Virtual Support for Real-World Movement

Using Chatbots to Overcome Barriers to Physical Activity

Sun, X.; Casula, D.; Navaratnam, A.; Popp, A.; Knopp, F.; Busini, G.; Wesołowski, J.; van Reeth, M.; Reich, E.; Wiers, R.; Bosch, J.A.

DOI
10.3233/FAIA230084

Publication date
2023

Document Version
Final published version

Published in
HHAI 2023: Augmenting Human Intellect

HHAI 2023: Augmenting Human Intellect
P. Lukowicz et al. (Eds.)

doi:10.3233/FAIA230084

Virtual Support for Real-World Movement: Using Chatbots to Overcome Barriers to Physical Activity

Xin Sun[a,1] and Dimosthenis Casula[a], Arathy Navaratnam[a], Anna Popp[a], Franziska Knopp[a], Giovanni Busini[a], Jan Wesołowski[a], Marie van Reeth[a], Elke Reich[a], Reinout Wiers[a], Jos A. Bosch[a]
[a]University of Amsterdam

Abstract. Conversational agents (CAs, aka chatbots) for behavioral interventions have great potential to improve patient engagement and provide solutions that can benefit human health. In this study, we examined the potential efficacy of chatbots in assisting with the resolution of specific barriers that people frequently encounter when doing behavioral interventions for the purpose of increasing physical activity (PA). To do this, six common barriers (i.e., things that stand in the way of increasing PA) were targeted (e.g., stress and fatigue), we adopted domain knowledge (i.e., psychological theories and behavioral change techniques) to design six interventions aimed at tackling each of these six barriers. These interventions were then incorporated into consultative conversations, which were subsequently integrated into a chatbot. A user study was conducted on non-clinical samples (n=77) where all participants were presented with three randomly but equally distributed chatbot interventions and a control condition. Each intervention conversation addressed a specific barrier to PA, while the control conversation did not address any barrier. The outcome variables were beliefs in PA engagement, attitudes toward the effectiveness of each intervention to resolve the barrier, and the overall chatbot experience. The results showed a significant increase in beliefs of PA engagement in most intervention groups compared to the control group, and positive attitudes toward the effectiveness of the interventions in reducing their respective barriers to PA, and positive chatbot experience. The results demonstrate that theory-grounded interventions delivered by chatbots can effectively help people overcome specific barriers to PA, thereby increasing their beliefs in PA engagement. These promising findings indicate that chatbot interventions can be an accessible and widely applicable solution for a larger population to promote PA.

Keywords. Domain knowledge and evidence grounded chatbot, behavioral intervention, barriers to physical activity, text classification

1. Introduction

Regular physical activity (PA) is essential for maintaining health and preventing diseases [28][29]. Despite numerous benefits of PA (e.g., improving both physical and mental health), many people struggle to engage in regular PA. Various barriers (e.g., lack of time and motivation) have been identified as reasons for low PA levels. Nowadays, chatbot-delivered interventions are mostly used in the context of health, such as mental

[1] Corresponding Author: Xin Sun, x.sun2@uva.nl.

health and chronic disease management. Although the use of chatbots in PA promotion is a promising area of research [6][8], there is limited evidence on their effectiveness in addressing perceived barriers to PA in individuals. The current study aims to fill this research gap by carrying out a user-centered study with a theory-grounded chatbot to evaluate the effectiveness of chatbot-delivered intervention in promoting PA.

