was no negative feedback regarding the second, emotionally enhanced, scenario. When completing the emotionally enhanced scenario some of the participants even had a complete change in heart regarding the realism and usefulness of online learning environments. All the nurses reported that the second scenario using the enhanced learning system was more realistic, helped them better realise how their actions and emotions can affect a patient and, thus, made them more empathetic and improved their learning experience.

Rich data was collected from the nine participants that support that the adapted scenario with the emotionally enhanced patients allowed them to:

- Experience a more realistic representation of carer/patient interaction
- Better recognise the patients' feelings;
- Better realise how feelings can affect patients' emotional state
- Empathise with the patients.

These results will be useful for researchers in designing and conducting future studies relevant to a broader range of healthcare providers. By creating more varied scenarios focusing on different target groups, e.g., people with learning disabilities, people that suffered from stroke etc., the architecture can be tested further and modified, as/if necessary, to better cater for different health conditions.

Previous research in this area has focused on:

1. Non interactive videos where the participants did not have a hands-on experience, as we mentioned in the introduction of this paper regarding the systematic review of nursing education teaching technologies (6).
2. scenarios focusing on only the emotions of the participants and not on both their emotions and their personality (40, 41).

Our work can be further adapted to be used in different and more varied fields, including crisis management and with both formal (soldiers, firefighters, law enforcement officers) and informal personnel. This can be done by using a different set of emotions for the participants; these can be negative, positive or include some of each type of emotion depending on the field and scenarios used. In future work, we plan to incorporate online assessments instruments into the virtual scenario for collecting quantitative data on how emotionally enhanced virtual agents could improve the learning experience.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by University of Wolverhampton ethics committee. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

ML: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing – original draft, Writing – review & editing. SA: Writing – review & editing. PL: Writing – review & editing. TH: Conceptualization, Methodology, Writing – review & editing. FL: Validation, Writing – review & editing. PK: Writing – review & editing. DS: Conceptualization, Formal Analysis, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The authors declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Yuan HB, William BA, Fang JB, Ye QH. A systematic review of selected evidence on improving knowledge and skills through high-fidelity simulation. *Nurse Educ Today*. (2012) 32(3):294–8. doi: 10.1016/j.nedt.2011.07.010

2. Arthur C, Levett-Jones T, Kable A. Quality indicators for design and implementation of simulation experiences: a delphi study. *Nurse Educ Today*. (2012) 33(11):1357–61. doi: 10.1016/j.nedt.2012.07.012

3. Lee SO, Eom M, Lee JH. Use of simulation in nursing education. *J Korean Acad Soc Nurs Educ*. (2007) 13(1):90–4. pISSN: 1225-9578. eISSN: 2093-7814.

4. Shin S, Park JH, Kim JH. Effectiveness of patient simulation in nursing education: meta-analysis. *Nurse Educ Today*. (2015) 35(1):176–82. doi: 10.1016/j.nedt.2014.09.009

5. Boots LMM, de Vugt ME, van Knippenberg RJM, Kempen GIJM, Verhey FRJ. A systematic review of internet-based supportive interventions for caregivers of patients with dementia. *Int J Geriatr Psychiatry*. (2013) 29(4):331–44. doi: 10.1002/gps.4016

6. Baysan A, Çonoğlu G, Özkütük N, Orgun F. Come and see through my eyes: a systematic review of 360-degree video technology in nursing education. *Nurse Educ Today*. (2023) 128:105886. doi: 10.1016/j.nedt.2023.105886

7. Nakazawa A, Iwamoto M, Kurazume R, Nunoi M, Kobayashi M, Honda M. Augmented reality-based affective training for improving care communication skill and empathy. *PLoS One*. (2023) 18(7):e0288175. doi: 10.1371/journal.pone.0288175

8. Evans-Martin FF. *Emotion and Stress*. New York, USA: Infobase Publishing (2007).

9. Rippon TJ. Aggression and violence in health care professions. *J Adv Nurs*. (2000) 31(2):452–60. doi: 10.1046/j.1365-2648.2000.01284.x

10. Kang SJ, Min HY. Psychological safety in nursing simulation. *Nurse Educ*. (2019) 44(2):E6–9. doi: 10.1097/NNE.0000000000000571

11. Schröder M, Cowie R. Developing a consistent view on emotion-oriented computing. *Lecture Notes in Computer Science*. (2006):194–205. doi: 10.1007/11677482_17

12. Maples B, Pea RD, Markowitz D. Learning from intelligent social agents as social and intellectual mirrors. In: Niemi H, Pea RD, Lu Y, editors. *AI in Learning: Designing the Future*. Cham: Springer (2023). p. 73–89. doi: 10.1007/978-3-031-09687-7_5

13. Colombo S, Rampino L, Zambrelli F. The adaptive affective loop: how AI agents can generate empathetic systemic experiences. In: Arai K, editor. *Advances in Information and Communication. FICC 2021*. Advances in Intelligent Systems and Computing, Vol. 1363. Cham: Springer (2021). p. 547–59. doi: 10.1007/978-3-030-73100-7_39

14. Zall R, Kangavari MR. Comparative analytical survey on cognitive agents with emotional intelligence. *Cogn Comput*. (2022) 14:1223–46. doi: 10.1007/s12559-022-10007-5

15. Prendinger H. Human physiology as a basis for designing and evaluating affective communication with life-like characters. *IEICE Trans Info Syst*. (2005) E88-D(11):2453–60. doi: 10.1093/ietisy/e88-d.11.2453

16. Spezialetti M, Placidi G, Rossi S. Emotion recognition for human-robot interaction: recent advances and future perspectives. *Front Robot AI*. (2020) 7. doi: 10.3389/frobt.2020.532279

17. McColl D, Hong A, Hatakeyama N, Nejat G, Benhabib B. A survey of autonomous human affect detection methods for social robots engaged in natural HRI. *J Intell Robot Syst*. (2016) 82:101–33. doi: 10.1007/s10846-015-0259-2

18. Cavallo F, Semeraro F, Fiorini L, Magyar G, Sinčák P, Dario P. Emotion modelling for social robotics applications: a review. *J. Bionic Eng*. (2018) 15:185–203. doi: 10.1007/s42235-018-0015-y

19. Erol BA, Majumdar A, Benavidez P, Rad P, Raymond Choo K, Jamshidi M. Toward artificial emotional intelligence for cooperative social human–machine interaction. *IEEE Trans Comput Soc Syst*. (2020) 7:234–46. doi: 10.1109/TCSS.2019.2922593

20. Andotra S. Enhancing human-computer interaction using emotion-aware chatbots for mental health support. [Preprint]. (2023). doi: 10.13140/RG.2.2.20219.90409. [Epub ahead of print].

21. Bilquise G, Ibrahim S, Shaalan K. Emotionally intelligent chatbots: a systematic literature review. *Hum Behav Emerg Technol*. (2020) 2022:9601630. doi: 10.1155/2022/9601630

22. Georgeff M, Pell B, Pollack M, Tambe M, Wooldridge M. The belief-desire-intention model of agency. *Lecture Notes in Computer Science*. (1999):1–10. doi: 10.1007/3-540-49057-4_1

23. Dörner D, Güss CD. PSI: a computational architecture of cognition, motivation, and emotion. *Rev Gen Psychol*. (2013) 17(3):297–317. doi: 10.1037/a0032947

24. Clore GL, Ortony A. Psychological construction in the OCC model of emotion. *Emot Rev*. (2013) 5(4):335–43. doi: 10.1177/1754073913489751

25. Jiang H, Vidal JM, Huhns MN. *EBDI: an architecture for emotional agents. Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems—AAMAS '07* (2007). doi: 10.1145/1329125.1329139

26. Sandercock J, Padgham L, Zambetta F. Creating adaptive and individual personalities in many characters without hand-crafting behaviors. *Lecture Notes in Computer Science*. (2006):357–68. doi: 10.1007/11821830_29

27. Chodorowski J. A case study of adding proactivity in indoor social robots using belief-desire-intention (BDI) model. *Biomimetics (Basel)*. (2019) 4(4):74. doi: 10.3390/biomimetics4040074

28. Bratman ME, Israel DJ, Pollack ME. Plans and resource-bounded practical reasoning. *Comput Intell*. (1988) 4(3):349–55. doi: 10.1111/j.1467-8640.1988.tb00284.x

29. Florea A, Kalisz E. *Embedding emotions in an artificial tutor. Seventh International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC'05)* (2005). doi: 10.1109/synasc.2005.34

30. Johansson M, Verhagen H, Eladhari MP. *Model of social believable NPCs for teacher training: using second life. 2011 16th International Conference on Computer Games (CGAMES)* (2011). p. 270–4. doi: 10.1109/cgames.2011.6000351

31. Zoumpoulaki A, Avradinis N, Vosinakis S. A multi-agent simulation framework for emergency evacuations incorporating personality and emotions. In: Konstantopoulos S, Perantonis S, Karkaletsis V, Spyropoulos CD, Vouros G, editors. *Artificial Intelligence: Theories, Models and Applications. SETN 2010*. Lecture Notes in Computer Science, Vol. 6040. Berlin, Heidelberg: Springer. (2010). p. 423–8.doi: 10.1007/978-3-642-12842-4_54

32. D'Haro LF, Kim S, Yeo KH, Jiang R, Niculescu AI, Banchs RE, et al. CLARA: a multifunctional virtual agent for conference support and touristic information. In: Lee G, Kim H, Jeong M, Kim JH, editors. *Natural Language Dialog Systems and Intelligent Assistants*. Cham: Springer (2015). p. 233–9. doi: 10.1007/978-3-319-19291-8_22

33. Philip P, Micoulaud-Franchi JA, Sagaspe P, Sevin E, Olive J, Bioulac S, et al. Virtual human as a new diagnostic tool, a proof of concept study in the field of major depressive disorders. *Sci Rep*. (2017) 7:42656. doi: 10.1038/srep42656

34. Collins EC, Prescott TJ, Mitchinson B. *Saying it with light: a pilot study of affective communication using the MIRO robot. Biomimetic and Biohybrid Systems: 4th International Conference, Living Machines 2015, Barcelona, Spain, July 28–31, 2015, Proceedings 4*. (2015). Vol. 9222. p. 243–55. doi: 10.1007/978-3-319-22979-9_25

35. McCrae RR, John OP. An introduction to the five-factor model and its applications. *J Pers*. (1992) 60(2):175–215. doi: 10.1111/j.1467-6494.1992.tb00970

36. Gallagher DJ. Extraversion, neuroticism and appraisal of stressful academic events. *Pers Individ Dif*. (1990) 11(10):1053–7. doi: 10.1016/0191-8869(90)90133-c

37. Sambol S, Suleyman E, Scarfo J, Ball M. Distinguishing between trait emotional intelligence and the five-factor model of personality: additive predictive validity of emotional intelligence for negative emotional states. *Heliyon*. (2022) 8(2):e08882. doi: 10.1016/j.heliyon.2022.e08882

38. Kumar VV, Tankha G. Association between the big five and trait emotional intelligence among college students. *Psychol Res Behav Manag*. (2023) 16:915–25. doi: 10.2147/PRBM.S400058

39. The Royal College of Nursing. RCN Nursing in Numbers 2023|Publications|Royal College of Nursing. (n.d.). Available online at: https://www.rcn.org.uk/Professional-Development/publications/rcn-nursing-in-numbers-english-uk-pub-011-188 (accessed June 2, 2023).

40. Mano LY, Mazzo A, Neto JR, Meska MH, Giancristofaro GT, Ueyama J, et al. Using emotion recognition to assess simulation-based learning. *Nurse Educ Pract*. (2019) 36:13–9. doi: 10.1016/j.nepr.2019.02.017

41. Abajas-Bustillo R, Amo-Setién F, Aparicio M, Ruiz-Pellón N, Fernández-Peña R, Silio-García T, et al. Using high-fidelity simulation to introduce communication skills about end-of-life to novice nursing students. *Healthcare*. (2020) 8(3):238. doi: 10.3390/healthcare8030238

10.3389/fdgth.2024.1307817

Appendix A: Questions from the semi-structured interviews

PART 1: Before completing the scenarios in the online environment
1. Please tell me a bit about yourself, what year you are in your training, and your main areas of interest?
2. What particular challenges do you face in your training when working with patients?
3. Do you feel that online training is something that could help you improve working with patients?
4. Have you done any online nursing before and, if yes, how did you find it?
5. Have you done any training with real actors before and, if yes, do you prefer training with real actors or virtual patients?
PART 2: After completing the scenarios in the online learning environment
1. How did you find them easy or difficult to go through? Please explain.
i. Did these feel realistic:
a. 3D environment. Please explain. b. Scenario/Questions. Please explain
2. Did reporting your emotions during the scenario help you understand better how these emotions can affect the patient? Please explain.
3. Did reporting your emotions during the scenario help you in thinking about and, subsequently, trying to control your emotions? Please explain.
1. Did the changes in the patient's mood in scenario 2 affect your feelings during the scenario? Please explain.
2. Did the changes in the patient's mood in scenario 2 make it clearer how your replies affected her? Please explain.
3. Did you find any of the two scenarios to:
a. Be more realistic. Please explain. b. Improve your learning more. Please explain. c. Make you more empathetic towards the patient. Please explain. d. Work better for you in any other way. Please explain.
4. What changes would you suggest we make to future online training scenarios?
5. Do you feel that more similar training could help you improve your nursing skills further? Would it help you become even more compassionate with patients? Please explain.
6. Would you be interested in taking part in more online training scenarios in the future?

Appendix B

Themes table

Themes related to the research objectives

Theme Number	Code	Theme
01	REAL-EXP	Realistic experience (RO1)
02	BET-RESP	Better responses (RO2)
03	IMP-LEXP	Improved Learning Experience (RO3)
04	INC-EMP	Increased Empathy (RO4)

Themes related to the Participant 1 (P1)

# of citations	Code	Theme
4	REAL-EXP	Realistic experience (RO1)
4	BET-RESP	Better responses (RO2)
4	IMP-LEXP	Improved Learning Experience (RO3)
6	INC-EMP	Increased Empathy (RO4)

Themes related to the Participant 2 (P2)

# of citations	Code	Theme
4	REAL-EXP	Realistic experience (RO1)
3	BET-RESP	Better responses (RO2)
3	IMP-LEXP	Improved Learning Experience (RO3)
5	INC-EMP	Increased Empathy (RO4)

Themes related to the Participant 3 (P3)

# of citations	Code	Theme
2	REAL-EXP	Realistic experience (RO1)
3	BET-RESP	Better responses (RO2)
3	IMP-LEXP	Improved Learning Experience (RO3)
4	INC-EMP	Increased Empathy (RO4)

Themes related to the Participant 4 (P4)

# of citations	Code	Theme
2	REAL-EXP	Realistic experience (RO1)
2	BET-RESP	Better responses (RO2)
2	IMP-LEXP	Improved Learning Experience (RO3)
2	INC-EMP	Increased Empathy (RO4)

Themes related to the Participant 5 (P5)

# of citations	Code	Theme
1	REAL-EXP	Realistic experience (RO1)
1	BET-RESP	Better responses (RO2)
2	IMP-LEXP	Improved Learning Experience (RO3)
1	INC-EMP	Increased Empathy (RO4)

Themes related to the Participant 6 (P6)

# of citations	Code	Theme
2	REAL-EXP	Realistic experience (RO1)
3	BET-RESP	Better responses (RO2)
1	IMP-LEXP	Improved Learning Experience (RO3)
1	INC-EMP	Increased Empathy (RO4)

Themes related to the Participant 7 (P7)

# of citations	Code	Theme
7	REAL-EXP	Realistic experience (RO1)
2	BET-RESP	Better responses (RO2)
1	IMP-LEXP	Improved Learning Experience (RO3)
2	INC-EMP	Increased Empathy (RO4)

Themes related to the Participant 8 (P8)

# of citations	Code	Theme
3	REAL-EXP	Realistic experience (RO1)
2	BET-RESP	Better responses (RO2)
3	IMP-LEXP	Improved Learning Experience (RO3)
4	INC-EMP	Increased Empathy (RO4)

Themes related to the Participant 9 (P9)

# of citations	Code	Theme
1	REAL-EXP	Realistic experience (RO1)
1	BET-RESP	Better responses (RO2)
4	IMP-LEXP	Improved Learning Experience (RO3)
3	INC-EMP	Increased Empathy (RO4)

frontiers | Frontiers in Digital Health

TYPE Original Research
PUBLISHED 22 April 2024
DOI 10.3389/fdgth.2024.1334058

OPEN ACCESS

EDITED BY
Edith Talina Luhanga,
Carnegie Mellon University Africa, Rwanda

REVIEWED BY
Victoria Tiase,
The University of Utah, United States
Elena Vlahu-Gjorgievska,
University of Wollongong, Australia

*CORRESPONDENCE
Yu-Peng Chen
✉ yupengchen@ufl.edu

RECEIVED 06 November 2023
ACCEPTED 05 April 2024
PUBLISHED 22 April 2024

CITATION
Chen Y-P, Woodward J, Shankar MN, Bista D, Ugwoaba U, Brockmann A, Ross KM, Ruiz J and Anthony L (2024) MyTrack+: Human-centered design of an mHealth app to support long-term weight loss maintenance. Front. Digit. Health 6:1334058. doi: 10.3389/fdgth.2024.1334058

MyTrack+: Human-centered design of an mHealth app to support long-term weight loss maintenance

Yu-Peng Chen[1]*, Julia Woodward[2], Meena N. Shankar[3], Dinank Bista[1], Umelo Ugwoaba[3], Andrea Brockmann[3], Kathryn M. Ross[3], Jaime Ruiz[1] and Lisa Anthony[1]

[1]Department of Computer and Information Science and Engineering, University of Florida, Gainesville, FL, United States, [2]Department of Computer Science and Engineering, University of South Florida, Tampa, FL, United States, [3]Department of Clinical and Health Psychology, University of Florida, Gainesville, FL, United States

A growing body of research has focused on the utility of adaptive intervention models for promoting long-term weight loss maintenance; however, evaluation of these interventions often requires customized smartphone applications. Building such an app from scratch can be resource-intensive. To support a novel clinical trial of an adaptive intervention for weight loss maintenance, we developed a companion app, MyTrack+, to pair with a main commercial app, FatSecret (FS), leveraging a user-centered design process for rapid prototyping and reducing software engineering efforts. MyTrack+ seamlessly integrates data from FS and the BodyTrace smart scale, enabling participants to log and self-monitor their health data, while also incorporating customized questionnaires and timestamps to enhance data collection for the trial. We iteratively refined the app by first developing initial mockups and incorporating feedback from a usability study with 17 university students. We further improved the app based on an in-the-wild pilot study with 33 participants in the target population, emphasizing acceptance, simplicity, customization options, and dual app usage. Our work highlights the potential of using an iterative human-centered design process to build a companion app that complements a commercial app for rapid prototyping, reducing costs, and enabling efficient research progress.

KEYWORDS

mHealth apps, human-centered design, weight management, adaptive interventions, self-monitoring, feedback, visualization, behavior change

1 Introduction

In recent years, there has been a growing interest in utilizing mobile health (mHealth) applications (apps) to support and enhance various aspects of healthcare (1). Obesity remains a substantial public health challenge in the United States (2), and extended care programs have proven effective in supporting long-term weight loss maintenance (3, 4). In the weight management arena, researchers have been using smartphone apps that allow individuals to track weight and weight-related behaviors (e.g., dietary intake and physical activity) to investigate novel adaptive intervention models, such that intervention may be "triggered" by different patterns in individual behavior (5–7). To provide extended-care support at times when individuals are at high risk for weight

regain, our team has designed and developed an adaptive weight maintenance intervention (8). The evaluation of such an intervention often requires development of an instrumented app with customized functionalities. However, building an app from scratch can be resource-intensive, especially when the main focus is on intervention development and outcomes assessment rather than comprehensive app development.

An alternative approach is to develop a companion app that harnesses the capabilities of the main commercial app through its open-source API. The commercially popular weight management app FatSecret (9) offers an API (10), logging interfaces, and comprehensive databases for tracking weight-related data; however, further instrumenting was needed to meet our specific goals, such as sending customized questionnaires. Thus, to support the implementation of our intervention and its evaluation in a randomized controlled clinical trial (8), we designed and developed a companion app, which we called the MyTrack+ app, to pair with the FatSecret app. Our main goal was to create a seamless logging experience for participants, ensuring smooth data collection. MyTrack+ integrates data from FatSecret and a BodyTrace smart scale (11), enabling participants to log and self-monitor their health data, while also incorporating customized questionnaires and timestamps to support our research aims. These ecological momentary assessment (12) questionnaires allow us to collect data necessary for implementing our intervention and enabling future exploratory studies. As a supplementary objective, we also integrated evidence-based behavior change techniques aimed at increasing participants' motivation, self-efficacy, and app engagement, thereby enhancing adherence to their health objectives.

To ensure the effectiveness and user-friendliness of MyTrack+, we followed an iterative, user-centered approach. Initially, we developed app mockups and created a high-fidelity prototype based on expert evaluations. Subsequently, we conducted a usability study with 17 university students and refined our app based on user feedback and insights from our health experts. Finally, we conducted an in-the-wild pilot study with 33 participants from the target population: adults from the general public who had reported recent weight loss and who were interested in weight loss maintenance. This pilot study provided valuable findings related to acceptability and usability, and we further refined our app based on these findings. This approach facilitated rapid prototyping, iterative testing, and refinement of the app, while incorporating user feedback and the expertise of health experts in weight loss maintenance. Currently, our app has been deployed in an ongoing clinical trial evaluating an adaptive intervention for supporting long-term weight loss maintenance (8).

Our work contributes to the literature by presenting an example of how user-centered design can benefit mHealth research. We document the iterative design process of a tailored mHealth app, specifically designed for research purposes, and its seamless integration with a commercial app. We discuss the implications of our design process, highlighting the potential benefits of leveraging commercial apps for rapid prototyping, thereby reducing implementation costs and enabling researchers to make efficient progress in their investigations for similar projects.

2 Related work

We present relevant prior work in the areas of (1) the design and use of mHealth apps for health behavior change, and (2) the use of user-centered design processes in mHealth apps in general.

2.1 mHealth apps for health behavior change

Smartphone app-based interventions have gained significant attention in recent years (1). Researchers (13, 14) have identified the benefit of various behavior change techniques, such as goal setting, self-monitoring, feedback, and social support, for improving dietary intake and physical activity within app-based interventions. Zhao et al. (13) conducted a literature review of 23 articles that focused on the use of mobile phone apps to promote health behavior changes in peer-reviewed journals. The authors found that out of the reviewed studies, 17 of them reported statistically significant outcomes indicating a positive influence on the desired behavior change. Notably, self-monitoring emerged as the most frequently utilized behavior change technique, being employed in 12 of the studies. Dounavi and Tsoumani (14) also conducted a systematic literature review with the specific goal of identifying the existing evidence on the effectiveness of mobile health technology in promoting weight management behaviors, such as physical activity and healthy eating. Out of the 39 analyzed studies, the authors found that high levels of engagement with a mobile health app led to satisfactory treatment adherence, resulting in successful weight loss and maintenance.