The objective of this work is to develop a chatbot to investigate its effectiveness in overcoming human perceived barriers to PA. To achieve this, a focus group was conducted in living lab [37] to investigate the barriers people encounter when trying to do PA. The evidence-based results of this focus group found six common barriers to PA: stress, fatigue, intrinsic motivation, time management, habit formation, and unsupportive social environment. These identified barriers represented the most common obstacles preventing people from engaging in PA, thus addressing them can maximize intervention effectiveness for more individuals. In the present study, interventions were developed as consultative conversations which were pre-scripted and grounded by domain knowledge (e.g., motivational interviewing (MI) [9] and graded exercise therapy (GET) [15]) to resolve these six barriers, and these conversations were then integrated into a chatbot. Therefore, the research questions of present study are as follow: Whether chatbot-delivered interventions can effectively help address specific barriers to PA? **(RQ1)** Whether chatbot-delivered interventions can increase participants' beliefs of PA engagement by means of reducing participants' perceived barriers to PA? **(RQ2)**

Findings from the current study show that the chatbot interventions were well-received by the participants with high levels of satisfaction. This provides valuable insights into the effectiveness of using a chatbot to overcome barriers to PA, and suggests that chatbots can be a promising tool for promoting PA by addressing specific barriers to PA. The **main contributions and novelties** of this work are: 1). The theory and evidence-grounded chatbot-delivered interventions designed by human experts to address six most common barriers to PA; 2). One of the first user studies to evaluate the effectiveness of a domain knowledge-grounded chatbot as a behavioral intervention to resolve specific barriers to PA; 3). Enable the exploratory vision that a chatbot can be integrated with AI techniques for automatically identifying individual barriers to PA during the conversation thereby providing tailored supports for PA promotion.

2. Background and Related work

Behavior change [30] is a complex process that is influenced by various factors, including individual differences and environmental factors. These factors can act as barriers to successful behavior change and can make it challenging for individuals to achieve their health goals. As a result, researchers have focused on identifying and addressing these barriers to improve the effectiveness of behavioral interventions. In recent years, studies have demonstrated that overcoming barriers to behavior change can improve the effectiveness of behavioral interventions. For instance, studies [1][2][3] found that addressing common barriers to PA (e.g., lack of motivation) was associated with significant improvements among individuals and can help individuals increase their likelihood of success in achieving their goals of behavior change.

In recent years, there has been increasing interest in using technology to help people overcome the barriers and promote PA [1][6][8]. In this context, chatbots have received considerable attention which can provide personalized advice to individuals, making it easier for people to engage in regular PA. Several studies have investigated the

effectiveness of chatbots in promoting PA [4][5][6][8][27]. From these studies, the chatbot was found to be effective in improving participants' self-reported PA levels. However, there are still several challenges that need to be addressed. For instance, there is a need for more studies that examine the effects of chatbot-delivered interventions on helping people overcome specific barriers during the behavior change, which is the research gap we explored in the present study. Incorporating natural language processing (NLP) techniques can enhance the capabilities of chatbots and make the chatbot more user-tailored and intelligent. For instance, in a study [6], a classification model was embedded in a chatbot to identify the barriers of smoking cessation. This model can classify users' responses into one of several predefined categories of barriers (e.g., stress or lack of motivation) and provide tailored responses to address such identified barrier.

3. Methods

3.1. Overview of the study design

This present study was designed to evaluate interventions to overcome different perceived barriers to PA. We employed a within-subjects design where all participants were presented with three random chatbot interventions and a control condition. Each intervention conversation addressed a specific barrier to PA, while the control conversation did not address any barriers. Three dependent variables were investigated in the study: beliefs in PA engagement, attitudes toward the effectiveness of the interventions in resolving their respective barriers to PA, and the overall chatbot experience. The study was approved by the Ethics Committee of Social and Behavioral science faculty in the University of Amsterdam. (ERB number: 2022-COP-15774)

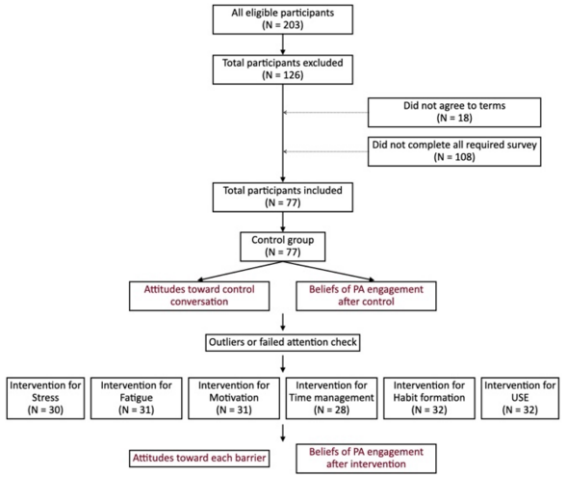

Figure 1. Flow of the study procedure and the recruitment process.

3.2. Participants

An a-priori statistical power analysis using G*Power revealed that the target was minimum 27 participants per intervention for a power of .8 to detect an effect size of .5 with a significance level of .05. Participants had to be at least 16 years old and proficient

in English and were required to have a smartphone on which the chatbot application (Telegram [33][33]) can be installed. Participants were recruited through social media, word of mouth, and Behavioral Science Lab of the University of Amsterdam. From January 6th 2023 to January 26th 2023, 77 participants were recruited who met the requirements, agreed to the terms of the study, and completed the study in its entirety.

3.3. Procedure

Participants started the study by reading an information letter in which they were informed on the goal and procedure of the study. If participants agreed to the terms of the consent form, they were presented with a link to the chatbot on the chatting application "Telegram". Otherwise, they were informed that they could withdraw from the study at any time. All participants went through both control conversation and three randomly assigned experimental intervention conversations. The control condition was always the first conversation that the participant came across followed by three experimental intervention conversations. The order of the experimental interventions was randomized by the survey tool Qualtrics to ensure the effects of each intervention were counterbalanced. After the control and each intervention conversation, participants answered surveys on Qualtrics about their attitudes toward the conversation for resolving the specific barrier and their beliefs of PA engagement. Lastly, participants were asked to evaluate their interactive experience with the chatbot. In all questionnaires attention checks were used to ensure that participants did not skip the questions. At the end of the study, all participants received a message thanking them for their participation.

3.4. Materials and Operationalizations

Chatbot mechanisms and embedded NLP techniques. The chatbot system was developed by the team leading the TIMELY project [34][36]. A technical scheme was proposed to efficiently and modularly develop such domain-specific conversational agents used for behavioral intervention and counselling, called 'Micro-conversation Scheme'. 'Micro-conversations' represent brief theory-grounded conversations pre-scripted by the domain experts. Each 'micro-conversation' has a conversational topic (e.g., intervention for fatigue, intervention for stress, goal setting) Then we used own developed dialogue management model to connect all these 'micro-conversations' by specific sequence and a central database to store all user information collected from ongoing 'micro-conversations' to make the chatbot contextual-aware. Additionally, the chatbot was enhanced by various NLP techniques (e.g., sentiment analysis, text classification and generation models) to give contextual-appropriate responses, for instance, a multi-tasks NLU layer to do both intent resolution and sentiment classification which enable the chatbot to give sentiment-appropriate responses. Besides, we embedded the generative model into the pre-scripted conversations as a flexible plugin module to generate the contextual reflections (e.g., affirmation and encouragement) on the user inputs, thereby facilitating users' motivation towards PA. Further, we propose that a classification model can be integrated to automatically detect the barrier(s) that people may have during the conversation, subsequently providing tailored intervention, which is a direction in the further study.

Conversational design and strategy. The conversations (micro-conversations) used in the present study were pre-scripted by human experts from the domain of Psychology and Behavioral Science. Conversations were grounded in the theories (e.g., Motivational

Interviewing (MI) [9] and various behavioral change techniques [1]) tailoring to resolve each specific barrier to PA identified in the aforementioned focus group with evidences and motivational techniques, such as MI, to facilitate users' intrinsic motivation towards PA. The pre-scripted conversational topics including greeting (to initially build relationship with users) and six interventions targeting six identified barriers to PA. We observed that users generally responded positively and engaged in meaningful dialogue with chatbot. Theory-grounded conversations were designed to address unique barriers ensuring that each dialogue was relevant and effective in helping individuals overcome their specific challenges. Focusing on one intervention per barrier simplified the chatbot development and provided the targeted solutions that directly addressed each barrier making it easier to evaluate the effectiveness of interventions. It also helped identify the successful interventions thereby optimizing chatbot's performance. The chatbot was able to adapt its responses based on users' input providing tailored support to address their concerns. Therefore, the chatbot's intelligence is not only based on the psychological and behavior change theories ensuring that our intervention content was evidence-based and theoretically grounded, but also enhanced by multiple additionally embedded NLU and NLG models to make the chatbot more intelligent and contextual-aware. Continuous evaluation and adaptation are essential for diverse populations and future research can explore other significant barriers by extending the micro-conversations.