Nevertheless, there are systematic literature reviews that find only modest evidence in support of the effectiveness of mobile apps in improving health behaviors or outcomes (15–17), in contrast to the positive evidence mentioned earlier. Based on a review of 27 studies, Schoeppe et al. (15) stated that multi-component interventions seem to be more effective than standalone app interventions and emphasized the need for further confirmation through controlled trials. Furthermore, current mHealth apps often provide extensive functionalities and complex interfaces, which may not necessarily contribute to their effectiveness (18–20). Lyzwinski et al. (18) conducted a literature review of qualitative studies focused on user perspectives and experiences with mHealth for weight loss. From their review of 20 articles, the authors identified that the most preferred apps were those that were simple and easy to use. Haggag et al. (19) performed an extensive analysis of mHealth app user reviews by extracting and translating over 5 million user reviews for 278 mHealth apps. The authors' findings revealed that providing users with more information or functionalities than necessary can result in user frustration and reduced app usage.

This background presents an opportunity for developing a minimalist app that only focuses on key factors for behavioral change, which we emphasize in our work. Together with trained interventionists and behavioral health experts, we are

using our app in an ongoing clinical trial to assess adaptive interventions aimed at supporting long-term weight loss maintenance.

2.2 User-centered design for mHealth apps

Previous research has emphasized the importance of *user-centered design*[1] processes in developing effective mHealth apps for various health domains, such as self-management of chronic conditions (21, 22), mental health (23), persons living with HIV (24, 25), women in substance use recovery (26), and participants with fall risk (27). Schnall et al. (24) conducted formative research, which included focus groups, participatory design sessions, and usability evaluations, to guide the development of a health management app for individuals living with HIV. Their review of 15 existing apps meeting their inclusion criteria revealed that none of them integrated all the functionalities identified during their formative work, pointing to a significant lack of clinically-backed design choices in current mHealth apps. The authors further utilized a user-centered model to iteratively develop and refine mock-ups, creating an mHealth app tailored for persons living with HIV (25). Their findings demonstrated that a user-centered approach offered a deeper understanding of their target users' specific requirements and facilitated the creation of an mHealth app that better aligned with user needs. Additionally, Eaves et al. (26) applied user-centered design in developing an mHealth app to support women in substance use recovery. Through an iterative design process, the authors showed that users' feedback helped tailor an mHealth app to maximize usability, access, and safety for this at-risk population. Lastly, Hsieh et al. (27) aimed to provide personalized fall risk screening for clinical populations, including older adults, individuals with Multiple Sclerosis, and wheeled-device users, by utilizing mHealth apps. The authors developed the interface of each app with a user-centered design approach through iterative usability testing and semi-structured interviews. The authors then tested their apps in real-world settings and demonstrated the effectiveness of their apps in measuring fall risk (comparable to clinical assessments) to enhance user safety.

To develop our minimalist weight management app, we employed a user-centered iterative design process to understand users' mental models, focusing solely on essential features. This approach enabled efficient progress in our research investigation and simpler data logging to support our research goals.

3 Designing MyTrack+: goals and features

We present the initial set of design goals and app features, brainstormed in concert with our behavioral health expert team members to ensure the MyTrack+ app would meet all the goals of the broader clinical trial.

3.1 Design goals

Our main goal was to support the implementation and evaluation of an adaptive weight maintenance intervention. Prior work had established that key data points required for such an intervention include participants' self-weighing frequency, self-monitored dietary intake and physical activity, and self-rated measures of weight-related variables (e.g., hunger) (28). Our clinical trial (8) serves as a testbed for an automated "trigger" algorithm which can detect when participants may be at risk of relapse based on their logging or other lapses (29). Thus, our overarching goal was to facilitate data collection during the clinical trial by creating a seamless and effortless logging experience for the participants. Additionally, we aimed to employ behavior change techniques to support participants in achieving their health objectives. By integrating evidence-based strategies, we aimed to enhance motivation, engagement, and self-efficacy, supporting sustainable behavior change and promoting successful long-term weight loss maintenance. Informed by guidance provided by health experts in long-term weight maintenance and prior work on applying behavior change techniques for health behavior change (30), we formulated the following design goals for our study.

3.1.1 Goal 1 (G1): enhance usability for effortless logging

We aimed to minimize the effort required for participants to log their data by designing an intuitive and user-friendly interface that is easily accessible and navigable. By reducing the effort needed to log data, participants are more likely to engage with the app consistently (14). When the interface is intuitive and user-friendly, it enhances the overall user experience and encourages active participation (14).

3.1.2 Goal 2 (G2): deliver necessary instruments

We aimed to deliver research-oriented questionnaires to collect data necessary for implementing the adaptive weight intervention and enabling further exploration. The ubiquitous nature of mobile devices enables low-effort self-reporting in real-time through ecological momentary assessment (EMA) (31). In mHealth apps, EMA measures (e.g., questionnaires assessing an individual's thoughts, feelings, and behaviors and the context in which these occur) are usually delivered repeatedly over time, in the natural environment (31). In our case, questionnaires prompt participants to self-rate, on 7-point Likert scales, factors that had been hypothesized previously to be associated with weight regain (28).

3.1.3 Goal 3 (G3): facilitate self-monitoring and provide feedback

In addition to collecting essential research data, the previous two goals serve the purpose of facilitating self-monitoring and self-reflection. We further aimed to deliver personalized health

[1]Also referred to as *human-centered design*.

information to participants in the form of summary graphs, providing tailored feedback on their progress. According to Bandura's social cognitive theory (SCT) (30), self-monitoring increases the user's awareness of their progress and feedback provides an opportunity for them to adjust their strategy (30). Self-monitoring refers to a person's action of keeping a record of details related to performance of the target behavior (e.g., logging the duration and intensity of performing certain physical activities). Such action is a part of the self-regulatory mechanism that is required for beneficial behaviors to be achieved and maintained (30). Moreover, for goals to be effective, summary feedback on self-reported details is also crucial. Such feedback provides an opportunity for individuals to adjust the level or direction of their effort or to adjust their strategies to match what the goal requires (30). Prior work showed that a system that facilitates the user's self-monitoring and provides feedback can effectively increase physical activity (32) and motivate healthy eating (33).

3.1.4 Goal 4 (G4): increase motivation and self-efficacy

Our secondary objective was to harness the power of behavior change techniques inspired by the principles of goal-setting theory to increase participants' motivation and self-efficacy. Locke and Latham's goal-setting theory (GST) (34) states that having a goal is a crucial cognitive determinant of human behavior and performance. According to the GST (34), there is a positive, linear relationship between goal difficulty and levels of effort and performance. The authors found that effective goals should be challenging enough to induce effort for an individual to be motivated but should not be so difficult that they cause repeated failures. Goal setting has been shown to be an important factor in supporting behavior change in various health fields (35, 36). Furthermore, Bandura's social cognitive theory (SCT) (30) holds that self-efficacy is a major determining factor for behavior change. Self-efficacy is an individual's belief in their ability to execute a certain behavior in a given situation. Repeated successes play a significant role in boosting self-efficacy, as past achievements have a notable influence on self-perception (30). Therefore, the keys to accomplishing long-term health behavior changes involve setting intermediate goals and making persistent efforts towards their achievement. Individuals can enhance their self-efficacy by actively monitoring their own behavior and receiving feedback that highlights progress towards goal attainment (34). Previous studies have focused on increasing the user's self-efficacy to motivate health behavior change (37, 38).

3.1.5 Goal 5 (G5): promote app engagement and adherence to health objectives

To achieve this goal, rooted in behavior change theories (30, 39), we aimed to encourage app engagement and adherence to health objectives by incorporating notifications and expert support into our system. The Fogg Behavior Model (39) emphasizes the importance of providing triggers to increase app engagement and adherence. Prior work has also pointed out that mHealth app notifications can aid in behavioral change through increasing user app engagement and adherence to health objectives (40). Additionally, the Supportive Accountability model proposed by Mohr et al. (41) argues that human support increases adherence to electronic health interventions through accountability to a human coach. Social persuasion provided by a human coach can also enhance an individual's self-efficacy, thereby increasing motivation (30). Being encouraged to perform certain behavior by others through suggestions can enhance an individual's belief in their capability of successfully executing such behavior (42). Previous studies found that users consider consultation and communication with health experts to be a crucial functionality in mHealth apps (24, 43).

3.2 Initial design features

To achieve our design goals, we considered the following initial design features in MyTrack+.

3.2.1 Interfaces for logging weight-related behaviors

While this section primarily focuses on designing MyTrack+, to achieve **G1** and also streamline the software engineering process, we utilized the existing features in a well-established commercial app, FatSecret, and a BodyTrace smart scale (Table 1). FatSecret provides an interface for logging consumed food and drinks, including portion sizes and nutritional information such as calories and macronutrient/micronutrient composition. It also includes features that enhance its usability, such as barcode scanning for food items and the "Recently Eaten" and "Most Eaten" options, reducing logging efforts. Users can also log exercises with duration and access a comprehensive database of physical activities. Additionally, our system incorporates the use of a BodyTrace smart scale that utilizes cellular network connectivity to transmit participant weight data to BodyTrace servers. This eliminates the need for participants to manually log their weight data, as developers can directly request the information from the servers.

TABLE 1 Our mHealth system incorporates three main components: (1) our MyTrack+ app, (2) the FatSecret commercial app, and (3) a BodyTrace smart scale.

Application/ Device	Data Type	Features
MyTrack+ app	Questionnaire	Summary graphs, Notifications, Goal setting, Health expert support
FatSecret app	Dietary intake, Physical activity	Self-logging user interface, Databases of nutritional and exercise information
BodyTrace smart scale	Weight	Automatic weight logging

This system was designed to facilitate the self-monitoring of health data relevant to weight management, including weight, dietary intake, physical activity, and weight-related factors (via questionnaires).

3.2.2 Navigation

Achieving effortless logging (**G1**) requires an intuitive app with easy navigation. We focused on facilitating both cross-app navigation (from MyTrack+ to FatSecret) and within-app navigation (between different pages within our app).

3.2.3 Research-oriented questionnaires

To achieve **G2**, we utilized MyTrack+ to conduct EMA through weekly and end-of-week check-in questionnaires. Questionnaire content was developed by our team of health experts, while the software engineering team dedicated their efforts to designing the user interface. This feature also supports the achievement of **G3** by allowing users to assess their progress through rating various factors that were previously hypothesized to be linked to weight regain (28).

3.2.4 Notifications

To achieve **G2**, we implemented notifications that appear both outside the app (i.e., push notifications) and inside the app (i.e., in-app notifications). These notifications serve as reminders for users to reflect on their weekly progress while also accomplishing **G5**. A survey found that designers use notifications as a way to facilitate behavioral change by increasing user engagement with the app and promoting adherence to health objectives (44).

3.2.5 Overall summary graphs

To achieve **G3**, we included summaries for calories remaining, physical activity, and weight in the form of summary graphs, as providing visual feedback on self-reported data is an effective strategy for boosting physical activity levels and promoting healthier eating habits (14). These graphs also aid in accomplishing **G4**, as they enhance self-efficacy by empowering the user to visualize and review their past achievements.

3.2.6 Goal setting and monitoring

To achieve **G4**, we implemented features that enable users to set their goals. We also implemented features that enable the user to track their progress towards intermediate sub-goals, contributing to both **G3** in providing feedback on self-reported data and **G4** by promoting self-efficacy through the accomplishment of repeated small successes.

3.2.7 Expert support

To achieve **G5**, we created a support page that facilitates communication between users and health experts, as accountability to a human coach has been demonstrated to enhance adherence (41). This feature also contributes to **G4** by promoting self-efficacy through social encouragement (30).

4 Iterative design of MyTrack+

To achieve our design goals and determine how to support the features we brainstormed, we followed a user-centered approach with iterative evaluation of prototypes and corresponding incremental changes, encompassing multiple stages (Figure 1). The process involved first developing app mockups and designing the backend system. We developed the first version of a high-fidelity prototype based on feedback from health experts in our team and through internal pilot tests with our human-computer interaction (HCI) researchers. Subsequently, we conducted a usability study with 17 university students to gain insights into ease of use, preferences, and technical issues. We refined our app based on feedback received from users and insights provided by our health experts. Finally, we conducted an in-the-wild pilot study involving 33 participants from the general public, and then further refined our app based on feedback from both participants and our health experts.

Our transdisciplinary team included faculty, staff, and graduate students with expertise in obesity treatment, clinical psychology, computer science, and human-computer interaction (HCI). The four-member health expert team comprised of a faculty clinical psychologist with 17 years of experience in obesity treatment, a registered dietitian with 21 years of research experience, and two Clinical and Health Psychology graduate students with training in obesity treatment. The six-member computer science team comprised of two faculty members with a combined total of 25 years of research experience in HCI and 15 years of industrial experience in software engineering, and four graduate students with training in computer science and HCI.

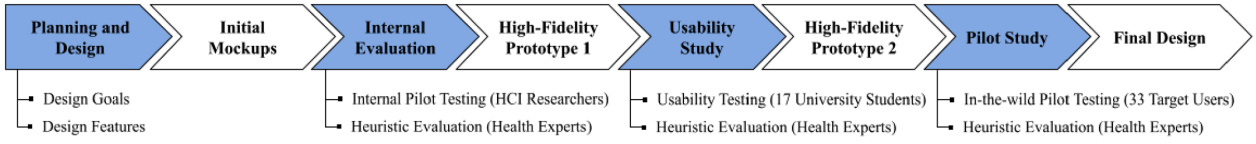


FIGURE 1
Our user-centered design process involving iterative evaluation and incremental refinement of app prototypes. The elements colored in blue indicate stages where we conducted brainstorming sessions, evaluations, and/or studies. Specifically, (1) HCI researchers were recruited for internal evaluation, (2) Computer Science students were recruited for the usability testing, and (3) the target users were recruited for the pilot study. We engaged our health experts for each of these stages to ensure that the incremental changes we made were aligned with recommendations from the weight management literature. The elements colored in white represent the result of the previous element.

4.1 Initial mockups and backend design

We designed the following mockup components based on our design goals and initial features (Figure 2). We provide a detailed explanation of each component in the same order as our initial design features listed in Section 3.2.

4.1.1 Navigation

At the bottom of MyTrack+ (Figure 2-1), we implemented a navigation panel that includes the "Home", "Diary", and "Support" tabs. These tabs allow the user to navigate to the Home screen, FatSecret, and the Support screen from any screen. We implemented the "Diary" tab to facilitate the navigation from MyTrack+ to FatSecret. While we lack control over the implementation of the FatSecret app, users can navigate backward using the built-in back navigation feature on both iOS and Android devices. We also implemented arrows within in-app notifications and summary graphs to visually indicate clickable components.

4.1.2 Questionnaires

We created a Questionnaire screen (Figure 2-2) that displays questions developed by our health experts. We implemented a 7-point Likert scale slider for each question, allowing users to self-rate the weight-related factors. Users can navigate to this screen through in-app notifications (Figure 2-3).

4.1.3 Notifications

We designed notifications based on empirically-derived notification design recommendations in mHealth apps, including position, aesthetics, and content (44). We placed the in-app notification at the top of the Home screen (Figure 2-3) to ensure that it is prominently visible and does not interfere with the main content of the app. We also implemented a transient warning message that appears at the bottom of the user's current screen and disappears after five seconds. This brief message can capture the attention of users who may not be actively looking at the top of the home screen or may be engaged in a different part of the app's interface. We included an arrow to indicate that the notification is clickable and a "warning" icon to indicate the importance of clicking on the notification. To create an effective visual effect, we chose orange as the background color for the in-app notification, leveraging its complementary nature to our main color theme of gray blue.

4.1.4 Summary graphs

Our summary visualization includes graphs for physical activity and weight data (Figure 2-4). Our app displays data summary on a week-to-week basis (Sunday to Saturday), while providing the functionality for the user to monitor tracking history of previous weeks. Regarding the type of weekly summary graphs, we explored line graphs (Figure 2-4-1) and bar charts (Figure 2-4-2). Additionally, for tracking physical activity, we explored different units, including total duration in minutes or total burned calories.

4.1.5 Dietary intake and goal indicators

Our app displays summary graphs for dietary intake (Figure 2-5). To display daily calorie goal and the weekly calorie intake, we considered two options: (1) display both graphs on the Home screen (Figure 2-5-1) or (2) display the daily calorie goal on the Home screen and link it to a separate screen for daily breakdown and weekly summary (Figure 2-5-2). For the daily calorie goal, we considered a balance equation, a circular progress bar, and a pie chart. For visualizing the detailed breakdown of daily and weekly calorie intake, we considered a bar chart and a pie chart, with text information displayed below and alongside. Additionally, we designed goal-monitoring features, including textual information in a daily summary indicating the calorie goal, total calories consumed, and remaining calories. In the weekly graph, a line representing the calorie goal was incorporated.

4.1.6 Support screen

On the Support screen (Figure 2-6), the app presents the message: "Fill out the form below and a specialist will contact you." We introduced a "specialist" character to indicate real human support, which has been shown to be effective in increasing engagement (40). Since our support is provided through phone, we designed open text input for the user to input their phone number and indicate their preferred call time.

4.1.7 Backend design

To ensure a seamless data collection process and smooth dual app usage, the MyTrack+ app focuses on integrating data from multiple sources: FatSecret and BodyTrace. Specifically, we implemented three backend components for this feature. We integrated the Firestore Database (45) into our system for data storage. We built Google Cloud Functions (46) for synchronization between our app and FatSecret. Google Cloud runs scheduled Cloud Functions to request data from FatSecret's server using the API they provide. We provided a "Diary" tab in the navigation panel at the bottom of the Home screen. This allows the user to navigate from MyTrack+ to FatSecret from any screen. In addition, the FatSecret API does not include timestamps for self-monitored weight, caloric intake, or physical activity. However, the inclusion of timestamps is crucial for supporting the implementation and evaluation of our adaptive weight maintenance intervention, as well as for conducting further exploratory study, such as developing novel models to proximally predict weight change and supporting future adaptive intervention development. Thus, in our backend system, we appended timestamps to the data collected in FatSecret and BodyTrace when requesting data from those applications' servers.

4.2 High-fidelity prototype 1

In our iterative process of improving the app to fulfill our design goals, we first sought guidance from our team of health experts and performed internal pilot tests with our team of HCI researchers. Based on feedback from these experts and researchers, we developed the initial high-fidelity prototype (Figure 3). We

10.3389/fdgth.2024.1334058

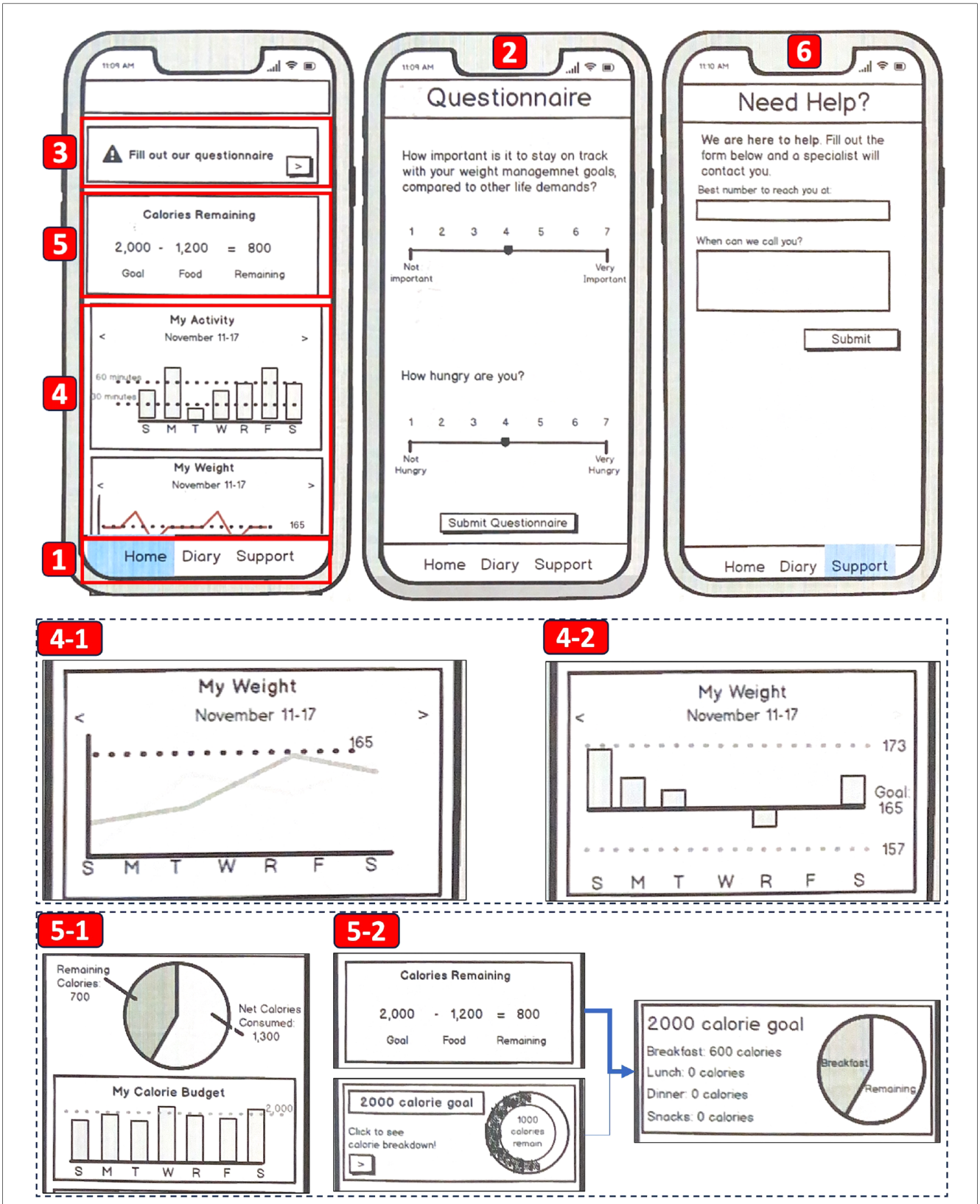


FIGURE 2
The app mockup components. The labels correspond to the descriptions provided in Section 4.1: (1) The navigation panel. (2) The Questionnaire screen. (3) The in-app notification. (4) The summary graphs for physical activity and weight data with two design options: (4-1) a line graph and (4-2) a bar chart. (5) The calorie intake summary with two design options: (5-1) display both daily and weekly summaries on the Home screen and (5-2) display the daily summary on the Home screen and link it to another screen for weekly summary. (6) The Support screen.