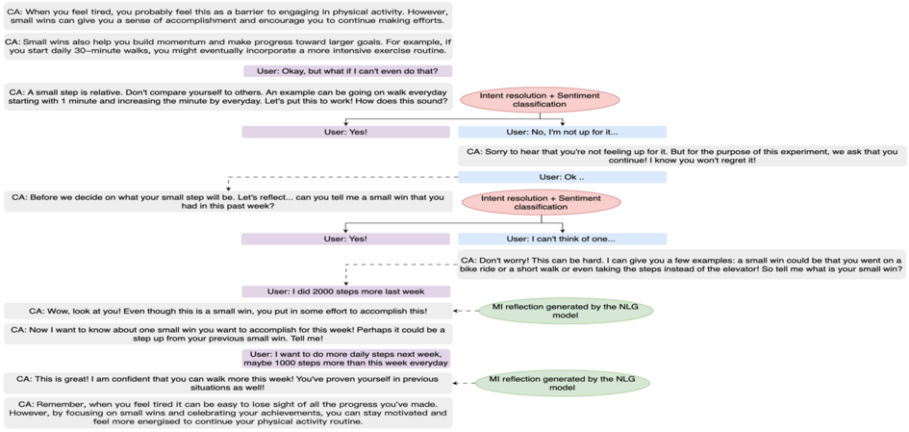

Figure 2. An example of conversational snippet using Behavioral Change Strategy (Small-Win) and MI to overcome barrier (Fatigue). Grey boxes represent CA's output utterances, while purple and blue box are user input (conversational paths depend on the result from sentiment model). The red oval highlights multi-task NLU layer (intent and sentiment classifications). The green oval is NLG model for generating MI reflections.

Control conversation. The control condition was chosen as a neutral conversational topic to establish a baseline for comparison with the experimental intervention conversations. Conversations for the control condition were designed to provide brief summaries of two books which described various strategies to increase PA. The barriers to PA were not addressed in the control conversation. The first book was "Not a Diet Book" by James Smith and the second book was "The Fitness Mindset" by Brian Keane. After the control conversation, the participants were asked to fill out the questionnaire about their attitudes toward the books and about their beliefs in PA engagement.

Intervention for reducing Stress. Previous research [11] concluded that stress could be a barrier to PA and mindfulness [12][32] was suggested to be able to significantly eliminate stress. For this information, the chatbot informed the participant that stress can be a barrier to PA, participants were then instructed to perform a two-minute mindfulness

exercise to reduce stress, which intended to be associated with an increased likelihood they would engage in PA. Following the mindfulness condition, participants were instructed to fill out a questionnaire regarding their attitudes toward the effectiveness of intervention to solve stress as a barrier to PA. This questionnaire was partly self-constructed and based on Perceived Stress Scale [13]. An example item of the questions was "Do you think the mindfulness exercise you just did would help you to become less stressed?" The items ranged from 1 ("strongly disagree") to 5 ("strongly agree").

Intervention for reducing Fatigue. Following the completion of the chatbot intervention, participants were asked to complete the modified fatigue assessment scale (FAS) questionnaire [14] to determine whether the intervention could alleviate symptoms of fatigue. The original FAS questionnaire [14] consisted of 10 items and the present study utilized six of those questions. An example item of the questions was "Do you think this intervention would help you with starting things?"All questions utilized a five-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree).

Intervention for enhancing Motivation. Participants had an intervention with the chatbot about intrinsic motivation for investigating the influence of mindfulness on PA engagement through intrinsic motivation as a mediator. The intervention was followed by a survey with 7 questions on the topics of intrinsic motivation. The questions on both topics were loosely based on specific items in the Exercise Motivation Scale [16]. An example item of the questionnaire was on the topics of intrinsic motivation, e.g., "Do you usually enjoy physical activity?" All questions were answered through a 5-point Likert scale, ranging from 1 "strongly disagree" to 5 "strongly agree".

Intervention for enhancing Time management. This intervention was used to improve time management by using implementation intentions. Participants were asked about their current level of PA and were informed about their barrier time management. Implementation intentions were presented as a strategy to overcome time management as a barrier by giving a definition and example of implementation intention to implement PA into their daily schedule. Time management was measured by a self-constructed five-item questionnaire based on the Assessment Time Management Skills (ATMS) questionnaire [19]. An example item was "Do you think this intervention would help you to manage your time well?" The items ranged from 1 ("strongly disagree") to 5 ("strongly agree"). A higher mean indicated higher time management skill.

Intervention for Habit formation. Participants in this intervention had a conversation designed to assist them in selecting relevant and stable environmental cues, e.g., having dinner that was expected to be followed by a 10-minute walk. They were also informed about important aspects of habit formation to provide more useful insight for the planning stage. After the intervention, participants were asked to answer five self-constructed questions (e.g., "Do you think that such an intervention about forming new habits can be useful for increasing physical activity?") that were supposed to measure the participants predictions about the usefulness of such an intervention in helping habit formation. The items ranged from 1 ("strongly disagree") to 5 ("strongly agree").

Intervention for resolving Unsupportive social environment (USE). In this intervention, participants had a conversation with the purpose to support the participant in identifying their tailored solution to create a supportive social environment for PA. Effectiveness of USE intervention was measured by a self-constructed eight-item questionnaire based on the interpersonal barriers of PA found in study [25]. Participants were asked for instance, "Do you think the plan you came up with will help you develop a supportive social environment for exercise?" The item was answered on a Likert scale ranging from 1 ("strongly disagree") to 5 ("strongly agree").

Belief of PA engagement. PA engagement was assessed using a self-constructed three-item questionnaire including items like "Do you think you would be more physically active after the conversation you just had?" The items ranged from 1 ("strongly disagree") to 5 ("strongly agree"). After each conversation (i.e., both control conversation and each intervention conversation), participants were asked to fill out the questionnaires. The present study used surveys to assess the subjective measurements of PA immediately after the interventions rather than the objective measurements of PA some time after the interventions, because current work aimed to explore the initial perceptions and beliefs of participants regarding the effectiveness of these interventions in overcoming barriers and promoting PA. The immediate post-intervention assessment of participants' beliefs and attitudes was a deliberate choice as it allowed us to capture their immediate reactions to the interventions. Future research could incorporate objective measurements of PA to complement subjective assessments.

Chatbot user experience. The user experience with the chatbot was measured using a self-constructed five-item questionnaire adapted from BUS questionnaire [10]. An example item included "Was the chatbot easy to understand?" The items ranged from 1 ("strongly disagree") to 5 ("strongly agree"). Participants were asked to fill out this questionnaire at the final stage of the study after all conversations had been completed.

4. Results

4.1. Main analysis

Stress. 30 participants were involved in the intervention for stress after all exclusion criteria. A one sample t-test was used to test whether expected believed stress levels (M=3.60, SD=.72) were significantly different than the neutral level of believed stress (M=3.00) after mindfulness condition. These results concluded that the difference in means was .60, $t(30) = 4.52$, $p < .001$. The questionnaire on believed stress was coded that higher scores on the questionnaire meant lower believed stress. A paired samples t-test was done to determine whether expected believed PA levels after control condition (M=3.40, SD =.69) were significantly different than the expected believed PA levels after mindfulness condition (M=2.97, SD=.96). Results of this test showed that the difference in means was .43, $t(30) = 3.03$, $p = .005$. Thus, expected believed PA levels after control condition were significantly higher than expected believed PA levels after mindfulness condition, which was not in line with our prediction. A correlational analysis was conducted to test whether, after mindfulness condition, expected believed stress levels (M=3.60, SD=.72) were significantly correlated to expected believed PA levels (M=2.97, SD=.96). The results determined that the correlation is $r(30) = .57$, $p = .001$. Hence, expected believed stress levels were significantly correlated to expected believed PA levels after mindfulness condition, which was commensurate with our prediction.