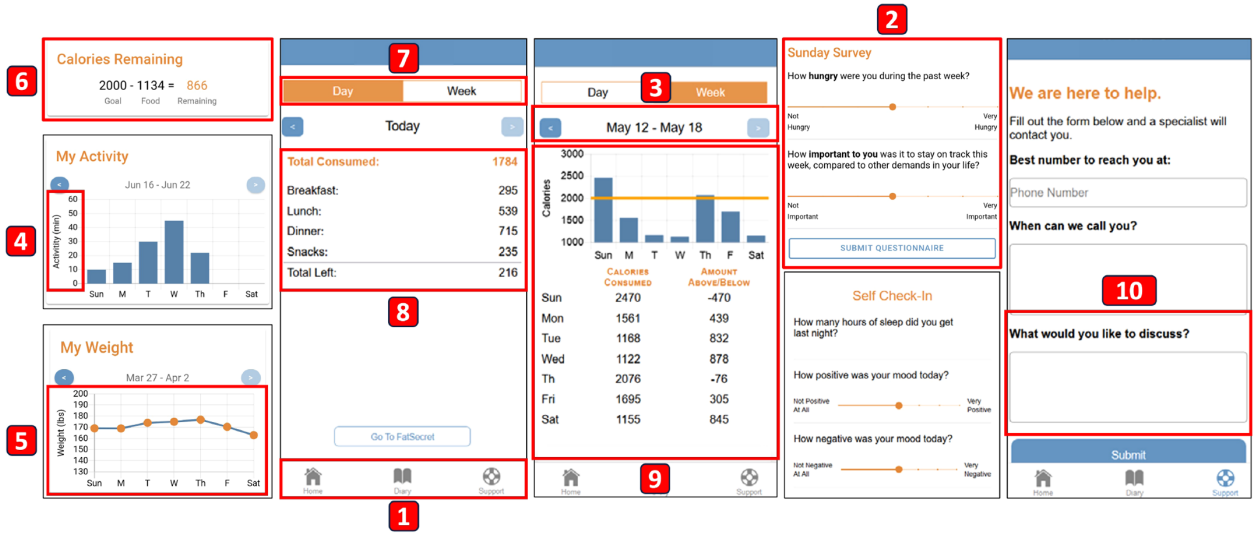


FIGURE 3
The first high-fidelity prototype. The leftmost screenshot displays the Home screen, with the components arranged from top to bottom following the same order as in the actual app. The labels' order follows our description in Section 4.2: (1) Navigation panel with icons. (2) End-of-week questionnaire. (3) Shading of the arrows in summary graphs. (4) Physical activity summary graph. (5) Weight summary graph. (6) Daily calorie goal. (7) Tabs for selecting types of details. (8) Daily dietary intake details. (9) Weekly dietary intake details. (10) Open text input for topics on the Support screen.

detail our iterative improvement process, highlighting our design goals.

To facilitate effortless logging (**G1**) through intuitive navigation and user-friendly interfaces, we refined navigation and questionnaire components, following suggestions from our HCI researchers who drew upon previous studies in the mHealth field (14). We integrated icons into the tabs of the navigation menu, providing visual cues for easy navigation (Figure 3-1). Additionally, we highlighted the keywords within the questions in the questionnaires to draw attention (Figure 3-2), aiming for **G2**. For example, we emphasized the "hungry" keyword using bold text in the question that prompts the participant to rate their hunger: "How **hungry** were you during the past week?" Lastly, we introduced shading to the arrows in summary graphs (Figure 3-3) to indicate whether the user can navigate and review data from past weeks (**G3**).

To enhance self-awareness (**G3**), boost motivation (**G4**), and increase engagement (**G5**), we incorporated feedback from our health experts to refine components such as summary graphs, dietary intake detail, and the Support screen. For the physical activity summary graph (Figure 3-4), we used minutes instead of "calorie burned" as the unit because our intervention goals are set in minutes, based on guidelines issued by both the American College of Sports Medicine (ACSM) (47) and the Centers for Disease Control and Prevention (CDC) (48). For the weight summary graph (Figure 3-5), we selected the line graph because it effectively shows the change of weight over time. For dietary intake summary (Figure 3-6), our app displays only the daily calorie goal summary on the Home screen to avoid redundancy. We opted for a balance equation format because it was perceived as more easily understandable and preferable at a glance, while a circular progress bar received unfavorable feedback. Clicking on the balance equation redirects the user to the dietary intake detail screen, which features two tabs at the top (Figure 3-7). For daily breakdown, based on feedback from our health experts, the pie chart was deemed confusing. Thus, we opted for a vertical subtraction expression to enhance clarity (Figure 3-8). This design presents the total calories consumed for the day at the top, followed by a breakdown of calories consumed by meals (e.g., breakfast, lunch, dinner, snacks) listed below, and concludes with the remaining calories at the bottom. For the weekly dietary intake summary, we opted for a bar chart representation, as it effectively emphasizes the line indicating the calorie goal (Figure 3-9). Supplementary numerical values are included below the bar chart, displaying the calories consumed and the deviation from the goal (i.e., daily calorie goal - calorie consumed) for each day of the week. Lastly, our health experts provided feedback indicating that incorporating an open text input for users to express the topics they would like to discuss is beneficial for both the user and the health professional (Figure 3-10).

4.3 Usability study

To iteratively assess and enhance the high-fidelity prototype of the MyTrack+ app, we performed a lab usability study involving 17 students majoring in Computer Science from our local university. Our goal is to evaluate MyTrack+'s general usability and identify basic bugs and major usability flaws. Although Computer Science students were not our main target users, according to Nielsen (49), involving students in the domain of interest from a local university in usability testing can still yield valuable feedback. We chose to recruit students as convenience samples (50) because (1)

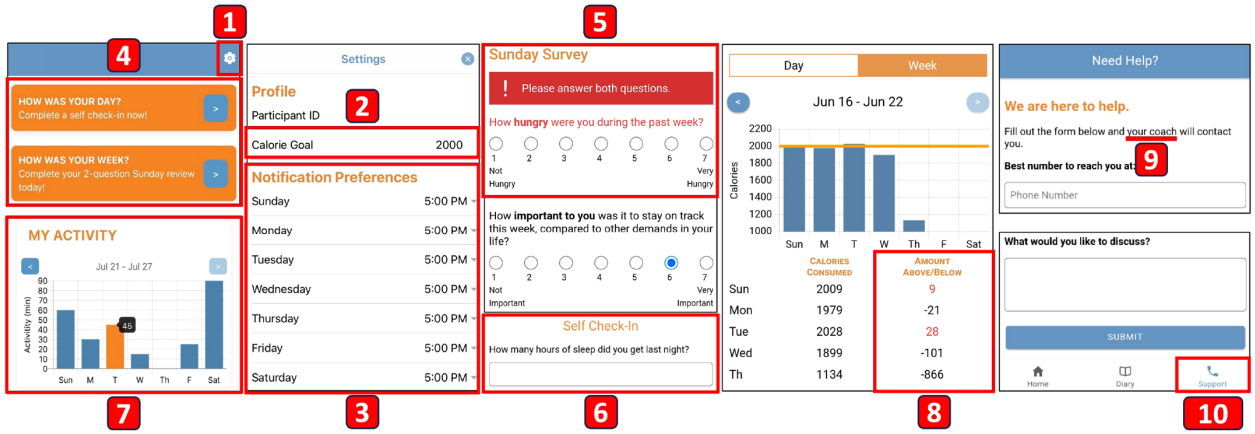


FIGURE 4
New features or changes in the second high-fidelity prototype: (1) The "gear" icon for navigating to Settings. (2) An input space where the user can set and change their calorie goal. (3) Notification preferences. (4) In-app notifications. (5) Unanswered question on the Questionnaire screen and radio buttons. (6) An open-response text box for the hours-of-sleep question. (7) Displaying numerical values when bars are clicked. (8) The values are calculated based on "consumed − goal". (9) The term "your coach" for tailored personification. (10) The "phone" icon for the Support tab.

they are easier to reach and more readily available and (2) they could also be potential users of weight management apps. We listed our study in the Computer Science department's Research Participation System to recruit students enrolled in Computer Science classes. We excluded one participant due to missing data. The mean (SD) age of the 16 participants was 21.7 (3.3) years and 3 participants (19.8%) were women.

To address potential limitations in our lab study, which may not reflect real-world stress, and because we recruited university students who may not have a weight maintenance goal after weight loss, we used four task scenarios in the usability study to evaluate our prototype:

- **Scenario 1** (**G1**, **G3**, and **G4**). We explained the context of use for this app being to help people manage their weight after weight loss. Our first scenario asked the participant to explore the Home screen with summary information and the dietary intake detail screen with this in mind.
- **Scenario 2** (**G5**). We described a scenario where the user was having trouble staying on track with their weight management program and would like someone to contact them in order to talk about it.
- **Scenario 3** (**G2**). We described a scenario where the user was experiencing a real-world stress (e.g., promotion at work) that demanded extra hours and attention. Participants were asked to fill out the end-of-week questionnaire, which is a part of their weight management program every Sunday. The end-of-week questionnaire includes two questions prompting the participant to rate two weight-related factors, which are necessary for our adaptive intervention.
- **Scenario 4** (**G2**). We then asked the participant to fill out the weekly check-in questionnaire, under the same circumstance (i.e., experiencing a real-world stress that demanded extra hours and attention). The weekly check-in questionnaire includes 12 questions prompting the participant to rate 12 weight-related factors, which are also necessary for our adaptive intervention. In our clinical trial, this questionnaire is sent on a random day between Monday and Saturday (with participants unaware of the day that the questionnaire will be asked).

We showed the participants the MyTrack+ app and read the description of our scenarios. They were then asked to perform tasks in our app and answer questions about the usability of the app while thinking aloud (51). They then completed a System Usability Scale (SUS) questionnaire (52) and received extra course credit as compensation. Our study protocol was approved by our university's Institutional Review Board (IRB).

4.4 High-fidelity prototype 2

The average SUS score was 84.53 (min = 60; max = 100; SD = 10.92), surpassing the widely accepted SUS score benchmark of 68 (53). Based on feedback from the usability study and our health experts, we further refined our app and developed the second high-fidelity prototype (Figure 4). Throughout this process, our main focus remained on achieving established design goals.

In the usability study, students noted their needs to input or change their calorie goal, mentioning that, "I was not sure how to input my target goal." (P17) and, "It feels like you should be able to click on the section headers and get taken to a configuration screen." (P05). Aligned with our design goal on providing goal-setting functionality to increase motivation (**G4**), we implemented a "gear" icon (Figure 4-1) on the Home screen and linked it to the Settings screen (Figure 4-2). Additionally, researchers have increasingly emphasized the importance of flexibility and customizability in mHealth apps to support user autonomy (54). Previous studies have found that customizable apps can increase users' engagement by providing a sense of control (55). Together with feedback from our health experts, we

implemented options for the user to specify their notification preferences (Figure 4-3), effectively achieving **G2** and **G5**. Based on the specified times, in-app notifications appear at the top of Home screen (Figure 4-4). The notification content was decided based on discussions among our HCI researchers and health experts, drawing from our expertise and prior work. Our goal was to deliver a motivational message ("How was your day/week?") with a sense of urgency in responding to the questionnaire, employing an assertive tone emphasized by an exclamation mark ("Complete a self check-in now!"), following the established recommendation (40).

Clicking on the in-app notifications redirects the user to a separate Questionnaire screen. Since the questions are critical for our research purposes, our app displays a warning if there is more than one unanswered question and changes the color of the unanswered question to red (Figure 4-5). Based on feedback from the students, mentioning that, "Sometimes, when I would try to scroll, my finger would accidentally move the sliders on the responses." (P05), which is consistent with the "fat fingers" issue for touchscreen gestures (56), we replaced the sliders with radio buttons. We also added an open-response text box to the hours-of-sleep question based on feedback from our team of health experts (Figure 4-6).

For summary graphs on the Home screen, students expressed difficulty in reading the chart, mentioning that, "If I was not paying attention I may not read the chart correctly." (P03). Our health experts also provided similar feedback, which prompted us to implement a feature wherein numerical values are displayed when bars or points in the graphs are clicked (Figure 4-7). This feature aims to facilitate self-reflection for increasing self-awareness (**G3**). For the numerical values on the weekly dietary intake screen (Figure 4-8), we changed the direction of subtraction from "goal − consumed" to "consumed − goal" based on feedback from our health experts, stating that negative signs should be used when someone is below their goal. Interestingly, students expressed confusion about this particular change, mentioning that, "At first sight the negative values seemed a little tricky to me. Only because thinking of something in the negative might ignite a feeling of inadequate performance?" (P08). Thus, we implemented visual indicators by marking values that exceeded the calorie goal in red, aiming to strike a balance between user preferences and health expert suggestions.

Lastly, students expressed confusion about the purpose of the Support screen, stating that, "[I'm] confused if the purpose of Support is supposed to connect you with someone to encourage you to reach your goals or if it's general support for the app overall." (P09). Our health experts supported this feedback and recommended replacing the "support buoy" icon with a "phone" icon (Figure 4-10). We also changed the term "specialist" to "your coach" (Figure 4-9), aligning with the established recommendation of incorporating tailored personification (40).

4.5 Pilot study

To iteratively assess and improve the second high-fidelity prototype of the MyTrack+ app, we conducted an in-the-wild pilot study involving 33 target users from our local community who had lost at least 5% of body weight during the past two years. We excluded two participants due to missing data. The mean (SD) age of the 31 participants was 40.1 (15.6) years and 21 participants (67.7%) were women. The following are details of our two-week pilot study procedure (Figure 5).

4.5.1 Recruitment

Participants were recruited using flyers, outreach to local employers/businesses, and newspaper ads; all such materials were approved by our IRB.

4.5.2 Orientation and informed consent process

Participants who met initial eligibility criteria as assessed via the phone screen were invited to attend an in-person orientation visit. The orientation visit included a discussion of the pros/cons of taking part in research and specific details of the current study. Potential participants were given the opportunity to privately ask study staff any remaining questions and then, if they remained interested, were asked to provide written informed consent.

4.5.3 Initial visit

Participants who met study eligibility criteria were provided a study smart scale developed by BodyTrace, Inc. (11). Study team members reviewed the MyTrack+ smartphone application and assisted participants with smartphone setup. Participants were asked to use the study smart scale and MyTrack+ application for the following two weeks before returning for a follow-up feedback visit. Participants were asked to weigh themselves each day, first thing in the morning after using the restroom but before eating/drinking, in nothing more than light indoor clothing (57).

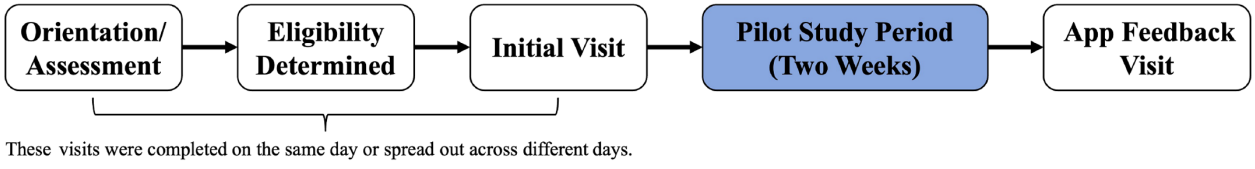


FIGURE 5
The procedure of our two-week, in-the-wild pilot study involving 33 target users.

4.5.4 Pilot study period

During the two-week study period, participants were asked to (1) use study smartphone apps to track weight, caloric intake, and physical activity each day, (2) use smart scale to measure weight each day, and (3) complete questionnaires when prompted via study smartphone app. MyTrack+ pushed two questionnaires to participants each week: one 12-item questionnaire (Weekly MyTrack+ Questionnaire) delivered on a random day each week, ranging from Monday through Saturday and one 2-item end-of-week check-in ("End of Week MyTrack+ Questionnaire").

4.5.5 App feedback visit

Participants were asked to return two weeks after their initial study visit to complete SUS questionnaires and provide feedback regarding the usability/acceptability of the MyTrack+ smartphone application. Study staff measured participant weight, and completed a semi-structured interview to assess participant perception of usability and acceptability of the MyTrack+ application. Audio recordings of the semi-structured interviews were collected to document suggestions for further app development. Participants were provided with a $30 honorarium for completing this study visit.

Our study protocol was approved by our university's IRB.

4.6 Final design

As the last phase of our iterative design process, we integrated the insights from the pilot study and gathered additional input from our health experts. The average SUS score of the second high-fidelity prototype was 70.48 (min = 27.5; max = 97.5; SD = 16.89), which remained higher than the typical SUS score benchmark of 68 (53). Focusing on our design goals, we continued to refine our app and created the final design of the MyTrack+ app (Figure 6), which is now in use in our ongoing clinical trial.

For the weight summary graph (Figure 6-2), our pilot participants expressed the need for more noticeable weight changes to enhance their motivation. Suggestions included providing a zoom-in functionality or reducing the unit range on the y-axis. To address the issue, we prioritized addressing the main range of weight changes while avoiding investing significant engineering effort in handling extreme edge cases. This was based on the understanding that in realistic scenarios, significant weight changes do not typically occur in a short period of time.

For dietary intake details (Figure 6-4), we included tailored feedback based on the total calories consumed, as suggested by our health experts. For example, if total calories consumed are too low below the goal, the page displays the message, "Based on your total calories consumed, you are likely not eating enough." Notably, nearly all participants expressed concerns regarding the delay in data synchronization between MyTrack+ and FatSecret. To address this issue, we modified our Cloud Functions to request data from FatSecret's server whenever the companion app is refreshed, effectively achieving **G1**.

5 Findings

We present four general themes identified throughout our design process.

5.1 Personalized preferences

Design preferences for an mHealth app tend to be highly personalized, varying from one user to another. Some users preferred the simplicity of MyTrack+, finding it effective in

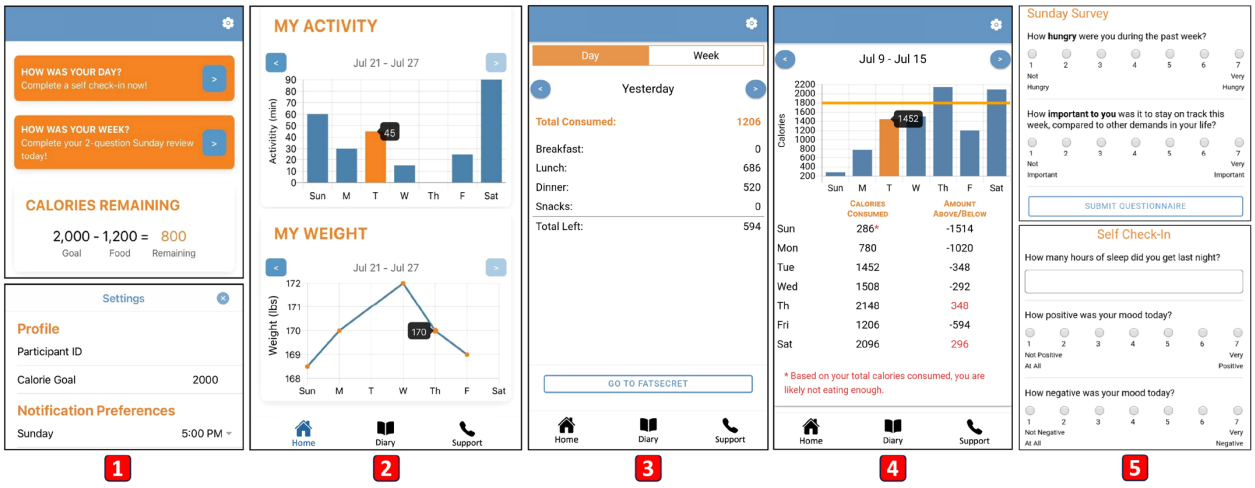


FIGURE 6
Main screens of our final design: (1) The upper part of our Home screen and the Settings screen. (2) The lower part of our Home screen with summary graphs. (3) Daily dietary intake summary. (4) Weekly dietary intake summary. (5) Questionnaires.

presenting relevant information clearly, stating that, "As it [MyTrack+] stands now, while there isn't much on the homepage, it's easy enough to navigate with a couple taps to any of the other information that someone would want to look at. So I think it's simple enough as it is now too jumbled to put too much on the home screen." (P418). Others, however, felt it lacked features and complexity compared to the FatSecret app, stating that, "The goal of [MyTrack+] is to show me that this is how many calories you've eaten during the day. It doesn't appeal to me. FatSecret is going to show me my macro breakdown for the day." (P427). Some users suggested adding more functionality to MyTrack+ based on their previous experiences with health-related technologies, such as providing detailed nutrition information for specific dietary restrictions and incorporating features like BMI tracking and water intake monitoring.

5.2 Effectiveness of behavior change techniques

Users found the process of setting a calorie goal and tracking their progress toward that goal to be highly beneficial. They appreciated visual representation of their performance, as it showed variations in calorie intake on different days, motivating them to be more mindful of their eating habits. Additionally, users emphasized the positive impact of the daily weight and weekly activity tracking features on their self-awareness. They also stated that these features empowered them to stay accountable by remembering whether they had achieved their exercise targets, leading to a positive impact on their adherence to their health goals. Furthermore, users found the in-app notification helpful in reminding them to answer questionnaires, which provided an opportunity for reflecting on health-related factors. Notably, some users extended their usage of the app and integrated it as an assistive tool in their daily lives. For example, one user made connections between their daily weight changes with their menstrual cycle.

5.3 User agency

Many users expressed their desire for control over how data is displayed and their need for flexibility to log for previous days. For example, when asked about the weight graph, one user mentioned that, "The increments are so minuscule that it looks like I have effectively no progress when I know I have." (P457). They then mentioned, "I did try to see if I could adjust the increments, but I didn't see a setting for that." (P457).

5.4 Expectations in apps and connected devices

Nearly all users expressed frustration with delays in synchronization and inaccuracies in data from different sources. A common practice among many users was to immediately verify the accuracy of their recorded weight and ensure that data was correctly synced between MyTrack+ and FatSecret after self-weighing or logging data.

6 Discussion

We discuss the implications derived from user's mental models and reflect on our design process.

6.1 Implications derived from users' mental models

Our studies and iterative design process enabled us to not only gather information about the usability of the app and make design decisions, but also revealed insights into potential target users' mental models of how mHealth apps like these should work. Our observations highlight the individualized nature of user preferences in mHealth app design, emphasizing the fact that there is no one-size-fits-all solution (58). This underscores the importance of adopting a user-centered design approach to understand the specific needs of the target users. We also observed that our app design has effectively enhanced users' motivation, self-efficacy, self-awareness, app engagement, and adherence to health goals. This echoes previous work on effectively applying behavior change techniques in mHealth apps to support users' health goals (13).

A noteworthy observation is that some participants extended the use of certain features, incorporating them as assistive tools in their daily lives. This action of adding flexibility to the use of MyTrack+ indicates the need of user autonomy for mHealth apps. In fact, researchers have emphasized the significance of user autonomy in mHealth technologies (54). Customizable apps have been shown to enhance user motivation and engagement by offering a sense of control (55). For example, systems that allow the user to manually add, edit, or delete personal health data can promote self-awareness and accountability (54). Customizable apps can also allow users to tailor app content to fit personal preferences and goals. Engaging in such customization can enhance user agency and increase the consumption of the customized content (55). As mentioned in our findings, participants expressed their desire for control over how data are displayed and their need for flexibility to log for previous days.

In the health IoT domain, interconnected smart devices form a cohesive system that empowers users to effortlessly achieve their health goals within their daily living environment. In this context, it becomes especially important to preserve human agency by helping users understand and giving them control over their AI-powered devices. This aligns with the primary goal of prioritizing issues of fairness, accountability, transparency, and ethics (FATE) for human-AI interaction (59), specifically highlighting the importance of providing transparency to enhance users' trust in AI-powered health applications. Our system design, including a companion app, a main commercial

app, and a smart scale, resembles interconnected smart devices, offering potential applications in the health IoT field. Users' expectations for a smooth connection between the two apps and instant synchronization of data from different sources highlight their needs for a seamlessly integrated system. More importantly, trust remains a significant concern for users transitioning from traditional devices to smart devices, emphasizing the importance of prioritizing FATE in AI-powered mHealth apps.

While prioritizing users' specific needs and ensuring human agency in AI-powered apps, it is equally crucial to evaluate the cost-effectiveness of additional implementations. Our observation uncovered users' legacy bias (60) from their experiences with other mHealth apps, leading them to expect our research-grade app to incorporate the comprehensive functionality and aesthetic of a commercial app. Not meeting these expectations may impact the app's usability, but trade-offs were necessary to achieve our main design goal: effectively supporting the implementation of our team's specific adaptive weight intervention and prioritizing the progress of our clinical trial aimed at evaluating the intervention's effectiveness. Constantly revisiting the main research purposes of the app and involving domain expertise in the design process are essential to balance this trade-off.

Overall, the implications derived from users' mental models emphasize the significance of understanding target users' specific needs, applying behavior change theories, providing user autonomy, ensuring seamless integration and trust in AI-powered devices, and recognizing users' legacy bias.

6.2 Reflection on our human-centered design process

Prior work (61) has demonstrated the effectiveness of including trained and seasoned experts throughout the design process of an mHealth app. By incorporating domain experts into the design process, designers can tap into specialized knowledge, gain deeper insights into the domain, and create solutions that effectively address user needs while also meeting domain-specific requirements. Our findings highlight the importance of acknowledging that user preferences may not always align with optimal design choices, particularly when target users may not possess the expertise in the specific field that the designer aims to contribute to. In our case, user preferences could have been influenced by their prior experiences with applications lacking clinical and behavioral health motivation. Therefore, while prioritizing the alignment of our design with user needs, we also consulted with our health experts to maintain adherence to the current best practices recommended in the weight maintenance field.

The iterative nature of a human-centered design process facilitated continuous improvement of our app. Through incremental small changes, we targeted specific aspects of the app, allowing us to closely identify and address user needs, gauge the impact of each update, and make necessary refinements. In our early usability study, we worked with students from our local university to gather general usability insights and make iterative improvements before progressing to testing with the target users, who are more costly and difficult to recruit. We aimed to wait until basic bugs and obvious usability flaws were fixed before going to our target user population. Although testing with a target audience is ideal, testing with students in the domain of interest (e.g., Computer Science) can still uncover high-level usability issues that are common across different user groups and provide valuable insights into the general user experience early in the design process (49). Specifically, pilot study participants appreciated the in-app notifications, the ability to access detailed information by clicking on the graphs, and the numerical data highlighting calorie differences.

It is important to note that this convenience sampling approach involves challenges to generalizability. It can be prone to high sampling bias and hence reduce representativeness (50). Specifically, since our major goal is to enable efficient research progress, we adopted Nielsen's concept of discount usability engineering (49) by creating real-world scenarios, instead of conducting an in-the-wild testing. However, a laboratory-based study may still not fully replicate real-world stress. Furthermore, when compared to our target user group, the students we recruited were younger and were not necessarily aiming to maintain their weight after a certain amount of weight loss. These differences in demographics and objectives could potentially introduce biases. As Nielsen (49) pointed out, when testing with students, researchers should consider whether the system is also intended to be used by older users. Also, prior work has indicated that when mHealth services are perceived as personalized, younger consumers tend to be more receptive to adopting them (62). This highlights the importance of continuously reassessing our design goals and critically evaluating whether specific features align with our objectives. We mitigated potential biases in our design process by engaging health experts to conduct heuristic evaluations that follow recommendations from the weight management literature for each incremental refinement.

Overall, the reflection on our human-centered design process emphasizes the importance of involving domain experts throughout the design process because user preferences may not always align with optimal design choices. We also highlight the value of working with a non-target population, which is less difficult to recruit, in the early design phase to address obvious issues, while considering potential biases stemming from contrasting preferences between user groups.

7 Limitations and future work

There are several limitations to our work. First, as mentioned previously, we recruited local university students instead of the target user group for the initial usability stage. While this decision had advantages (rapidly advancing the design process), it also carried the risk of falling to a local optimum without feedback from target users. To mitigate this risk, we employed mental walkthroughs of target scenarios during the usability study and actively engaged health experts in the decision-making process for refining our design. Future work can consider involving target users at earlier stages to better tailor designs to their needs.

Second, it is important to note that we did not perform a comprehensive analysis of the conversations in our semi-structured interviews and measures of usability from questionnaire responses. However, the focus of our work was to effectively utilize an iterative human-centered design process to implement a companion app that supports the implementation and evaluation of our adaptive weight maintenance intervention, rather than developing a comprehensive mHealth system for weight maintenance for all cases. Thus, we decided to prioritize the advancement of our clinical trial for evaluating the adaptive weight intervention. Future work can conduct a comprehensive analysis of the conversations in our semi-structured interviews and incorporate measures of usability from questionnaire responses to establish generalizable design guidelines on weight management app implementation, specifically for dual app use cases.

8 Conclusion

This paper presents the results of a user-centered design process with iterative evaluation and incremental refinements of prototypes to develop a companion app for supporting a weight loss and weight maintenance clinical trial. Our process included initial design brainstorming with our team of HCI and health experts, a basic usability study with 17 university students, and a 2-week pilot deployment with 33 target users. Overall, our project aimed to accelerate the implementation of our clinical trial on adaptive weight maintenance interventions by leveraging existing commercial apps and developing a secondary app to meet our specific research requirements, such as instrumentation needed to collect relevant data. Our primary focus was on facilitating effortless activity, food, and weight logging by the target users. We have also explored strategies to enhance motivation, self-efficacy, self-awareness, app engagement, and adherence to health goals, drawing upon behavior change theory as a guiding framework. We demonstrated the effectiveness of utilizing an iterative human-centered design process to implement a companion app for supporting the implementation and evaluation of our adaptive weight maintenance intervention.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by University of Florida Institutional Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

Y-PC: Conceptualization, Data curation, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing; JW: Writing – original draft, Writing – review & editing, Conceptualization, Data curation, Investigation, Methodology, Project administration; MN.S: Writing – original draft, Writing – review & editing, Conceptualization, Data curation, Investigation, Methodology, Project administration; DB: Writing – original draft, Writing – review & editing; UU: Writing – review & editing, Conceptualization, Data curation, Investigation; AB: Writing – review & editing, Conceptualization, Data curation, Investigation; KM.R: Funding acquisition, Resources, Supervision, Writing – original draft, Writing – review & editing, Conceptualization, Data curation, Investigation, Methodology, Project administration; JR: Funding acquisition, Resources, Software, Supervision, Writing – review & editing, Conceptualization, Data curation, Investigation, Methodology, Project administration; LA: Writing – original draft, Writing – review & editing, Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article.

This work was supported by the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health Grant under award #R01DK119244. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2024.1334058/full#supplementary-material

References

1. Laranjo L, Ding D, Heleno B, Kocaballi B, Quiroz JC, Tong HL, et al. Do smartphone applications and activity trackers increase physical activity in adults? systematic review, meta-analysis and metaregression. *Br J Sports Med*. (2021) 55:422–32. doi: 10.1136/bjsports-2020-102892

2. Hales CM, Carroll MD, Fryar CD, Ogden CL. *Prevalence of obesity and severe obesity among adults: United States, 2017–2018*. NCHS Data Brief, no 360. Hyattsville, MD: National Center for Health Statistics (2020). https://www.cdc.gov/nchs/products/databriefs/db360.htm

3. Ross Middleton KM, Patidar SM, Perri MG. The impact of extended care on the long-term maintenance of weight loss: a systematic review and meta-analysis. *Obes Rev*. (2012) 13:509–17. doi: 10.1111/j.1467-789X.2011.00972.x

4. Dombrowski SU, Knittle K, Avenell A, Araújo-Soares V, Sniehotta FF. Long term maintenance of weight loss with non-surgical interventions in obese adults: systematic review and meta-analyses of randomised controlled trials. *BMJ*. (2014) 348:g2646. doi: 10.1136/bmj.g2646

5. Carpenter CA, Ugwoaba UA, Cardel MI, Ross KM. Using self-monitoring technology for nutritional counseling and weight management. *Digit Health*. (2022) 8:20552076221102774. doi: 10.1177/20552076221102774

6. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. *Ann Behav Med*. (2018) 52:446–62. doi: 10.1007/s12160-016-9830-8

7. Forman EM, Goldstein SP, Crochiere RJ, Butryn ML, Juarascio AS, Zhang F, et al. Randomized controlled trial of OnTrack, a just-in-time adaptive intervention designed to enhance weight loss. *Transl Behav Med*. (2019) 9:989–1001. doi: 10.1093/tbm/ibz137

8. National Institutes of Health. Evaluation of an adaptive intervention for weight loss maintenance (2019). Identifier NCT04116853. Available online at: https://clinicaltrials.gov/study/NCT04116853 (Accessed November 1, 2023).

9. FatSecret. FatSecret - calorie counter and diet tracker for weight loss (2023). Available online at: https://www.fatsecret.com/ (Accessed July 13, 2023).

10. FatSecret Platform. The largest global nutrition database, recipe and food API (2023). Available online at: https://platform.fatsecret.com/api/ (Accessed July 13, 2023).

11. BodyTrace Inc. Accurate, real-time weight tracking (2023). Available online at: https://www.bodytrace.com/medical/ (Accessed July 13, 2023).

12. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. *Br J Health Psychol*. (2010) 15:1–39. doi: 10.1348/135910709X466063

13. Zhao J, Freeman B, Li M. Can mobile phone apps influence people's health behavior change? an evidence review. *J Med Internet Res*. (2016) 18:e287. doi: 10.2196/jmir.5692

14. Dounavi K, Tsoumani O. Mobile health applications in weight management: a systematic literature review. *Am J Prev Med*. (2019) 56:894–903. doi: 10.1016/j.amepre.2018.12.005

15. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. *Int J Behav Nutr Phys Act*. (2016) 13:127. doi: 10.1186/s12966-016-0454-y

16. Direito A, Carraça EV, Rawstorn JC, Whittaker R, Maddison R. mHealth technologies to influence physical activity and overall effectiveness: behavior change techniques, systematic review and meta-analysis of randomized controlled trials. *Ann Behav Med*. (2017) 51:226–39. doi: 10.1007/s12160-016-9846-0

17. Milne-Ives M, Lam C, De Cock C, Van Velthoven MH, Meinert E. Mobile apps for health behavior change in physical activity, diet, drug and alcohol use, and mental health: systematic review. *JMIR mHealth uHealth*. (2020) 8:e17046. doi: 10.2196/17046

18. Lyzwinski LN, Caffery LJ, Bambling M, Edirippulige S. Consumer perspectives on mHealth for weight loss: a review of qualitative studies. *J Telemed Telecare*. (2018) 24:290–302. doi: 10.1177/1357633X17692722

19. Haggag O, Grundy J, Abdelrazek M, Haggag S. A large scale analysis of mHealth app user reviews. *Empir Softw Eng*. (2022) 27:196. doi: 10.1007/s10664-022-10222-6

20. Schwarz A, Winkens L, de Vet E, Ossendrijver D, Bouwsema K, Simons M. Design features associated with engagement in mobile health physical activity interventions among youth: systematic review of qualitative and quantitative studies. *JMIR mHealth and uHealth*. (2023) 11:e40898. doi: 10.2196/40898

21. McCurdie T, Taneva S, Casselman M, Yeung M, McDaniel C, Ho W, et al. mHealth consumer apps: the case for user-centered design. *Biomed Instrum Technol*. (2012) Suppl:49–56. doi: 10.2345/0899-8205-46.s2.49

22. Beck AK, Kelly PJ, Deane FP, Baker AL, Hides L, Manning V, et al. Developing a mHealth routine outcome monitoring and feedback app ("SMART Track") to support self-management of addictive behaviours. *Front Psychiatry*. (2021) 12:677637. doi: 10.3389/fpsyt.2021.677637

23. Wies B, Landers C, Ienca M. Digital mental health for young people: a scoping review of ethical promises and challenges. *Front Digit Health*. (2021) 3:697072. doi: 10.3389/fdgth.2021.697072

24. Schnall R, Mosley JP, Iribarren SJ, Bakken S, Carballo-Diéguez A, Brown Iii W. Comparison of a user-centered design, self-management app to existing mHealth apps for persons living with HIV. *JMIR mHealth and uHealth*. (2015) 3:e91. doi: 10.2196/mhealth.4882

25. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). *J Biomed Inform*. (2016) 60:243–51. doi: 10.1016/j.jbi.2016.02.002

26. Eaves ER, Doerry E, Lanzetta SA, Kruithoff KM, Negron K, Dykman K, et al. Applying user-centered design in the development of a supportive mHealth app for women in substance use recovery. *Am J Health Promot*. (2023) 37:56–64. doi: 10.1177/08901171221113834

27. Hsieh KL, Frechette ML, Fanning J, Chen L, Griffin A, Sosnoff JJ. The developments and iterations of a mobile technology-based fall risk health application. *Front Digit Health*. (2022) 4:828686. doi: 10.3389/fdgth.2022.828686

28. Ross KM, Qiu P, You L, Wing RR. Week-to-week predictors of weight loss and regain. *Health Psychol*. (2019) 38:1150–8. doi: 10.1037/hea0000798

29. Ross KM, You L, Qiu P, Shankar MN, Swanson TN, Ruiz J, et al. Predicting high-risk periods for weight regain following initial weight loss. *Obesity*. (2024) 32:41–9. doi: 10.1002/oby.23923

30. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev*. (1977) 84:191–215. doi: 10.1037/0033-295X.84.2.191

31. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annu Rev Clin Psychol*. (2008) 4:1–32. doi: 10.1146/annurev.clinpsy.3.022806.091415

32. Miyake A, Takahashi M, Hashimoto R, Nakatani M. Stepup forecast: predicting future to promote walking. In: *Proceedings of the 23rd International Conference on Mobile Human–Computer Interaction*. New York, NY, USA: Association for Computing Machinery (2021). MobileHCI '21, p. 1–12. doi: 10.1145/3447526.3472020.

33. Dragoni M, Donadello I, Eccher C. Explainable ai meets persuasiveness: translating reasoning results into behavioral change advice. *Artif Intell Med*. (2020) 105:101840. doi: 10.1016/j.artmed.2020.101840

34. Locke EA, Latham GP. Building a practically useful theory of goal setting and task motivation: a 35-year odyssey. *Am Psychol*. (2002) 57:705–17. doi: 10.1037/0003-066X.57.9.705

35. Pearson ES. Goal setting as a health behavior change strategy in overweight and obese adults: a systematic literature review examining intervention components. *Patient Educ Couns*. (2012) 87:32–42. doi: 10.1016/j.pec.2011.07.018

36. Shilts MK, Horowitz M, Townsend MS. Goal setting as a strategy for dietary and physical activity behavior change: a review of the literature. *Am J Health Promot*. (2004) 19:81–93. doi: 10.4278/0890-1171-19.2.81

37. Konrad A, Bellotti V, Crenshaw N, Tucker S, Nelson L, Du H, et al. Finding the adaptive sweet spot: balancing compliance and achievement in automated stress reduction. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. New York, NY: Association for Computing Machinery (2015). p. 3829–38. doi: 10.1145/2702123.2702512.

38. Rabbi M, Pfammatter A, Zhang M, Spring B, Choudhury T. Automated personalized feedback for physical activity and dietary behavior change with mobile phones: a randomized controlled trial on adults. *JMIR mHealth and uHealth*. (2015) 3:e42. doi: 10.2196/mhealth.4160

39. Fogg B. A behavior model for persuasive design. In: *Proceedings of the 4th International Conference on Persuasive Technology*. New York, NY, USA: Association for Computing Machinery (2009), Persuasive '09, p. 1–7. doi: 10.1145/1541948.1541999.

40. Muench F, Baumel A. More than a text message: dismantling digital triggers to curate behavior change in patient-centered health interventions. *J Med Internet Res*. (2017) 19:e147. doi: 10.2196/jmir.7463

41. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. *J Med Internet Res*. (2011) 13:e30. doi: 10.2196/jmir.1602

42. Festinger L. A theory of social comparison processes. *Hum Relat*. (1954) 7:117–40. doi: 10.1177/001872675400700202

43. Hou IC, Lan MF, Shen SH, Tsai PY, Chang KJ, Tai HC, et al. The development of a mobile health app for breast cancer self-management support in taiwan: design thinking approach. *JMIR mHealth and uHealth*. (2019) 8:e15780. doi: 10.2196/15780

44. Woodward J, Chen YP, Jurczyk K, Ross KM, Anthony L, Ruiz J. A survey of notification designs in commercial mHealth apps. In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery (2021), CHI EA '21, p. 1–7. doi: 10.1145/3411763.3451639.

45. Firebase. Firestore (2023). Available online at: https://firebase.google.com/docs/firestore (Accessed July 13, 2023).

46. Google Cloud. Cloud functions (2023). Available online at: https://cloud.google.com/functions (Accessed July 13, 2023).

47. American College of Sports Medicine. Trending topic — physical activity guidelines (2023). Available online at: https://www.acsm.org/education-resources/trending-topics-resources/physical-activity-guidelines (Accessed November 1, 2023).

48. Centers for Disease Control and Prevention. How much physical activity do adults need? (2022). Available online at: https://www.cdc.gov/physicalactivity/basics/adults/index.htm (Accessed November 1, 2023).

49. Nielsen J, *Usability Engineering*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. (1994).

50. Olson JS, Kellogg WA, *Ways of Knowing in HCI*. Springer Publishing Company: Incorporated (2014).

51. Greenberg S, Carpendale S, Marquardt N, Buxton B, *Sketching User Experiences: The Workbook*. 1st ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. (2011).

52. Brooke J. SUS: a 'Quick and Dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. *Usability Evaluation in Industry*. London: CRC Press (1996). p. 184–96.

53. Hyzy M, Bond R, Mulvenna M, Bai L, Dix A, Leigh S, et al. System usability scale benchmarking for digital health apps: meta-analysis. *JMIR mHealth uHealth*. (2022) 10:e37290. doi: 10.2196/37290

54. Zhang R, E Ringland K, Paan M, C Mohr D, Reddy M. Designing for emotional well-being: integrating persuasion and customization into mental health technologies. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery (2021), CHI '21, p. 1–13. doi: 10.1145/3411764.3445771.

55. Sundar SS, Marathe SS. Personalization versus customization: the importance of agency, privacy, and power usage. *Hum Commun Res*. (2010) 36:298–322. doi: 10.1111/j.1468-2958.2010.01377.x

56. Woodward J, Shaw A, Luc A, Craig B, Das J, Hall P, et al. Characterizing how interface complexity affects children's touchscreen interactions. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery (2016), CHI '16, p. 1921–33. doi: 10.1145/2858036.2858200.

57. Krukowski RA, Ross KM. Measuring weight with electronic scales in clinical and research settings during the coronavirus disease 2019 pandemic. *Obesity*. (2020) 28:1182–3. doi: 10.1002/oby.22851

58. Birkmeyer S, Wirtz BW, Langer PF. Determinants of mHealth success: an empirical investigation of the user perspective. *Int J Inf Manage*. (2021) 59:102351. doi: 10.1016/j.ijinfomgt.2021.102351

59. Liao QV, Subramonyam H, Wang J, Wortman Vaughan J. Designerly understanding: information needs for model transparency to support design ideation for ai-powered user experience. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery (2023), CHI '23, p. 1–21. doi: 10.1145/3544548.3580652.