Fatigue. 31 participants were involved in the intervention for fatigue after all exclusion criteria. A one-sample t-test was performed to determine whether participants' attitudes toward the effectiveness of this intervention can be helpful to reduce fatigue. The results showed that the mean fatigue score (M=3.32, SD=.60) was significantly higher than the neutral value (M=3.00), $t(31) = 3.60$, $p < .001$. A paired samples t-test was performed to determine whether there was higher difference in the beliefs of PA engagement in the experimental condition compared to the control condition. The higher the scores meant the more that participants felt that they would want to be more

physically active after the intervention. There was a significant difference in beliefs of PA engagement between the conditions, namely the experimental condition (M=3.27, SD=.83) had a mean of 0.37 higher than the control group (M=2.90, SD=1.04), $t(31) = 1.90$, $p = .033$. A Pearson's correlation analysis was performed and a significantly positive correlation was found between the mechanism of action and beliefs of PA engagement in the experimental condition, $r(31) = .66$. A moderate positive correlation was found between the mechanism of action and beliefs of PA engagement in the control condition, $r(31) = .23$.

Motivation. 31 participants were involved in this intervention after all exclusion criteria as the experimental condition to overcome the barrier of lacking motivation. The results of the one sample t-test showed that the mean motivation score for the participants (M=3.54, SD=.49) was significantly higher than the population mean of 3.00, $t(31) = 7.69$, $p < .001$. This shows that the mindfulness-based intervention used in the chatbot conversation had a positive influence on the participants' scores on the intrinsic motivation questionnaire. The results of the paired samples t-test showed that the mean score on the beliefs of PA engagement in the experimental condition (M=3.44, SD=.99) was significantly higher than the control condition (M=2.95, SD=.91); $t(31) = -3.51$, $p = .005$). A Pearson's correlation analysis was performed to compare the relationship between the scores on the intrinsic motivation questionnaire and respectively scores on the beliefs of PA engagement questionnaire. The results show that there was an important difference in the correlation with the mean score on the beliefs of PA between the control group ($r=.04$, $p=.81$) and the experimental group ($r=.62$, $p<.001$), which means, the intervention group showed a significant positive correlation.

Time management. 28 participants were involved in this intervention after all exclusion criteria. The one sample t-test revealed that time management (M=3.28, SD=.76) after the intervention differ from the neutral value (M=3.00), $t(28) = 1.94$, $p = .063$, indicating an improvement of time management after the intervention conversation. The paired samples t-test showed that participants reported higher beliefs of PA engagement after the intervention conversation (M=3.40, SD=.78) than after the control conversation (M=3.23, SD=.85), $t(28) = -1.19$, $p = .285$, indicating potential effectiveness of the intervention on increasing belief of PA engagement. The correlation analyses showed that all measured variables were positively correlated. Time management significantly predicted beliefs of PA engagement after the intervention conversation, $r = .54$, $p = .002$, and after the control condition, $r = .46$, $p = .013$.

Habit formation. 32 participants were involved in this intervention after all exclusion criteria. One sample t-test was performed to examine participants' opinions about whether a chatbot conversation about habit formation can be useful for helping people with habit formation (M=3.82, SD=.58), $t(32) = 7.94$, $p < .001$, compared to the population value (M=3.00). Paired samples t-test was performed. In line with our expectations, participants after the habit formation chatbot conversation scored significantly higher (M=3.57, SD=.44) on beliefs to engage in more PA, than after control chatbot conversation (M=3.28, SD=.72), $t(32) = -2.46$, $p = .024$. A correlational analysis showed that participants' opinions about whether a chatbot conversation about habit formation can be useful after the habit formation condition were significantly correlated to belief of PA engagement, $r(32) = .68$.