60. Morris MR, Danielescu A, Drucker S, Fisher D, Lee B, Schraefel MC, et al. Reducing legacy bias in gesture elicitation studies. *Interactions*. (2014) 21:40–5. doi: 10.1145/2591689

61. Neuhauser L, Kreps GL, Morrison K, Athanasoulis M, Kirienko N, Van Brunt D. Using design science and artificial intelligence to improve health communication: ChronologyMD case example. *Patient Educ Couns*. (2013) 92:211–7. doi: 10.1016/j.pec.2013.04.006

62. Guo X, Zhang X, Sun Y. The privacy–personalization paradox in mHealth services acceptance of different age groups. *Electron Commer Res Appl*. (2016) 16:55–65. doi: 10.1016/j.elerap.2015.11.001

frontiers | Frontiers in Digital Health

TYPE Brief Research Report
PUBLISHED 10 May 2024
DOI 10.3389/fdgth.2024.1335713

OPEN ACCESS

EDITED BY
Nikki Holliday,
Coventry University, United Kingdom

REVIEWED BY
Namkee Choi,
The University of Texas at Austin, United States
Roberta Bevilacqua,
National Institute of Science and Health for Aging (IRCCS), Italy

*CORRESPONDENCE
Eline H. G. M. Collombon
✉ eline.collombon@ou.nl

RECEIVED 09 November 2023
ACCEPTED 30 April 2024
PUBLISHED 10 May 2024

CITATION
Collombon EHGM, Bolman CAW, de Bruijn G-J, Peels DA and Lechner L (2024) Recruitment strategies for reaching adults aged 50 years and older with low socioeconomic status for participation in online physical activity interventions. Front. Digit. Health 6:1335713. doi: 10.3389/fdgth.2024.1335713

Recruitment strategies for reaching adults aged 50 years and older with low socioeconomic status for participation in online physical activity interventions

Eline H. G. M. Collombon[1]*, Catherine A. W. Bolman[1], Gert-Jan de Bruijn[2], Denise A. Peels[1] and Lilian Lechner[1]

[1]Faculty of Psychology, Open Universiteit, Heerlen, Netherlands, [2]Department of Communication Studies, University of Antwerp, Antwerp, Belgium

Background: Generally, the health condition of those with higher socioeconomic status (SES) is better compared to those with lower SES. The application of appropriate strategies to reach low SES populations with electronic health (eHealth) interventions is thus of major importance to reduce health inequalities. eHealth-studies providing detailed information on recruitment strategies are scarce, despite the fact that this information is crucial for comparable research and implementation.
Objective: To provide insight into the reach, sample characteristics and costs of three pre-planned strategies for recruiting adults aged 50 years and older with low SES for participation in an online physical activity intervention, as part of a field study.
Methods: Recruitment took place via (1) invitation letters via a municipality, (2) gyms and (3) social media advertisements, aiming to include 400 participants. Additional procedures were followed to reach specifically the low SES group. Response rates, sociodemographic characteristics and costs per strategy were assessed.
Results: The highest response was shown for the municipality approach ($N = 281$), followed by social media ($N = 71$) and gyms ($N = 45$). Ten participants were recruited via family/friends. The most low-educated participants were reached via the municipality ($N = 128$) followed by social media ($N = 9$), gyms ($N = 8$) and family/friends ($N = 5$). Recruitment costs were with €2,142.37 the highest for the municipality compared to €96.81 for social media and no costs for gyms.
Conclusions: Recruitment via invitation letters through a municipality has the highest potential for reaching low SES participants of the three applied strategies, although the higher recruitment costs need to be taken into account.

KEYWORDS

older adults, low education, vulnerable populations, eHealth, mHealth

Abbreviations
eHealth, electronic health; PA, physical activity; RCTs, randomized controlled trials; RQ, research question; SES, socioeconomic status.

1 Introduction

Socioeconomic status (SES) is a term used to describe an individual's affluence or social standing, referencing factors such as wealth, educational level and occupation (1). The relationship between SES and health has been well-established, and shows that the health condition of those with higher SES is better compared to those with lower SES (2, 3). The influence of SES on health is closely related to people's lifestyle which includes health risk behaviors such as smoking and health promoting behaviors such as physical activity (PA) (4). Given this, it is remarkable that lifestyle-related electronic health (eHealth) and mobile health (mHealth) interventions are often not used by the low SES population (5). Although challenging, reaching participants with low SES with electronic health (eHealth) interventions is, particularly in the light of the current digital era, of major importance to reduce health inequalities due to SES instead of widening the gap.

An important step for reaching participants with low SES is the application of well-planned, appropriate and inclusive recruitment strategies. eHealth-studies providing detailed information on strategies applied to reach their participants are scarce, despite the fact that this information is crucial for comparable research. Available literature on recruitment strategies often focuses on clinical settings, randomized controlled trials (RCTs), not specifically on eHealth research and/or the SES group (6–9). Therefore, the aim of the current study was to provide insight into the reach, sample characteristics and costs of three pre-planned strategies for recruiting adults aged 50 years and older with low SES for participation in an online PA intervention consisting of three computer-based tailored PA advices combined with a mobile-based activity tracker (10), as part of a field study. To accomplish this, the following research questions (RQ) were investigated: (RQ1) Which recruitment strategy results in the highest and fastest response?, (RQ2) Which recruitment strategy is most suitable for reaching the low SES population?, (RQ3) Which recruitment strategy is most suitable for reaching populations with a specific gender, age or health status?, (RQ4) Which recruitment strategy is most beneficial with regard to costs?.

2 Methods

2.1 Participants

The aim for the field study was to include 400 participants aged 50 years and older with a focus on the low SES population, although other SES groups were not excluded from participation. An additional aim was that the sample consisted of 200 participants without a (chronic) disease and 200 participants with a (chronic) disease, since tailoring procedures within the online PA intervention were optimized prior to the field study specifically for these subgroups. With this study population distribution, detailed insights on use and appreciation of the intervention for subgroups with and without (chronic) diseases could be obtained, which is described elsewhere [in preparation] A criterium for being classified into the (chronic) disease group was that participants were limited in being physically active as a result of their disease. Additional inclusion criteria applicable to the total sample were: (1) able to use a computer, laptop or tablet, (2) having an e-mail address, (3) having a smartphone, (4) not previously participated in a study of the Active4Life project (11). Eligibility to participate was assessed during the online registration procedure. The aim of this paper was to provide insight into the recruitment strategies applied during the field study with a specific focus on reaching low SES groups.

2.2 Recruitment procedures

Three different pre-planned recruitment strategies were deployed parallel to each other, namely recruitment via (1) a municipality, (2) gyms and (3) social media. These strategies were selected and considered appropriate and feasible based on previous eHealth studies conducted within our research group (12, 13).

Firstly, residents aged 50 years and older of three low SES neighbourhoods in a municipality received a personal invitation letter by post on behalf of the municipality, VIE (a regional organization that stimulates lifestyle and vitality) and the university for participation in the online PA intervention. The letters were basically the same for all invited residents, but differed in detail for the age groups 50–64 years and 65 + years. The 50–64 letter focused more on the healthy population, whereas the 65 + letter focused more on the population with health complaints. However, in both letters it was emphasized that anyone aged 50 years and older could participate despite any health problems. Interested invitees could register via internet by entering the hyperlink stated in the paper-based invitation letter.

Secondly, recruitment took place via gyms affiliated with project partner NL Actief, the Dutch trade association for sports organizations. Gyms were invited by NL Actief to participate in the recruitment procedures based on their location in a low SES region. Gyms signing up to participate, received an online flyer to distribute among their (potential) members aged 50 years and older. Methods for distribution of the flyer were determined by gyms themselves, although options were provided by the researchers for guidance. Some gyms included the flyer for example in their newsletter, whereas others posted the flyer on their social media channels. Since the flyers were online, interested people could be directly forwarded to the information website and registration portal by clicking on the link in the flyer.

The third strategy was recruitment via social media advertisements on Facebook. In order to reach participants aged 50 years and older with low SES, three targeting settings on age, educational level and location were added to the advertisements. With regard to age, the advertisement was only shown to adults aged 50 years and older. For educational level, the advertisement was not shown to those who added to their Facebook profile that they were attending higher education or a master degree or those who received their higher education certificate, university degree, master degree or PhD. The advertisements were only shown within pre-selected Dutch regions based on the number of low-

educated persons living in the area combined with the degree of ageing (14). Also here, interested people were directly forwarded to the information website and registration portal by clicking on the link in the online flyer.

2.3 Measures and statistics

Recruitment method and the sociodemographics gender, age, educational level and (chronic) disease were assessed during the registration procedure. Educational level was categorized into low (i.e., primary, basic vocational or lower general school), middle (i.e., medium vocational school, higher general secondary education and preparatory academic education) and high (i.e., higher vocational school or university level) according to the Dutch educational system (15). Participants were classified into the (chronic) disease subgroup when they indicated during the registration procedure that they were limited in being physically active as a result of a (chronic) disease. Performance statistics of social media advertisements such as costs, reach and clicks on links were derived from the ad center of Facebook. Information on costs and reach of the personal invitation letters was provided by the municipality. Chi-squares and one-way analyses of variances (ANOVAs) were performed to test on an exploratory level for differences on the above mentioned sociodemographics between the recruitment strategies ($P \leq .05$).

3 Results

3.1 Which recruitment strategy results in the highest and fastest response?

Recruitment via the three pre-planned strategies was initiated in week 2. The participants that were recruited in week 1 ($N = 2$) were reached via social media, likely via advertisements for previous already completed studies of the Active4Life project where the same registration portal was used (11). The most participants were also recruited in the second week when pre-planned strategies were initiated ($N = 263$). Of those 263 participants, the majority was reached via the municipality ($N = 229$). In addition to the three pre-planned recruitment strategies, 10 participants indicated that they came into contact with the online PA intervention by family, friends or acquaintances. This unplanned and naturally developed recruitment method is reported as the fourth strategy. A schematic overview of the number of recruited participants per week in total and separated per strategy is provided in Figure 1. In the third week, 69 participants were recruited with the majority coming from gyms ($N = 29$). Afterwards, the number of recruited participants per week decreased gradually. With only 1 recruited participant in both the eighth and ninth week, it was decided that a new action was needed to boost recruitment. Since only 30 additional participants were needed to reach the planned sample size of 400, solely recruitment via social media was boosted through a new advertisement. No new actions were performed within the municipality and gyms approach, to avoid an overload of recruited participants. As a result, 37 participants were recruited via social media in the tenth week resulting in a total of 407 registered adults aged 50 years or older. No selection of recruited participants was needed to reach the pre-defined goal of including 200 participants with a (chronic) disease and 200 participants without a (chronic) disease. This distribution within the study population arose naturally during recruitment. The municipality approach delivered the most participants ($N = 281$, 69.0% of the total study population), followed by social media ($N = 71$, 17.4%), gyms ($N = 45$, 11.1%) and family + friends ($N = 10$, 2.5%).

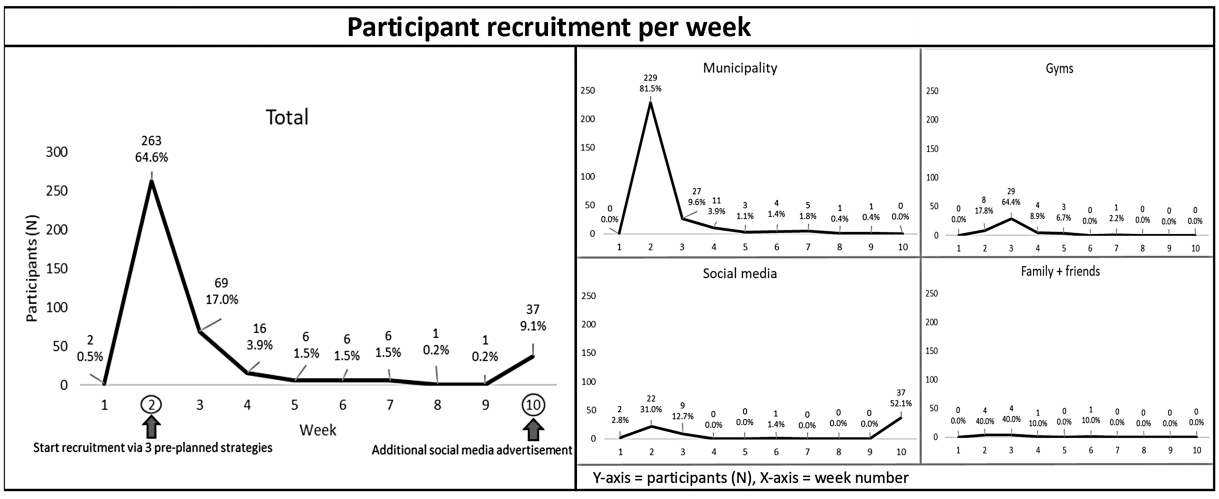


FIGURE 1
Overview of number of recruited participants per week.

3.2 Which recruitment strategy is most suitable for reaching the low SES population?

Sociodemographic characteristics of recruited participants in total and per strategy are shown in Table 1. In total, more low- ($N = 150$, 36.9% of the total study population) and middle-educated participants ($N = 149$, 36.6%) were reached than high-educated participants ($N = 108$, 26.5%). Significantly more low-educated participants were reached via the municipality ($N = 128$, 45.6% of the municipality group) than via gyms ($N = 8$, 17.8%) and social media ($N = 9$, 12.7%) ($\chi^2 = 50.429$, $P < .001$).

3.3 Which recruitment strategy is most suitable for reaching populations with a specific gender, age or health status?

Table 1 shows that with regard to gender, an almost equal amount of males ($N = 142$, 50.5% of the municipality group) and females ($N = 139$, 49.5%) was reached via the municipality. Significantly more males were reached via the municipality compared to gyms ($N = 15$, 33.3%) and social media ($N = 8$, 11.3%) ($\chi^2 = 60.546$, $P < .001$). Further, participants reached via social media were significantly younger than participants reached via the municipality (mean age of 60.0 vs. 64.0 years) ($F = 4.812$, $P = .003$). Proportionally more participants without a (chronic) disease were reached via gyms (60.0% of the gyms group), social media (59.2%) and family + friends (60.0%), whereas proportionally more participants with a (chronic) disease were reached via the municipality (54.8%). However, this difference did not reach significance ($\chi^2 = 7.056$, $P = .070$).

3.4 Which recruitment strategy is most beneficial with regard to costs?

An overview of the number of participants reached per strategy in relation to recruitment costs is provided in Table 2. Costs for recruitment via the municipality were with a total of €2,142.37 ($2,341.06) and €7.62 ($8.33) per yielded participant the highest, which comprised the letter/envelope-, printing- and delivery costs for 3,417 personal invitation letters. With regard to social media, the first advertisement ran for 7 days, cost €48.95 ($53.74), reached 7,316 Facebook users and resulted in 167 clicks on the link leading to the registration website. The second advertisement ran for 5 days, cost €50.09 ($54.99), reached 5,624 Facebook users and resulted in 179 clicks on the link leading to the registration website. Total costs for the social media recruitment were thus €96.81 ($105.79) and €1.36 ($1.49) per yielded participant. The ratio between the number of potential participants reached with the strategy and the number that actually registered to participate was with 8.2% for the municipality higher compared to 0.5% for social media. No data were available for the number of potential participants reached via the other two recruitment strategies.

4 Discussion

4.1 Principal findings

The aim of the current study was to provide insight into the reach, sample characteristics and costs of three pre-planned strategies for recruiting adults aged 50 years and older with low SES for participation in an online PA intervention including an activity tracker. The accompanying research questions can be answered based on the results of the study. Firstly, it can be concluded that recruitment via personal invitation letters through a municipality results in the highest and fastest response when compared to recruitment via gyms and social media. The observed highest reach/registered ratio for recruitment via the municipality is in line with other studies, showing that the application of more personalized approaches leads to higher enrollment rates (16) (RQ1). Additionally, the paper-based recruitment approach of sending personal invitation letters via the municipality was the most suitable for reaching the low SES population based on the strategies applied during the study

TABLE 1 Sociodemographic characteristics of recruited participants.

	Municipality $N = 281$	Gyms $N = 45$	Social media $N = 71$	Family + friends $N = 10$	P	Total $N = 407$
Gender					<.001*	
Male, N (%)	142 (50.5)[a]	15 (33.3)[a]	8 (11.3)[b]	3 (30.0)[a,b]		168 (41.3)
Female, N (%)	139 (49.5)[a]	29 (64.4)[a]	63 (88.7)[b]	6 (60.0)[a,b]		237 (58.2)
Other, N (%)	0 (0.0)[a]	1 (2.2)[a,b]	0 (0.0)[a]	1 (10.0)[b]		2 (0.5)
Age in years, mean (SD)	64.0 (8.6)[a]	61.7 (8.5)[a,b]	60.0 (7.0)[a]	61.3 (9.5)[a,b]	.003*	63.0 (8.4)
Educational level					<.001*	
Low, N (%)	128 (45.6)[a]	8 (17.8)[b,c]	9 (12.7)[c]	5 (50.0)[a,b]		150 (36.9)
Middle, N (%)	103 (36.7)[a]	15 (33.3)[a]	29 (40.8)[a]	2 (20.0)[a]		149 (36.6)
High, N (%)	50 (17.8)[a]	22 (48.9)[b]	33 (46.5)[b]	3 (30.0)[a,b]		108 (26.5)
(Chronic) disease					.070	
Yes	154 (54.8)	18 (40.0)	29 (40.8)	4 (40.0)		200 (49.1)
No	127 (45.2)	27 (60.0)	42 (59.2)	6 (60.0)		207 (50.9)

[a,b,c]Each superscript letter denotes a subset of group categories whose column proportions do not differ significantly from each other at the .05 level.
*$P \leq .05$.

10.3389/fdgth.2024.1335713

TABLE 2 Overview number of participants reached in relation to recruitment costs.

	Municipality	Gyms	Social media	Family + friends
Total recruitment costs				
In Euros (€)	2,142.37	0	96.81	0
In Dollars ($)[a]	2,341.06	0	105.79	0
Potential participants reached, *N*	3,417	Unknown	12,940	Unknown
Total participants, *N*	281	45	71	10
Ratio reached/ registered, %	8.2	Unknown	0.5	Unknown
Costs per participant				
In Euros (€)	7.62	0	1.36	0
In Dollars ($)	8.33	0	1.49	0
Low educated participants				
N	128	8	9	5
% of total	45.6	17.8	12.7	50.0
Costs per low educated participant				
In Euros (€)	16.73	0	10.76	0
In Dollars ($)	18.28	0	11.76	0

[a]Values in Euros converted to Dollars based on exchange rate of 06/20/2023.

(RQ2).. No paper-based recruitment approaches were applied for the gyms and social media since only online advertisements were used. Although the low SES population was thus quickly reached via the municipality, higher costs were involved for this paper-based strategy compared to the other completely online strategies. These higher costs need to be taken into account when selecting a recruitment strategy (RQ4). Figure 2 provides a decision tree for selecting the appropriate recruitment strategy in order to reach populations with specific sociodemographics based on the relative results of the study (RQ3). In particular, the result that relatively a larger amount of males (50.5%) was reached via the municipality approach is relevant since previous eHealth-studies have shown that males are more difficult to reach for participation in lifestyle-related interventions compared to females (17). The same applies to the findings on educational level, since low-educated participants are more difficult to recruit than middle- and high-educated participants (18). The decision tree guides future eHealth-studies in selecting an appropriate recruitment strategy.

4.2 Strengths

The application of different recruitment strategies parallel to each other can be considered a strength of the study. With this method, the target sample size of 400 participants aged 50 years and older was achieved within 10 weeks. This is a short period of time when compared to other eHealth-studies, where problems with reaching the sample size are frequently reported (19). The fact that participants received an intervention consisting of three tailored online PA advices combined with an activity tracker, which they were allowed to keep after completion of the study, contributed possibly to quickly reaching the desired sample size. However, considerations for participants to register were not investigated. More insight into reasons to participate could be valuable.

Further, the selected recruitment strategies can be considered successful since the population with low educational level was reached within this study. In particular, relatively large amounts of low educated participants were recruited via the municipality approach. Since the selected municipality has one of the lowest SES scores within the Netherlands (20), it is expected that also the low SES population was reached during recruitment. The additional procedures that were followed within the different

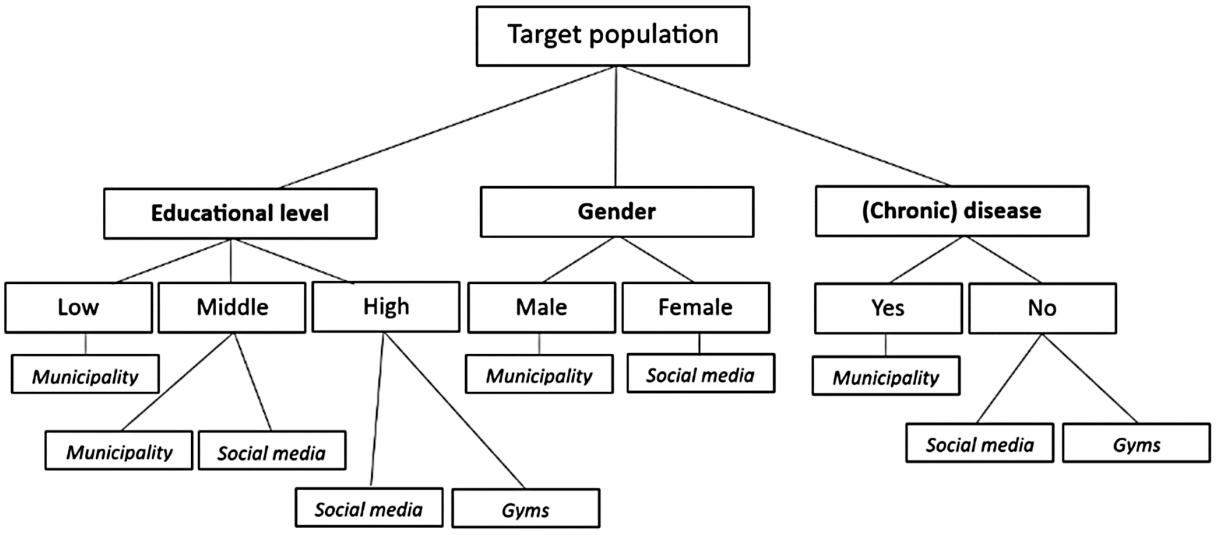


FIGURE 2
Decision tree for selecting a recruitment strategy to reach a specific target population of adults aged 50 years and older.

recruitment strategies likely contributed to successfully reaching the low SES population as well. Among others, recruitment was deployed in low SES regions and targeting variables were added to social media advertisements. These additional actions are considered essential for reaching the low SES population. The sample characteristics of our preceding RCT underline this need, since mainly high-educated participants were reached while applying a more general recruitment strategy [submitted as Collombon EHGM, Bolman CAW, de Bruijn GJ, Peels DA, Verboon P, Lechner L: The efficacy of online physical activity interventions with added mobile elements within adults aged 50 years and over: a randomized controlled trial.], which is in line with other studies (13).

Lastly, the practical study design of this field study and the accompanying high external validity can be considered a strength. Commonly, recruitment studies present results of experimental study designs (7) which impedes the application of these results in practice.

4.3 Limitations

During this study, only educational level was assessed as outcome measure of SES. A more extensive insight into SES would have been obtained by assessing other factors as well (1, 21). An example is financial status, although it is expected that questions related to financial situation are not accepted by participants, as shown in previous studies (22). Since recruitment was employed specifically in low SES regions (municipality and gyms approach) it is expected that the low SES population was also reached while reaching low educated participants. However, this assumption should be interpreted with caution.

Further, it should be taken into account that this study was conducted in the Netherlands. The findings on recruitment strategies are only generalizable to other countries with caution. Comparable follow-up studies in other countries are recommended to confirm or refute our findings.

Although this study provides guidelines for reaching the low SES population, successful recruitment strategies alone are not sufficient. After registration, it is important that participants actually use an eHealth intervention. To encourage this, it is important that the characteristics and needs of the low SES population are taken into account during the design process of an intervention (23). It has namely been shown that this population faces more frequently low eHealth-literacy which comprises "a set of skills and knowledge that are essential for productive interactions with technology-based tools" (24–26). Populations with low eHealth-literacy are often not involved in research (27). Additionally, the low SES group has different life situations and eHealth expectations compared to the high SES group (28). Neglecting these differences during eHealth design processes can even exacerbate the digital divide and health inequalities instead of bridging this gap (29).Involving the target population in intervention design processes presents a solution for this (30).

4.4 Future recommendations

Based on the results of this study, future eHealth-studies are recommended to recruit via personal invitation letters through a municipality in order to reach the low SES population. It remains unclear whether solely the sending of paper-based personal invitation letters, solely approaching the target population via a municipality or the combination of both was responsible for successfully reaching the low SES group via this recruitment strategy. Future research could elucidate this. The degree of personalization could be further optimized by having contact with potential participants via telephone or face-to-face instead of sending personal invitation letters. Future studies are recommended to investigate whether more personalized approaches yield more low SES participants. However, it has to be taken into account that this is more time-consuming and costly compared to solely sending personal invitation letters.

Recruitment via family + friends was not a pre-planned strategy during this study and arose naturally. Future studies are recommended to investigate whether a pre-planned strategy via family and friends through snowball sampling has potential for reaching low SES participants. Although snowball sampling is commonly used within qualitative research (31), it might also be useful for reaching potential participants for participation in an online PA intervention.

4.5 Conclusions

In conclusion, recruitment via personal invitation letters through a municipality has the highest potential for reaching low SES participants of the three strategies applied during this study. However, higher costs are involved compared to recruitment via gyms and social media. Revealed insights on the sociodemographics gender, educational level and health status per recruitment strategy can guide future eHealth-studies to select appropriate strategies for reaching their specific target population of adults aged 50 years and older.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by central ethical review committee of the Open Universiteit. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

EC: Conceptualization, Data curation, Methodology, Writing – original draft, Writing – review & editing. CB: Conceptualization, Methodology, Writing – review & editing. G-JB: Conceptualization,

Methodology, Writing – review & editing. DP: Conceptualization, Methodology, Writing – review & editing. LL: Conceptualization, Methodology, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article.

This study was funded by the Netherlands Organization for Health Research and Development (ZonMW, 546003005).

Acknowledgments

The authors would like to thank the municipality of Kerkrade, VIE and NL Actief for their contribution to recruitment procedures.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Baker EH. *Socioeconomic Status, Definition*. The Wiley Blackwell encyclopedia of health, illness, behavior, and society (2014). p. 2210–4. doi: 10.1002/9781118410868.wbehibs395

2. Harper S, Lynch J. Trends in socioeconomic inequalities in adult health behaviors among US states, 1990–2004. *Public Health Rep*. (2007) 122(2):177–89. doi: 10.1177/003335490712200207

3. Mackenbach JP, Stirbu I, Roskam A-JR, Schaap MM, Menvielle G, Leinsalu M, et al. Socioeconomic inequalities in health in 22 European countries. *N Engl J Med*. (2008) 358(23):2468–81. doi: 10.1056/NEJMsa0707519

4. Simandan D. Rethinking the health consequences of social class and social mobility. *Soc Sci Med*. (2018) 200:258–61. doi: 10.1016/j.socscimed.2017.11.037

5. Schneider F, van Osch L, Schulz DN, Kremers SP, de Vries H. The influence of user characteristics and a periodic email prompt on exposure to an internet-delivered computer-tailored lifestyle program. *J Med Internet Res*. (2012) 14(2):e40. doi: 10.2196/jmir.1939

6. Healy P, Galvin S, Williamson PR, Treweek S, Whiting C, Maeso B, et al. Identifying trial recruitment uncertainties using a james lind alliance priority setting partnership–the PRioRiTy (prioritising recruitment in randomised trials) study. *Trials*. (2018) 19(1):1–12. doi: 10.1186/s13063-017-2423-4

7. Treweek S, Mitchell E, Pitkethly M, Cook J, Kjeldstrøm M, Johansen M, et al. Strategies to improve recruitment to randomised controlled trials. *Cochrane Database Syst Rev*. (2010) 4:MR000013. doi: 10.1002/14651858.MR000013.pub5

8. Stuber JM, Middel CN, Mackenbach JD, Beulens JW, Lakerveld J. Successfully recruiting adults with a low socioeconomic position into community-based lifestyle programs: a qualitative study on expert opinions. *Int J Environ Res Public Health*. (2020) 17(8):2764. doi: 10.3390/ijerph17082764

9. Stanczyk N, Bolman C, Smit E, Candel M, Muris J, De Vries H. How to encourage smokers to participate in web-based computer-tailored smoking cessation programs: a comparison of different recruitment strategies. *Health Educ Res*. (2014) 29(1):23–40. doi: 10.1093/her/cyt104

10. Collombon EH, Peels DA, Bolman CA, de Bruijn G-J, Lechner L. Adding mobile elements to online physical activity interventions for adults aged over 50 years: prototype development study. *JMIR Formative Research*. (2023) 7:e42394. doi: 10.2196/42394

11. Collombon EH, Bolman CA, Peels DA, de Bruijn G-J, De Groot RH, Lechner L. Adding mobile elements to online physical activity interventions targeted at adults aged 50 years and older: protocol for a systematic design. *JMIR Res Protoc*. (2022) 11(7):e31677. doi: 10.2196/31677

12. Peels DA, Bolman C, Golsteijn RHJ, De Vries H, Mudde AN, van Stralen MM, et al. Differences in reach and attrition between web-based and print-delivered tailored interventions among adults over 50 years of age: clustered randomized trial. *J Med Internet Res*. (2012) 14(6):e2229. doi: 10.2196/jmir.2229

13. Coumans J. MyLifestyleCoach: Development and evaluation of a web-based computer-tailored intervention for diet and physical activity guided by self-determination theory and motivational interviewing. (2022).

14. Blaauwberg. Available online at: https://blaauwberg.nl/analyse-hoe-veerkrachtig-is-uw-regio/ (cited June 21, 2023).

15. Available online at: https://www.cbs.nl/nl-nl/nieuws/2019/33/verschil-levensverwachting-hoog-en-laagopgeleid-groeit/opleidingsniveau (cited August 3, 2023).

16. Al-Dhahir I, Reijnders T, Faber JS, van den Berg-Emons RJ, Janssen VR, Kraaijenhagen RA, et al. The barriers and facilitators of eHealth-based lifestyle intervention programs for people with a low socioeconomic status: scoping review. *J Med Internet Res*. (2022) 24(8):e34229. doi: 10.2196/34229

17. Muellmann S, Forberger S, Möllers T, Bröring E, Zeeb H, Pischke CR. Effectiveness of eHealth interventions for the promotion of physical activity in older adults: a systematic review. *Prev Med*. (2018) 108:93–110. doi: 10.1016/j.ypmed.2017.12.026

18. Nicholson LM, Schwirian PM, Klein EG, Skybo T, Murray-Johnson L, Eneli I, et al. Recruitment and retention strategies in longitudinal clinical studies with low-income populations. *Contemp Clin Trials*. (2011) 32(3):353–62. doi: 10.1016/j.cct.2011.01.007

19. Chesser A, Burke A, Reyes J, Rohrberg T. Navigating the digital divide: a systematic review of eHealth literacy in underserved populations in the United States. *Informatics for Health and Social Care*. (2016) 41(1):1–19. doi: 10.3109/17538157.2014.948171

20. Available online at: https://www.cbs.nl/nl-nl/achtergrond/2022/14/statusscore-per-wijk-en-buurt-o-b-v-welvaart-opleidingsniveau-en-arbeid (cited April 10, 2024).

21. voor de Statistiek CB. Bevolking; hoogstbehaald onderwijsniveau en onderwijsrichting (in Dutch) (2018). Available online at: https://opendatacbs nl/statline/#/CBS/nl/dataset/82816NED/table (Accessed June 21, 2023).

22. Stanczyk N, De Vries H, Candel M, Muris J, Bolman C. Effectiveness of video- versus text-based computer-tailored smoking cessation interventions among smokers after one year. *Prev Med*. (2016) 82:42–50. doi: 10.1016/j.ypmed.2015.11.002

23. Bukman AJ, Teuscher D, Feskens EJ, van Baak MA, Meershoek A, Renes RJ. Perceptions on healthy eating, physical activity and lifestyle advice: opportunities for adapting lifestyle interventions to individuals with low socioeconomic status. *BMC public Health*. (2014) 14:1–11. doi: 10.1186/1471-2458-14-1036

24. Chan CV, Kaufman DR. A framework for characterizing eHealth literacy demands and barriers. *J Med Internet Res*. (2011) 13(4):e1750. doi: 10.2196/jmir.1750

25. Choi NG, DiNitto DM. The digital divide among low-income homebound older adults: internet use patterns, eHealth literacy, and attitudes toward computer/internet use. *J Med Internet Res*. (2013) 15(5):e93. doi: 10.2196/jmir.2645

26. Western MJ, Armstrong ME, Islam I, Morgan K, Jones UF, Kelson MJ. The effectiveness of digital interventions for increasing physical activity in individuals of low socioeconomic status: a systematic review and meta-analysis. *Int J Behav Nutr Phys Act*. (2021) 18(1):1–21. doi: 10.1186/s12966-021-01218-4

27. Cheng C, Beauchamp A, Elsworth GR, Osborne RH. Applying the electronic health literacy lens: systematic review of electronic health interventions targeted at socially disadvantaged groups. *J Med Internet Res*. (2020) 22(8):e18476. doi: 10.2196/18476

28. Faber JS, Al-Dhahir I, Reijnders T, Chavannes NH, Evers AW, Kraal JJ, et al. Attitudes toward health, healthcare, and eHealth of people with a low

socioeconomic status: a community-based participatory approach. *Front Digit Health*. (2021) 3:690182. doi: 10.3389/fdgth.2021.690182

29. Crawford A, Serhal E. Digital health equity and COVID-19: the innovation curve cannot reinforce the social gradient of health. *J Med Internet Res*. (2020) 22(6):e19361. doi: 10.2196/19361

30. Hesse BW, Shneiderman B. Ehealth research from the user's perspective. *Am J Prev Med*. (2007) 32(5):S97–S103. doi: 10.1016/j.amepre.2007.01.019

31. Parker C, Scott S, Geddes A. *Snowball Sampling*. SAGE research methods foundations (2019).

frontiers | Frontiers in Digital Health

TYPE Original Research
PUBLISHED 12 September 2024
DOI 10.3389/fdgth.2024.1343965

OPEN ACCESS

EDITED BY
Kim Cornelia Marie-Louise Bul,
Coventry University, United Kingdom

REVIEWED BY
Elissa Z. Faro,
University of Iowa Health Care, United States
Ben Singh,
University of South Australia, Australia
Miriam Nkangu,
University of Ottawa, Canada

*CORRESPONDENCE
Priscilla Maliwichi
pmaliwichi@must.ac.mw

RECEIVED 24 November 2023
ACCEPTED 19 August 2024
PUBLISHED 12 September 2024

CITATION
Sowon K, Maliwichi P, Chigona W and Malata A (2024) The role of the community of purpose in maternal mHealth interventions in Sub-Saharan Africa context.
Front. Digit. Health 6:1343965.
doi: 10.3389/fdgth.2024.1343965

The role of the community of purpose in maternal mHealth interventions in Sub-Saharan Africa context

Karen Sowon[1], Priscilla Maliwichi[1,2]*, Wallace Chigona[1] and Address Malata[3]

[1]Department of Information Systems, University of Cape Town, Cape Town, South Africa, [2]Department of Computer Science and Information Technology, Malawi University of Science and Technology, Limbe, Malawi, [3]Office of the Vice Chancellor, Malawi University of Science and Technology, Limbe, Malawi

Background: mHealth has increasingly been touted as having the potential to help Sub-Saharan Africa achieve their health-related sustainable development goals by reducing maternal mortality rates. Such interventions are implemented as one-way or two-way systems where maternal clients receive pregnancy related information via SMS. While such technologies often view the users (the maternal health client) as having agency to adopt, we know from pregnancy literature that the pregnancy experience in Africa and other developing countries is often more collective. In addition to the maternal health client, other members of the community have high stakes in the pregnancy, and this often affects maternal healthcare-seeking behavior.
Objective: The aim of this paper, therefore, is to understand the pathways through which these other members of the community affect mHealth use.
Methods: The study used a qualitative approach and a case study research design. We analyzed two mHealth cases from Kenya and Malawi. In the Kenyan case, maternal health clients had mobile phones to receive pregnancy-related messages, while in the Malawi case, maternal health clients did not have mobile phones. Data were collected through interviews and focus group discussions. The study used an inductive thematic analysis to analyze the data.
Results: The findings show that maternal stakeholders form a community of purpose (CoP) that plays a crucial role in the implementation, uptake, and use of mHealth. The CoP influences maternal health clients through a diverse range of mechanisms ranging from sensitization, bridging the digital literacy gap and legitimization of the intervention. The nature of influence is largely dependent on the contextual socio-cultural nuances.
Conclusion: Our results provide useful insights to mHealth implementers to know how best to leverage the CoP for better mHealth uptake and usage. For example, engaging healthcare providers could champion adoption and use, while engaging other family-related stakeholders will ensure better usage and compliance, encourage behavior change, and reduce mHealth attrition.

KEYWORDS

maternal healthcare seeking, digital health, mHealth, community of purpose, social capital, mHealth adoption and use, mobile health, mHealth intervention

1 Introduction

Maternal mortality is still disproportionately high in countries in Sub-Saharan Africa (SSA) compared to more developed nations. The risk of a woman dying is estimated to be 1 in 6 in poorer economies compared to 1 in 30,000 in places such as Northern Europe (1). Developing countries also account for almost all maternal deaths, with almost 70% concentrated in SSA alone (2, 3). The global community has therefore endorsed the reduction of maternal mortality as a critical development goal–with an ambitious target of a mortality rate that is less than 70 per 100,000 by 2030 (4).

Concurrently, the proliferation of mobile phones has led to the rise of mHealth as a potential tool to overcome traditional healthcare barriers and challenges and to meet health-related sustainable development goals. mHealth is characterized as health interventions that employ mobile technologies, including mobile phones, wearable devices, personal digital assistants, tablet PCs, and similar devices. For this study, our definition is limited to the use of mobile phones. The technology leverages the access and portability of mobile phones to provide healthcare services to healthcare consumers (5). The maternal health landscape is filled with projects and studies that demonstrate various applications of mHealth to facilitate point of care, data collection, patient monitoring, and the delivery of health information (6–8). Given the limited technological context of many emerging economies, most mHealth interventions are implemented to disseminate pregnancy-related information, as well as reminders of antenatal care (ANC) visits via short message service (SMS) in one-way or two-way implementations (9). SMS is also the most preferred method to ensure the inclusion of both basic and smartphone users – especially in SSA where approximately 51 percent of mobile phones are smartphones, while 41 percent are basic mobile phones (10). Most of these smartphone owners live in urban areas, while the majority of rural mobile owners own a basic mobile phone (11).

While the use of mHealth shows promise, the first challenge is that interventions are often designed with a single user in mind – a perspective that assumes a one-to-one relationship between an individual and a device such as a mobile phone or a health tracker (12, 13). However, in most contexts in SSA, the ownership of devices does not always follow a one-to-one paradigm. In 2022, the region had a mobile penetration rate of 43 percent (10). This shows that the continent is still lagging in terms of mobile phone ownership, a problem that is often addressed by phone sharing. Furthermore, the gender gap in mobile phone ownership is more pronounced, with women being 13 percent less likely to own a mobile phone than men (10). For example, 44.9 percent of men in Malawi own a mobile phone compared to 37.7 percent of women (14). This gap is especially significant in rural areas, where only 26 percent of women own a mobile phone compared to 47 percent of men (14). In such situations where maternal health clients do not own a mobile phone, access to pregnancy-related information can be through a shared mobile phone in which the owners of these mobile phones can be community health workers (CHWs), family members, friends, community members, or community volunteers (15, 16).

Similarly, even if we were to keep the assumption that all users have mobile phones, studies of health behavior indicate a trend in developing countries to seek advice on treatment from elderly family members (17). In maternal health, the importance of female relatives becomes particularly pronounced, as authority and decision making in pregnancy matters are often socially designated within the female sphere (18, 19). Beyond elderly female family members, the involvement of partners or husbands is significant in the pregnancy journey. In a Muslim society, it was observed that Muslim women were required to seek permission from their husbands for healthcare decisions (20). As decision makers and primary controllers of household resources, men play an essential role in influencing women's healthcare-seeking behavior (17, 21). Hence, the second challenge lies in the assumption that the maternal health user has complete autonomy to use, or not to use mHealth as would be expected in a more Western context.

These diverse stakeholders that offer social and cultural support to pregnant women form a Community of Purpose (CoP), a term that is inspired by the concept of community of practice that defines a group of people united by a common interest in a specific domain of knowledge (22). The CoP is defined as a community of people working toward a common goal, purpose, or objective (23)—in maternal health, this would be the common objective of a healthy pregnancy (24). The influence that such a group has is often based on the tangible and intangible resources that they share such as shared trust, adherence to group norms and sanctions. In essence, these define social capital (25). In broader terms, social capital entails any instance in which people cooperate for common ends on the basis of shared informal norms and values–which in the context of pregnancy entail all the social and cultural beliefs and norms around pregnancy.

Although behavioral health interventions have often focused on individual-level change, some empirical work has underscored the importance of social capital to health outcomes (26, 27). Given that the importance of other community members to maternal healthcare-seeeking is well established (17, 28, 29), it seems reasonable to conjecture that these influences go beyond traditional healthcare-seeking and also influence the use of mHealth in maternal health contexts. We posited that these relationships are likely to affect mHealth adoption and use. Therefore, research on the dynamics of behavior change in mHealth contexts where social structures such as the CoP are carried out within other sociocultural realities is needed. This understanding is particularly crucial, because, as described before, the pregnancy experience is more collective in Africa than it is in western contexts. Hence, the objective of this research is to examine the impact of the CoP as constructed and determined by gendered norms and sociocultural rules on the success of mHealth use in contexts where users may own phones and where they do not. The questions we seek to answer are:

- RQ1: What role does the CoP play in the adoption and use of maternal mHealth interventions in a Sub-Saharan Africa context?

- RQ2: How does the role of the CoP compare between contexts where users have mobile phones and where they do not?

We used a case study approach to answer these questions with two case studies from Kenya and Malawi. These countries were chosen for two reasons: (1) they both still have high maternal mortality rates and (2) they represent two countries with diverse mobile phone penetration realities which likely influence the use of mHealth interventions. While Kenya's mobile penetration rate is 117.2 percent, Malawi's is 57.2 percent (30).

2 Methods

The data used in this paper were part of two larger qualitative studies that we conducted separately in Kenya and Malawi. Both studies adopted a case study approach. The Kenyan arm of the study was completed between January and May 2019. The study in Malawi was completed between January and August 2020. This paper reports only on data related to the role of the CoP.

2.1 Study design: context and case descriptions

The two cases that we focused on in Kenya and Malawi respectively were the PROMPTS mHealth intervention and Chipatala Cha Pa Foni (CCPF). Given the diverse realities of mobile phones described previously, maternal health clients in Kenya accessed PROMPTS with their own phones. PROMPTS provided staged pregnancy-related information via two-way SMS.

In Malawi, the adoption and use of mHealth in rural settings is hampered by the ownership of mobile phones (11). The price of mobile phones, even the most basic ones, makes it difficult for most people in rural communities to afford a mobile phone due to limited financial resources (31). Other barriers to technology use include poor battery life of mobile phones, network coverage problems, and malfunctioning mobile phone keypad (32). The barrier of non-ownership of mobile phones makes mHealth beneficiaries use borrowed mobile phones to use mHealth interventions. In Malawi therefore, CCPF was targeted at both women with, and without phones. For the purpose of this study, we were interested in how the women without phones in Malawi used mHealth. The following subsections elaborate further on these two case studies:

2.1.1 Case study 1: PROMPTS Maternal mHealth Intervention—Kenya

PROMPTS was initially developed as a postnatal checklist intervention that community health workers were responsible for administering to women after birth, with the aim to encourage the uptake of postpartum care services (33). However, due to the limitations of in-person home visits in a staff-constrained environment, the program evolved into a mobile phone SMS service to reach more women. The postnatal checklist messages were adapted for use as SMS. The messages were also further developed and refined in consultation with the maternal health clients who participated in Focus Group Discussions (FGDs). These sessions aimed to solicit details of the information that women felt they needed after delivery. This was combined with information prioritized by the healthcare provider which they considered pertinent throughout the pregnancy continuum.

The emerging version of PROMPTS was implemented as a toll-free text messaging platform to send staged messages and was combined with a clinician-supported helpdesk to answer questions. Messages were periodically adjusted based on maternal health clients' feedback and common questions. At the time of the study, the program had since expanded to four other counties (administrative locations in Kenya) had enrolled more than 25,000 women and answered more than 30,000 questions. To create awareness about the intervention, the PROMPTS implementers put up posters in the waiting bays at the respective facilities. Periodically, specific personnel who had been employed by the mHealth implementers visited the facilities to conduct information sessions where they explained to women how the intervention worked and helped those who were interested to register. Registration was done by sending a toll-free SMS with the word 'MIMBA' a Kiswahili word for pregnancy, to a specific short code.

2.1.2 Case study 2: the CCPF maternal mHealth intervention - Malawi

The Chipatala Cha Pa Foni (CCPF) intervention which translates to "Health Centre by Phone" in Chichewa was implemented to provide maternal health clients in Malawi with pregnancy tips and reminders. In addition, women could call a toll-free hotline for health information and advice (34). Like PROMPTS in Kenya, CCPF was open to pregnant women, and mothers of infants or children below five years of age who used CCPF while pregnant. At the time of the study, CCPF had enrolled more than 7,500 women and answered more than 2,000 calls per month. Maternal health clients registered for the intervention using their mobile phones, while those without mobile phones registered using mobile phones of family members, community members, or community volunteers. The intervention provided mobile phones to community volunteers, who served as agents of the intervention in their communities. The community volunteers trained the women who did not own mobile phones, on how to access the interventions using the project's mobile phones. At the time of the study, CCPF was available country-wide and owned by the Malawi Government. More details of the operations of CCPF can be found in (16).

2.2 Sampling and recruitment

We used purposeful sampling to recruit study participants. We sampled 40 respondents in Kenya and 31 respondents in Malawi. In Kenya, maternal health clients were recruited by visiting local health facilities, and in addition to participating in interviews, those who agreed to participate in the FGD were also enrolled

for the same. In Malawi, two CCPF volunteers were contacted to refer women who accessed CCPF via their phones. Consenting women were contacted mostly through their husbands phones since the women did not own any phones. Similar to Kenya, some women agreed to participate in both interviews and FGDs.