Unsupportive social environment. 32 participants were involved in this intervention after all exclusion criteria. One sample t-test showed that mean score of attitudes towards reducing USE (M=3.43, SD=.75) was higher than the population mean 3.00, $t(32) = 3.26$, $p < .001$. A paired samples t-test was conducted to compare the mean scores on beliefs

in PA engagement. On average participants reported slightly stronger beliefs in PA engagement after the intervention (M=3.08, SD=.86), in comparison to after the control (M=3.02, SD=.97). In addition, the correlations were conducted to see whether those differ in predictability of belief in PA engagement. The result indicates that there was a significant correlation between attitude toward the intervention to resolve the barrier USE and beliefs in PA engagement after experimental condition, $r = .71$, $p < .001$.

Table 1. Descriptive statistics on three outcome variables. BPAE represents the belief of PA engagement.

Variables	n	Min	Max	Mean	SD
Attitudes toward reducing Stress	30	1.00	4.00	3.60	.72
BPAE after Experimental condition	30	1.00	4.00	2.97	.96
BPAE after Control condition	30	1.00	4.00	3.40	.69
Attitudes toward reducing Fatigue	31	1.33	4.00	3.32	.60
BPAE after Experimental condition	31	1.00	4.00	3.27	.83
BPAE after Control condition	31	1.00	4.33	2.90	1.04
Attitudes toward enhancing Motivation	31	2.14	4.29	3.54.	.49
BPAE after Experimental condition	31	1.00	4.67	3.44	.99
BPAE after Control condition	31	1.00	4.33	2.95	.91
Attitudes toward enhancing Time Management	28	2.00	4.40	3.28	.76
BPAE after Experimental condition	28	1.67	4.67	3.40	.78
BPAE after Control condition	28	1.00	4.33	3.23	.85
Attitudes toward Habit Formation	32	1.00	5.00	3.82	.58
BPAE after Experimental condition	32	1.00	4.25	3.57	.44
BPAE after Control condition	32	2.00	4.33	3.28	.72
Attitudes toward resolving USE	32	1.88	4.75	3.43	.75
BPAE after Experimental condition	32	1.67	4.33	3.08	.86
BPAE after Control condition	32	1.00	4.33	3.02	.97
Overall chatbot experience	77	1.00	5.00	3.43	.78

Table 2. One sample t-test on the attitudes toward the effectiveness of each intervention to resolve specific barrier. VS. the population mean (M = 3.00).

Variables	Difference in mean	t	p
Attitudes toward reducing Stress	.60	4.52	< .001
Attitudes toward reducing Fatigue	.32	3.60	< .001
Attitudes toward enhancing Motivation	.50	7.69	< .001
Attitudes toward enhancing Time Management	.28	1.94	.063
Attitudes toward Habit Formation	.82	7.94	< .001
Attitudes toward resolving USE	.43	3.26	< .001
Overall chatbot experience	.50	5.81	< .001

Table 3. Paired samples t-test on belief of PA engagement (BPAE) between control and experimental condition.

Variables	Difference in mean	t	p
Stress: BPAE control VS. BPAE experimental	.43	3.03	.005
Fatigue: BPAE control VS. BPAE experimental	-.37	1.90	.033
Motivation: BPAE control VS. BPAE experimental	-.57	-3.51	.005
Time Management: BPAE control VS. BPAE experimental	-.17	-1.19	.285
Habit Formation: BPAE control VS. BPAE experimental	-.29	-2.46	.024
USE: BPAE control VS. BPAE experimental	-.06	-.37	.726

Table 4. Correlation analysis between three outcome variables.

Variables	1	2	3
Attitudes toward reducing Stress	1		
BPAE after Experimental condition	.57	1	
BPAE after Control condition	.54	