To recruit partners, the researchers reached out to the husbands of consenting women to ask if they were interested in participating in the study. Partners to women in Kenya participated in the FGD only, while in Malawi, partners were interviewed. Although we provide complete sample information for all study participants in Kenya and Malawi Tables A1, A2, we have focused the demographic information on the various individuals who constituted the CoP in the two mHealth contexts Tables A3, A4.

2.3 Data collection

Both studies adopted a variety of methods to collect data. However, for this paper, we used data from qualitative interviews and FGDs with various stakeholders in the maternal health context. All interviews lasted 45–60 min, and we audio-recorded them with permission from the participants. The FGDs with 5-8 participants lasted for 60–90 min. Interviews and discussions were conducted in the respective languages: Kiswahili and English in Kenya and Chichewa and English in Malawi. We piloted all data collection instruments with respondents similar to the participant sample and made relevant changes to enhance the clarity of the questions. We do not report on the pilot data in this paper.

2.3.1 Interviews

Researchers from Kenya and Malawi designed the interview questions for their respective countries. In Kenya, the interviews with maternal health clients centered around the following topics:

- their decision-making and considerations to use PROMPTS
- their healthcare-seeking behavior and practices before and after registering for PROMPTS
- what roles other community members played during pregnancy

In Malawi, the interviews with maternal health clients covered the following topics:

- maternal clients' motivation to use borrowed mobile phones to access the intervention
- cultural issues surrounding pregnancy especially when using a borrowed mobile phone
- the maternal clients' relationship with mobile phone owners and the other type of support the mobile phone owners offered to maternal clients
- what roles other community members played during pregnancy

The remaining stakeholders in Malawi, that is, community volunteers, members, and health officials, were engaged on the following issues: why they lend maternal clients their mobile phones to use the intervention, what else they did to support maternal clients when using their mobile phones, and how they influenced maternal clients to use the maternal mHealth intervention (Appendix A).

2.3.2 Focus group discussions

The partners of the maternal health clients in Kenya participated in a group discussion where the questions centered around the following themes (Appendix B): their perceptions and attitudes about women using the mHealth intervention, their beliefs and perceptions about pregnancy care and support for pregnant women, as well as their perceptions on the impact of the SMS intervention on women's lives.

The FGDs with the maternal health clients in Malawi centered on how they used the intervention. The FGDs in both cases were useful in gaining additional insight into culture, social norms, values, and power relationships with respect to maternal healthcare and pregnancy.

2.3.3 Data collection procedures

We sought informed consent before all interviews and FGDs that were conducted in a private space. For the FGDs, the researchers explained the limitations of group confidentiality, but encouraged participants to maintain this confidentiality beyond the meeting. We also assigned pseudonyms to all users to enhance their confidentiality. We were aware of the risk of interviewing pregnant women, that there could be an emergency during interviews and FGDs. Thus, in Malawi, we did not include pregnant women in our sample. However, in Kenya, where most of the maternal health participants were pregnant or early postpartum, the interviews were conducted in a private space at the health facility where the women visited for antenatal care (ANC) or at their homes if they preferred.

2.4 Data analysis

The researchers translated and transcribed any non-English data into English. We deidentified all data before analysis as an extra precaution in case participants' personal information had been captured in the course of the study. The transcripts were then uploaded to Nvivo 12 for analysis. We employed an inductive approach (35) to find patterns in the data (codes) following which we grouped these into higher-level themes and sub-themes (Table 1). Given that the original studies were completed at two different times, the respective researchers individually coded the data with frequent discussions with one other joint researcher to discuss emerging themes and to consider alternative interpretations. These peer discussions helped limit researcher bias. The final themes from this analysis process have been presented in the findings section with selected supporting excerpts of respondents' verbatim quotes. The deidentified verbatim excerpts are presented using the following pseudonym structure: "*CaseTypeofParticipant numeric number*," for example, PROMPTSClient1.

The two studies adopted Lincoln and Guba's (36) model of trustworthiness to ensure rigor. In addition to triangulating the data collection methods, we adopted peer debriefing with a

TABLE 1 Themes and their descriptions.

Main theme	Sub-theme	Description
Persuasion		The action of encouraging someone to do something
Persuasion factors	Obeying authority	Tendency people have to try to please those in charge
	Maintaining harmony	The act of avoiding fighting or arguing but rather living peacefully
	Peer influence	Doing something because your friend or other people in the community are doing it
Training		Teaching or developing other people's technical mobile phone usage skills
Technology access		Making a technological device such as a mobile phone available to someone
Sensitization		The process of letting someone know about an event or things that are happening in a community
Technology legitimization		The act of checking if the information is aligned with the Ministry of Health information of someone's country

senior researcher to review the inquiry process and cross-check inferences from the data analysis process. We also provide the study context and case description to ensure transferability to other similar contexts.

2.5 Ethics approvals

Both studies were conducted in one institution. Therefore, we obtained both institutional and national ethics approval before data collection. In Kenya, we obtained permission from the Amref Research Ethics Committee as well as from the National Council of Science and Technology (NACOSTI). In Malawi, we obtained permission from the Malawi Ministry of Health and Balaka District Health Office. Furthermore, we obtained ethical clearance from the National Health Sciences Research Committee (Malawi). We also sought permission from the intervention implementing agencies in both countries to use these interventions as case studies.

3 Results

We interviewed both maternal health clients and various members of the community that make up the maternal health CoP. These include other family members, community members, community volunteers, and health surveillance assistants (HSAs). Tables A3, A4 shows the demographic characteristics of CoP members in the two countries. Our analysis resulted in 5 major themes that highlight how the CoP influences maternal healthcare clients to use maternal mHealth interventions. The five mechanisms that we find are:

1. Persuading maternal clients to use the intervention
2. Training maternal clients on how to use digital health intervention
3. Provision of mobile phone access to maternal clients and
4. Sensitizing maternal clients about the interventions
5. Legitimizing the intervention by corroborating the health information

We further find that certain factors mediate the role of the CoP as illustrated in Figure 1. Persuasion was further mediated by factors that, in combination with trust between maternal clients and members of the CoP, played a role in influencing the use of mHealth.

3.1 Persuading maternal clients to use mHealth

Maternal clients in this study were convinced to register for mHealth interventions by other members of the community, such as HSAs, community volunteers, and their husbands. Although maternal clients in Malawi were persuaded to use the CCPF mHealth intervention by HSAs and community volunteers who visited them at their homes or met at social gatherings in their communities, the persuasion of women in Kenya happened through the implementation representatives of the intervention who visited the health facilities to provide information sessions.

> "We had to convince [maternal health clients] that the intervention was legitimate and that it is something approved by the county… their trust issue, they said, was that different people, or conmen or scammers send them messages [referring to general mobile phone scams] of which they don't know if they are true or not." [PROMPTSInformant 3]

Persuasion is a fundamental element in changing human behaviour and attitude. During persuasion, maternal clients were informed about the mHealth intervention and its benefits, which in turn led to a willingness to adopt.

> "I joined [CCPF] because of the advice I received from the HSA, that I can be helped while at home. Also, I can be listening to messages about my pregnancy…and when I deliver; I can also follow how my baby is growing" CCPFClient 13.

> PROMPTSClient 20 said: "What motivated me to register was that I might be in the house, and I encounter a particular challenge, and maybe I don't have money at that moment… to reach the hospital. [So], while I am sitting in my house, I can send a message telling them I'm experiencing this and the other."

The findings also indicate that before the CoP could encourage mothers to use mHealth interventions, they themselves had to be convinced of its perceived usefulness. For example, in the case of the mHealth intervention in Kenya, one partner referring to the wife said: "I was very happy because she told me that she knew about [the intervention] at the hospital and that made me comfortable because I knew that the hospital could not put up posters if the information was not genuine" PROMPTSMenFGD.

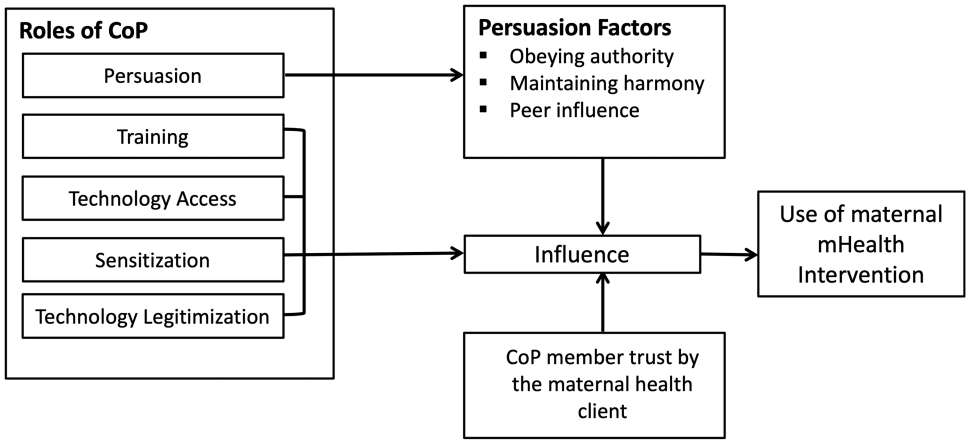


FIGURE 1
Community of Purpose framework for influencing mHealth intervention use.

Previous research (9, 37) has shown that when such health interventions are associated with a trusted entity, they engender a higher level of trust for both the CoP and the maternal health client and subsequently encourages adoption. One participant said "My husband asked me about these messages. I told him it's something like an online clinic through our phone, and that the doctors were from XYZ Hospital. And then he said it was okay." PROMPTSClient 10. In both Kenya and Malawi, the health facilities, healthcare providers and community volunteers represented trusted entities as seen in what one participant shared: "I remember that my husband was concerned [about me meeting you]. 'Where will you be meeting?' I told him that it would be at the hospital. And he said that was okay" [PROMPTSClient 15].

Trusting the intervention subsequently influenced the the actions of other stakeholders. A partner to one of the women explained how the mHealth intervention changed their perceptions and behavior. "The other thing is about nutrition, because maybe we had that old mentality about what pregnant women should eat, but now after interacting with this SMS service, you get to know other nutritional meals that pregnant women can eat" PROMPTSPartner2.

We identified three persuasion factors through which persuasion operates: (i) obeying authority, (ii) maintaining harmony, and (iii) peer influence.

3.1.1 Obeying authority

Our findings highlight that maternal health clients in both Kenya and Malawi easily deferred to figures in positions of authority. For example, since HSAs and healthcare providers were generally considered highly respected, maternal clients accepted and appropriated CCPF and PROMPTS when recommended by these higher figures. One key informant explained this better. "If a provider tells [the mother], you're good, go home, it's unfortunate that still very few people can oppose a provider or give their own opinion even if they feel it's not right…" [PROMPTSInformant 1]. The power wielded by healthcare providers was also captured by PROMPTSClient21 who said: "[My husband] used to tell me to go to the older woman and then I asked in the SMS and they told me that wasn't useful, the doctor has the ability··· [my husband] just agreed and said that the doctor is educated than the older women."

The male partners in the study corroborated the idea of submitting to authority. In addition to healthcare providers, other stakeholders, such as older women, also exercised authority. These women generally provided care and oversight of a woman's pregnancy, their advice was usually upheld. "We get [information] from those [older mothers] who have delivered and know more than you. Maybe they have other children and they have experienced this before and so they know what you don't know" PROMPTSClient13. The men in FGD in the Kenyan case said that when it came to pregnancy-related advice that was given by these older women, they "Did not [question] the reasons, but followed the [advice] because there could be a reason why and you don't want to go against what you have been warned" [PROMPTSMenFGD].

3.1.2 Maintaining harmony

Some maternal health clients were forced to use interventions because they wanted to maintain harmony with the respective members of the CoP. "My husband did the registration, so I didn't know what they were talking about. However, when the messages came, we were reading them together…" [CCPFClient 1]. Such compliance to maintain harmony may have been necessitated by the fact that most of the maternal clients depended on their partners for financial support. The women in the FGD group in Kenya said: "They [husbands] are the heads of the family, and we need financial assistance, so you inform them so that they can give you money."

"We're finding that the [husbands] are the ones who are ensuring that the woman is eating the right diet during pregnancy because most of the times husbands are the providers, especially in our setting… where most women stay at home, and

the husbands go to work. So, they're basically the providers" [PROMPTSInformant 1].

3.1.3 Peer influence

Maternal clients were convinced to use CCPF and PROMPTS because other women were using it. One participant shared that she saw other maternal healthcare clients using CCPF and heard about other maternal clients' experiences with CCPF which was good. This is similar to findings of other studies who found that mentors and role models in communities have the potential to persuade other users to accept and use the mHealth intervention (38).

3.2 Training maternal clients on how to use the mHealth intervention

During the feasibility study before CCPF was implemented, the project implementers found that the women in Malawi had lower literacy than their male counterparts (39). Hence, community members were trained as community volunteers so that they could provide training to maternal clients on how to use the intervention. Training maternal clients on how to use the mHealth intervention promoted social learning and social acceptability of the intervention among members of the community. With the lack of digital skills likely to disproportionately affect users without technologies such as mobile phones, such community-enabled digital skills training was central to mHealth adoption and use. Thus, training these maternal clients on how to use CCPF on a mobile phone was the first step to influence the use of the mHealth intervention. Otherwise, the non-usage of mHealth interventions could lead to their failure to achieve their purpose (40), which is to improve the maternal healthcare-seeking bahavior.

3.3 Technology access

Maternal clients who stand to benefit from mHealth could be disproportionately disadvantaged by the lack of mobile phones. Access to a mobile phone is an important prerequisite for the success of mHealth use. Although maternal clients in Malawi did not own mobile phones, they accessed a mobile phone through family circles such as their husbands and mothers-in-law, as well as community volunteers and community members. "For CCPF, I use the community member's mobile phone. I dial the CCPF number myself and I can talk to the hotline worker. I can talk privately on the mobile phone. I am used to this mobile phone for CCPF because this mobile phone is always available to me. . ." [CCPFClient 15].

A community volunteer mentioned that her mobile phone was available for maternal clients to use for CCPF and she even visited registered and unregistered maternal clients to see if maternal clients needed to use her mobile phone to access CCPF. Another volunteer mentioned that after the mobile phone that she was given during the pilot phase of the intervention malfunctioned, she bought another phone and continued her role as a community volunteer in her village. "The mothers live far from where I live, so I just choose a weekday that I can visit the mothers. . ." [CCPFCommunity Volunteer 1].

One CCPF community member, a wife of a village headman, said she allowed maternal clients to use her mobile phone to access the intervention. In the rural context of Malawi, a village headman presided over various aspects of life, especially cultural and social issues in the community (41). Village headmen were also expected to advise and support their community, as well as to implement policies handed down from the national level (42).

3.4 Sensitizing maternal clients

In the case of CCPF, community volunteers and HSAs were mandated to sensitize maternal healthcare clients about the maternal mHealth intervention in their communities. Community volunteers paid special attention to maternal clients who did not own a mobile phone to alert them about CCPF. "The health counselor in the village told us that there is CCPF, and you can call them anytime to ask about any pregnancy-related problems. . ." [CCPFClient 8]. Furthermore, husbands of maternal clients attested that when they received an SMS about CCPF, they informed their wives about it. In this study, male participation in maternal-related issues was more prevalent, especially in sensitizing maternal clients about CCPF and persuading maternal clients to use CCPF.

In the case of PROMPTS, the sensitization was performed by PROMPTS personnel who visited the specified clinics on certain days of the week to explain the intervention to the waiting ANC clients while encouraging them to enroll. In this case, these personnel also helped reduce any uncertainties by explaining to the women what the intervention was about and explaining that it was free to use.

3.5 Technology legitimization

The CoP facilitated the legitimization of the mHealth intervention by offering a means for maternal health clients to corroborate the information they received from both PROMPTS and CCPF. Given the diverse uncertainties around pregnancy emanating from the cultural context, maternal healthcare clients established the credibility of pregnancy-related information to establish trust. They achieved this by comparing the mHealth-related information with the advice of trusted stakeholders, including trusted older female relatives and sometimes healthcare providers, when they attended ANC clinics. The legitimization itself was achieved by: (i) Comparing new information with existing knowledge and (ii) Comparing information across multiple sources.

Legitimacy is defined as "a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs and definitions" (43). The women determined the reliability of the information against a priori knowledge. "At times there are some questions you sort of already have the answers to, but you just ask to make a comparison" [PROMPTSClient 27]. When one

trustworthy source was not available, the women established the credibility of the information by comparing across multiple sources: "I had two sources of information, the SMS and the doctors at the clinic. After receiving the messages I would verify some of that information by asking the doctor the same questions, which led to total trust because I saw that the information was accurate" [PROMPTSClient 10].

Other researchers (44, 45) have noted similar behavior of health clients engaging multiple sources of information and care-seeking to address health concerns.

3.6 How does the role of the CoP compare between settings where users have mobile phones compared to where they do not?

Our findings show that the CoP plays an important role both when users have a mobile phone and when they do not. However, the exact nature of their role is sometimes dependent on the unique circumstances of the maternal clients. For example, in Malawi where mHealth users did not have phones, in addition to providing technology access, the CoP sensitized women on the existence of the intervention and acted as important influencers to adoption and use. Thus, the CoP played a more direct role. In Kenya, where women already had phones, the CoP had a more indirect influence. For example, by providing opportunities for women to corroborate the information they received from the mHealth platform, CoP acted as indirect facilitators of adoption. We therefore observe that although the specific mechanisms of the CoP's influence varies between contexts, their centrality within the mHealth appropriation and use domain remains the same.

4 Discussion

The research community has called on the need for studies to provide a detailed understanding of how social capital can be manipulated to influence health results (46). We evaluate the role of the community of purpose as a vehicle through which social capital can manifest. The findings of this study suggest that CoP played different roles in influencing maternal clients to use maternal mHealth interventions. Collectively, we draw three important lessons from our findings: (1) The extent of CoP influence is inextricably linked to the cultural context, (2) The involvement of the CoP catalyzes health behavior change by providing multiple layers of influence, and (3) While the CoP as a whole is central to maternal health, the respective constituents do not wield equal power.

4.1 Lesson 1: the degree of influence of CoP is inextricably linked to the cultural context

In collectivist cultures, the community is valued over the individual. "Umuntu ngumuntu ngabantu" which loosely translates as "[a] person is a person bacause of other people" (47) gives priority to empathy, caring and understanding, and the contributions of individuals in a community are valued and cherished (31). Furthermore, sharing and neighborly assistance are part of an African identity (48), and are attributed to the Ubuntu philosophy practiced in the communities we studied. Although project implementers in Malawi encouraged maternal clients to use their family member's mobile phones, as well as friends' mobile phones, this practice was acceptable due to the Ubuntu value system. These contributed to establishing an inclusive community that helped promote the participation of all members of the community and create a sense of belonging (49, 50). Inclusive communities acknowledge that all members of the community, including the marginalized, should participate in everything that happens in their community (51). The most important thing in an inclusive community is to recognize that people have different needs and that diversity should be valued (51). Therefore, inclusive communities played a vital role in identifying and removing barriers to community participation. When communities are inclusive, sharing of resources, such as mobile phones, helps to reduce the digital divide gap (52), thus making mHealth interventions work in resource-constrained communities.

In addition to the Ubuntu value system that facilitated certain sharing-based technology outcomes, the persuasion factors are rooted in culture. Obedience to authority, seeking harmony, and peer influence are dependent on aspects of culture such as power distance and norms in collectivist cultures. Hence, we do not expect that the same influence factors will work in the same way in, say, more individualistic cultures.

4.2 Lesson 2: the CoP acts as agents of change in the context of maternal mHealth interventions

By influencing the use of mHealth among the women, offering training, increasing the accessibility and reach of mHealth and facilitating the legitimization of mHealth for continued use, the CoP acts as agents of change. It has been previously noted that the low literacy of maternal healthcare clients inhibits maternal mHealth intervention designed for poor-resource settings (53, 54). ICT capabilities are essential for the uptake of many digital technologies including mHealth interventions. In this study, to promote ICT usage skills, community volunteers trained maternal healthcare clients on how to access CCPF using a mobile phone. By providing the technology by which women could have access to maternal health information, the CoP played an important role of being infomediaries.

Maternal mHealth interventions also translate to desired health outcomes only as health-related advice is followed. In Malawi, CCPF community volunteers visited maternal clients in their homes to inform and encourage its use. In Kenya, once legitimized and women's use of the intervention became a culturally appropriate behavior, other members of the CoP were able to facilitate follow-through with information that would previously be impossible, such as those related to food taboos that husbands would not purchase for their wives. Other studies have similarly shown that

various actors play an important role as agents of behavior change (55, 56). Community members are particularly effective because they live within the proximity of maternal clients. This means that they can more easily monitor the behaviors of their maternal healthcare clients to ensure that they are following the information appropriately. Other studies have called these CoP members 'watchdog-oriented community members' because they play an important supervisory function in society by monitoring the health of consumers in their communities (57).

Similarly, initial trust is necessary for the adoption of maternal mHealth services, especially given that the maternal healthcare-seeking context contributes to the increased liability for the newness of new interventions which negatively impacts adoption (24, 37). Various design and implementation characteristics that help minimize perceived risks and uncertainties about using a new mHealth intervention can help to engender initial trust (37). By providing an avenue through which women legitimize the intervention and therefore overcome uncertainties related to the newness of the intervention, the CoP acted as agents of change to increase adoption.

4.3 Lesson 3: although CoP as a whole is central to maternal health, the respective constituents do not wield equal power

The findings confirm that the CoP is central to the experience of using maternal mHealth interventions. However, we find that CoP members play different roles, and we posit that some of these differences arise from their respective social capital as individual constituents. For example, health care workers wield more power as "gate-keepers" of medical care based on their expertise, training, and education. However, older women provide emotional and domestic support during pregnancy, and the power they exercise is based on age and perceived experience with pregnancy-related issues. Finally, other stakeholders like partners are often considered the main heads of their households, decision makers, and financial controllers. Some of these roles have been highlighted before in literature on maternal health, but our study illustrates that these different aspects come together to form a complex web of dependencies that influence mHealth usage. In a sense, we show that these aspects of power and social structures also translate to the adoption and use of digital health technology.

4.4 Study limitations

Although we had planned to engage a diverse group of other members of the community closely involved in woman's pregnancy in Kenya such as their female relatives: mothers, grandmothers and mothers-in-law, this was not possible because most women in the urban areas lived away from their next of kin most of whom were in their rural places of origin. We believe that a broader range of insights from these key stakeholders would be valuable in enriching the findings of this study. The context of pregnancy may also be nuanced when compared to other health domains, and therefore we claim no generalizability of the findings to other areas where change in health behavior may be a desired outcome. Other limitations relate to the inherent nature of qualitative research. First, given the small sample size, our results may not be generalizable to entire populations of maternal mHealth users. Second, some nuances may also have been lost in translation of the interviews between the languages, which we mitigated by having the researchers from Malawi and Kenya do the transcription to English. Third, the self-reported nature of the interviews may add additional biases. However, we believe that the participation of multiple participants facilitates a better triangulation and validity of the findings.

5 Recommendations and future directions

Our study offers an in-depth understanding of the centrality of other community members in the context of a woman's pregnancy (which we have referred to as Community of Purpose -CoP) in the adoption and use of digital health technologies among maternal health clients. The cases we have drawn from are maternal health interventions that were implemented in Kenya and Malawi among low socioeconomic women and among women without phones, respectively. Our study is qualitative and we do not make any claims on generalizability of the findings. However, we hypothesize that the insights generated may be relevant to other African countries that share similar contexts both economically and culturally.

We offer the following two recommendations.

5.1 Recommendation 1: design and implementation of mHealth and other similar digital health technologies with the different roles of the CoP in mind

The study underscores the need for implementers and developers of such technology to consider more holistically the various relationships and the different roles they play and design and implement technology in a manner that will maximize on outcomes by leveraging their respective roles and minimize unnecessary adoption and usage challenges. For example, in the CCPF, the recruitment and training of community volunteers proved to be invaluable. Although this may have been necessary due to the unique context of mHealth users without phones, we posit that this would be equally useful in contexts where users have phones. The findings of both cases demonstrate that associating mHealth interventions with an entity that the community already trusts (for example, hospitals and healthcare workers) contributes to the initial legitimization of mHealth interventions, which facilitates the adoption of mHealth. Since the CoP is also central to legitimizing information by offering a means for maternal health client to corroborate the information they receive, involving them in the design and implementation of digital health interventions will

prove extremely useful in contexts such as those presented in this study. Although the involvement of healthcare providers in the implementation of mHealth can be beneficial in encouraging adoption, it has potential drawbacks. Patient dissatisfaction with providers involved in health interventions could negatively affect their use. Possible solutions involve managing the intervention's operations independently, but involving healthcare workers to create awareness. This will also ensure that providers in resource-constrained settings do not feel burdened by the additional processes of an intervention, and allow patients who prefer the anonymity of mHealth interventions to leverage on such affordances. Involving other CoP members, such as volunteers for women without phones, would also have to think about how to increase motivation. This might entail some form of compensation for their time as seen in Larsen-Cooper et al., (58). Given that our study was limited to the domain of maternal health, we encourage other researchers to explore the role of the CoP in other health domains and for various health outcomes.

5.2 Recommendation 2: conduct further research to establish when group targeting of interventions is beneficial to individual targeting

Given that many maternal healthcare-seeking behaviors occur within the purview of social and cultural norms within the society, interventions at the relevant community level might yield more significant outcomes than individual-focused interventions. Our research supports the idea of group (CoP) targeted interventions in maternal health in Africa. More empirical research would be useful in establishing this. Future research could contribute to building a tool that can be used for context assessment and as a guide to know when and where group or individual targeting would be more desirable.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Amref ESRC, (Kenya) National Health Sciences Research Committee (Malawi). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

KS: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing. PM: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing. WC: Supervision, Validation, Writing – review & editing. AM: Supervision, Writing – review & editing.

Funding

The authors declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Katusiime J, Pinkwart N. Supporting maternal health education in developing countries using mobile phones-results of a pilot study. In: *Proceedings of the First African Conference on Human Computer Interaction* (2016). p. 48–57.

2. Jones M. *How much has maternal health improved?*. Tech. rep., World Economic Forum (2015).

3. WHO. *Maternal morality: key facts*. Tech. rep., World Health Organization (2023).

4. WHO. *World health statistics 2018 – monitoring health for the sdgs, sustainable development goals*. Tech. rep., World Health Organization (2018).

5. Akter S, Ray P. mhealth – an ultimate platform to serve the unserved. *Yearb Med Inform*. (2010) 19:94–100. doi: 10.1055/s-0038-1638697

6. Chen H, Chai Y, Dong L, Niu W, Zhang P. Effectiveness and appropriateness of mhealth interventions for maternal and child health: systematic review. *JMIR mHealth uHealth*. (2018) 6:e8998. doi: 10.2196/mhealth.8998

7. Poorman E, Gazmararian J, Parker RM, Yang B, Elon L. Use of text messaging for maternal and infant health: a systematic review of the literature. *Matern Child Health J*. (2015) 19:969–89. doi: 10.1007/s10995-014-1595-8

8. Tamrat T, Kachnowski S. Special delivery: an analysis of mhealth in maternal and newborn health programs and their outcomes around the world. *Matern Child Health J*. (2012) 16:1092–101. doi: 10.1007/s10995-011-0836-3

9. Sowon K, Maliwichi P, Chigona W. The influence of design and implementation characteristics on the use of maternal mobile health interventions in kenya: systematic literature review. *JMIR mHealth uHealth*. (2022) 10:e22093. doi: 10.2196/22093

10. GSMA. *The mobile economy Sub-Saharan Africa 2023*. Tech. rep., GSMA (2023).

11. Maliwichi P, Mthoko H, Chigona W, Mburu C, Densmore M. Does mobile phone ownership matter? insights on engagement in mhealth and e-government

interventions from Southern Africa. *Int J Health Promot Educ*. (2021) 1–22. doi: 10.1080/14635240.2021.1995775

12. Naughton F, Ward E, Khondoker M, Belderson P, Minihane AM, Dainty J, et al. Health behaviour change during the uk covid-19 lockdown: findings from the first wave of the c-19 health behaviour and well-being daily tracker study. *Health Psychol*. (2021) 26:624–43. doi: 10.1111/bjhp.12500

13. McCarthy H, Potts HW, Fisher A. Physical activity behavior before, during, and after covid-19 restrictions: longitudinal smartphone-tracking study of adults in the United Kingdom. *J Med Internet Res*. (2021) 23:e23701. doi: 10.2196/23701

14. NSO. *Survey on access and use of information and communications technologies by households and individuals in Malawi 2019*. Tech. rep., NSO (2020).

15. Blauvelt C, West M, Maxim L, Kasiya A, Dambula I, Kachila U, et al. Scaling up a health and nutrition hotline in Malawi: the benefits of multisectoral collaboration. *BMJ*. (2018) 363:k4590. doi: 10.1136/bmj.k4590

16. Maliwichi P, Chigona W. Towards a framework on the use of infomediaries in maternal mHealth in rural Malawi. *Interdiscip J Inf Knowl Manage*. (2022) 17:387–411. doi: 10.28945/5015

17. Kaiser JL, Fong RM, Hamer DH, Biemba G, Ngoma T, Tusing B, et al. How a woman's interpersonal relationships can delay care-seeking and access during the maternity period in rural zambia: an intersection of the social ecological model with the three delays framework. *Social Sci Med*. (2019) 220:312–21. doi: 10.1016/j.socscimed.2018.11.011

18. Ensor T, Cooper S. Overcoming barriers to health service access: influencing the demand side. *Health Policy Plan*. (2004) 19:69–79. doi: 10.1093/heapol/czh009

19. Mumtaz Z, Salway SM. Gender, pregnancy and the uptake of antenatal care services in Pakistan. *Sociol Health Illness*. (2007) 29:1–26. doi: 10.1111/j.1467-9566.2007.00519.x

20. Serizawa A, Ito K, Algaddal A, Eltaybe R. Cultural perceptions and health behaviors related to safe motherhood among village women in eastern sudan: Ethnographic study. *Int J Nurs Stud*. (2014) 51:572–81. doi: 10.1016/j.ijnurstu.2013.08.007

21. Shaikh BT, Hatcher J. Health seeking behaviour and health service utilization in Pakistan: challenging the policy makers. *J Public Health (Bangkok)*. (2005) 27:49–54. doi: 10.1093/pubmed/fdh207

22. Wenger E. *Communities of practice: a brief introduction*. Tech. rep., University of Oregon (2011)

23. Bhattacharyya S, Basak J, Bhaumik P, Bandyopadhyay S. Cultivating online virtual community of purpose to mitigate knowledge asymmetry and market separation of rural artisans in India. In: *Evolving Perspectives on ICTs in Global Souths: 11th International Development Informatics Association Conference, IDIA 2020, Macau, China, March 25–27, 2020, Proceedings 11*. Springer (2020). p. 201–15.

24. Cheruto SK. *How the healthcare-seeking socio-cultural context shapes maternal health clients' mHealth utilisation in a Kenyan context* (PhD thesis). Faculty of Commerce, Department of Information Systems. (2021). http://hdl.handle.net/11427/34016

25. Mahmood K. Social capital: from concept to theory. *Pak J Sci*. (2015) 67(1):114–9. doi: 10.57041/pjs.v67i1.378

26. Lagaert S, Snaphaan T, Vyncke V, Hardyns W, Lieven PJR, Willems S. A multilevel perspective on the health effect of social capital: evidence for the relative importance of individual social capital over neighborhood social capital. *Int J Environ Res Public Health*. (2021) 18:1526. doi: 10.3390/ijerph18041526

27. Shiell A, Hawe P, Kavanagh S. Evidence suggests a need to rethink social capital and social capital interventions. *Social Sci Med*. (2020) 257:111930. doi: 10.1016/j.socscimed.2018.09.006

28. Yuan M, Yue-qun C, Hao W, Hong X. Does social capital promote health? *Soc Indic Res*. (2022) 162:501–24. doi: 10.1007/s11205-021-02810-8

29. Nagy E, Moore S, Silveira PP, Meaney MJ, Levitan RD, Dubé L. Low socioeconomic status, parental stress, depression, and the buffering role of network social capital in mothers. *J Ment Health*. (2022) 31:340–7. doi: 10.1080/09638237.2020.1793118

30. Kemp S. *Digital 2023*. Tech. rep., DataReportal (2023).

31. Zamani ED, Sbaffi L. *In Information and Communication Technologies for Development: 16th IFIP WG 9.4 International Conference on Social Implications of Computers in Developing Countries, ICT4D 2020, Manchester, UK, June 10–11*. Springer (2020). Vol. 16. p. 93–101.

32. Duclos V, Yé M, Moubassira K, Sanou H, Sawadogo NH, Bibeau G, et al. Situating mobile health: a qualitative study of mhealth expectations in the rural health district of Nouna, Burkina Faso. *Health Res Policy Syst*. (2017) 15:41–53. doi: 10.1186/s12961-017-0211-y

33. McConnell M, Ettenger A, Rothschild CW, Muigai F, Cohen J. Can a community health worker administered postnatal checklist increase health-seeking behaviors and knowledge?: evidence from a randomized trial with a private maternity facility in Kiambu County, Kenya. *BMC Pregnancy Childbirth*. (2016) 16:1–19. doi: 10.1186/s12884-016-0914-z

34. VillageReach. *A national mobile health hotline in Malawi: one step closer*. Tech. rep., VillageReach (2017).

35. Clarke V, Braun V. Thematic analysis. *J Posit Psychol*. (2017) 12:297–8. doi: 10.1080/17439760.2016.1262613

36. Lincoln Y, Guba E. *Naturalistic Inquiry*. London: Sage Publications (1985).

37. Sowon K, Chigona W. Trust in mHealth: how do maternal health clients accept and use mhealth interventions? In: *Conference of the South African Institute of Computer Scientists and Information Technologists 2020* (2020). p. 189–97.

38. Besral B, Misrawati M, Afiyanti Y, Ismail RI, Arifin H. Miesra mHealth: marital satisfaction during pregnancy. *PLoS One*. (2023) 18:e0289061. doi: 10.1371/journal.pone.0289061

39. Nyemba-Mudenda M, Chigona W. mHealth outcomes for pregnant mothers in Malawi: a capability perspective. *Inf Technol Dev*. (2018) 24:245–78. doi: 10.1080/02681102.2017.1397594

40. Khatun F, Heywood AE, Ray PK, Bhuiya A, Liaw ST. Community readiness for adopting mhealth in rural bangladesh: a qualitative exploration. *Int J Med Inform*. (2016) 93:49–56. doi: 10.1016/j.ijmedinf.2016.05.010

41. Gondwe DK. The role of traditional leaders in rural development, a case study of Karonga district in Malawi. *Int J Sci Res*. (2020) 9:1584–9. doi: 10.21275/SR20423141736

42. Walsh A, Matthews A, Manda-Taylor L, Brugha R, Mwale D, Phiri T, et al. The role of the traditional leader in implementing maternal, newborn and child health policy in Malawi. *Health Policy Plan*. (2018) 33:879–87. doi: 10.1093/heapol/czy059

43. Suchman MC. Managing legitimacy: strategic and institutional approaches. *Acad Manage Rev*. (1995) 20:571–610. doi: 10.2307/258788

44. de Graft Aikins A. Food beliefs and practices during pregnancy in ghana: Implications for maternal health interventions. *Health Care Women Int*. (2014) 35:954–72. doi: 10.1080/07399332.2014.926902

45. Graner S, Klingberg-Allvin M, Duong LQ, Krantz G, Mogren I. Pregnant women's perception on signs and symptoms during pregnancy and maternal health care in a rural low-resource setting. *Acta Obstet Gynecol Scand*. (2013) 92:1094–100. doi: 10.1111/aogs.12170

46. Villalonga-Olives E, Wind T, Kawachi I. Social capital interventions in public health: a systematic review. *Social Sci Med*. (2018) 212:203–18. doi: 10.1016/j.socscimed.2018.07.022

47. Bandawe CR. Psychology brewed in an african pot: indigenous philosophies and the quest for relevance. *Higher Educ Policy*. (2005) 18:289–300. doi: 10.1057/palgrave.hep.8300091

48. Van Dyk GA, Matoane M. Ubuntu-oriented therapy: prospects for counseling families affected with HIV/AIDS in Sub-Saharan Africa. *J Psychol Afr*. (2010) 20:327–34. doi: 10.1080/14330237.2010.10820382

49. Mohd NM, Mamat M, Mahamood S, Mohd NM. Asean as an inclusive community: identifying and overcoming the challenges. In: *Charting a Sustainable Future of ASEAN in Business and Social Sciences* (2020). p. 215–24.

50. Scholtz B, Jonamu B, Calitz AP. Information management in an African higher education institution a gap model for environmental. In: *Proceedings of the 8th IDIA International Development Informatics Association Conference, Port Elizabeth, South Africa* (2014). p. 3–4.

51. Mpofu J, Sefotho MM, Maree JG. Psychological well-being of adolescents with physical disabilities in zimbabwean inclusive community settings: an exploratory study the Zimbabwe situation. *Afr J Disabil*. (2017) 6:1–7. doi: 10.4102/ajod.v6i0.325

52. Bertot JC. Building digitally inclusive communities: the roles of public libraries in digital inclusion and development. In: *ICEGOV'15-16: Proceedings of the 9th International Conference on Theory and Practice of Electronic Governance* (2016). p. 95–102.

53. Dasuki SI, Zamani ED. Assessing mobile phone use by pregnant women in nigeria: a capability perspective. *Electron J Inf Syst Dev Ctries*. (2019) 85:e12092. doi: 10.1002/isd2.12092

54. Maliwichi P, Chigona W, Sowon K. Appropriation of mhealth interventions for maternal healthcare in Sub-Saharan Africa: hermeneutic review. *JMIR mHealth uHealth*. (2021) 9:e22653. doi: 10.2196/22653

55. Lund S, Nielsen BB, Hemed M, Boas IM, Said A, Said K, et al. Mobile phones improve antenatal care attendance in Zanzibar: a cluster randomized controlled trial. *BMC Pregnancy Childbirth*. (2014) 14:1–10. doi: 10.1186/1471-2393-14-29

56. Ngabo F, Nguimfack J, Nwaigwe F, Mugeni C, Muhoza D, Wilson DR, et al. Designing and implementing an innovative SMS-based alert system (rapidSMS-MCH) to monitor pregnancy and reduce maternal and child deaths in Rwanda. *Pan Afr Med J*. (2012) 13:1–15. PMID: 23330022; PMCID: PMC3542808

57. Frig M, Fougère M, Liljander V, Polsa P. Business infomediary representations of corporate responsibility. *J Bus Ethics*. (2018) 151:337–51. doi: 10.1007/s10551-016-3200-0

58. Larsen-Cooper E, Bancroft E, O'Toole M, Jezman Z. Where there is no phone: the benefits and limitations of using intermediaries to extend the reach of mhealth to individuals without personal phones in Malawi. *Afr Popul Stud*. (2015) 29:1628–42. doi: 10.11564/29-1-714

Appendix A. FGD interview questions for the CoP in Kenya

The questions below will be used to engage the SOs to gain further insights on power dynamics and decision making process in the mHealth environment that affect utilization of mHealth interventions

Describe your understanding of the pregnancy experience:

- What is the definition of a good/successful pregnancy experience?
- What rules and guidelines exist in your community/family/society to make this goal possible?

Please share about the traditional methods/your previous ideas of pregnancy-related support

- Before learning from the SMSs or from the health facility, what were your ideas on:
  - the role of the healthcare provider during pregnancy
  - when a pregnant woman should start seeing the healthcare provider
  - how many visits to the doctor during pregnancy?
  - the procedure for having her pregnancy related questions answered: who can she talk to and what informs who she can talk to?
- When in the pregnancy are these people involved? (pregnancy stage or circumstances necessitating involvement)
  - Among the people who can be involved, who is responsible for what?
- Who is responsible for the decision of when, where and how a pregnant woman seeks care and support? (if not husband, what is the role of the husband in this decision making process?)
- What is the decision making role/authority based on?

How else can a mother within your society achieve a successful pregnancy?

- What is the role of other alternative sources of care like TBAs in the pregnancy care system?
- What do think about your partner following all that the intervention prescribes? Are there things you felt it necessary to seek a second opinion to confirm what you received from the SMSs?
- Please describe what you see as your role in the decisions regarding following/adhering to what the intervention though the SMSs would prescribe.

What challenges have you experienced with the used of the SMS service for pregnancy support? (Any conflicts in the nature of information, the asking/answering of questions etc.)

- How agreeable are you to her following everything the intervention says?
- Please describe your process of resolving conflict if what the SMS says is different from what you knew before.

How do you feel the SMS program has changed the healthcare seeking experience for you and your partner?

- Did the use of the SMSs by your partners change your place and role in the healthcare seeking decision making process? How? How did you feel about this?
- How did the SMS program change your involvement and engagement with the pregnancy?
- How did the SMSs change your interaction with the healthcare providers?

How else would you say the SMSs have transformed the pregnancy experience for you as a partner?

Appendix B. Semi-structured interview for the CoP (mobile phone owners) in Malawi

The questions below will be used as a guide for the interview with mobile phone owners (CHWs, family members, friend). The question may be rephrased and probed in various ways.

Relationship with the maternal healthcare client

1. What is the relationship with the maternal healthcare client?
   a. Client
   b. Mother
   c. Wife
   d. Friend
   e. Other
2. Besides the intervention, what other purposes do you allow people to use your phone?
3. How do you decide which people to use your phone?
4. How long is the maternal healthcare client being using your phone?
5. What other roles did you do to make sure that mothers use the intervention?

TABLE A1 Study sample information for PROMPTS intervention.

	Women	Significant other	mHealth official	Healthcare providers
Population and sample size (n)	30	5	3	2
Study procedures	Interviews and FGDs	FGD	Interviews	Informal discussion

TABLE A2 Study sample information for CCPF intervention.

	Women	Significant other	mHealth official	Hotline worker and community volunteer
Population and sample size (n)	20	6	2	3
Study procedures	Interviews and FGDs	Interviews	Interviews	Interviews

TABLE A3 CCPF CoP demographics.

Mobile phone owner	Age	Gender	No of maternal clients who used the phone
Husband 1	39	M	1
Husband 2	20	M	1
Husband 3	37	M	1
Husband 4	37	M	1
Mother-in-law 1	35	F	1
Community volunteer 1	42	F	8
Community member 1	43	F	7
Officer 1–3	–	M	–

TABLE A4 PROMPTS CoP demographics.

Respondent	Partner to	Age	Education	Work status
Partner 1	Mother 20	27	Primary	Casual laborer
Partner 2	Mother 24	29	Primary	Casual laborer
Partner 3	Mother 17	35	Tertiary	Teacher
Partner 4	Mother 12	30	Secondary	Self employed
Partner 5	Mother 9	28	Secondary	Casual laborer
mHealth officials 1–3	–	–	–	–
Healthcare providers 1–2	–	–	–	–

Motivation to let the maternal healthcare client their mobile phone

6. Why do you allow maternal healthcare client to use your phone?
7. Describe the process how maternal healthcare client asks to use your phone?
8. Describe the process how someone asks to use the phone for normal use?

Benefits and challenges

9. Why do you think other people do not allow maternal healthcare clients to use their phones?
10. What do you get in return when maternal healthcare clients are using your phone?
11. What challenges do you face when maternal healthcare clients are using your phone?

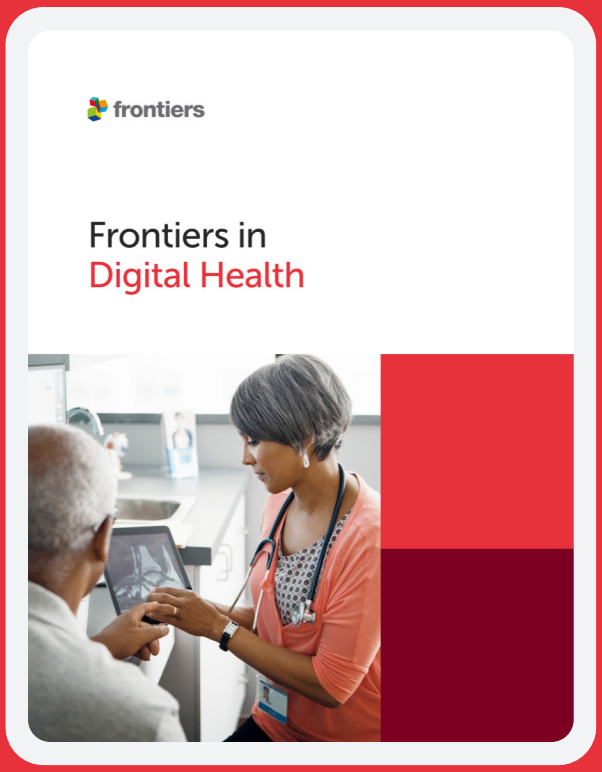
Frontiers in
Digital Health

Explores digital innovation to transform modern healthcare

A multidisciplinary journal that focuses on how we can transform healthcare with innovative digital tools. It provides a forum for an era of health service marked by increased prediction and prevention.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact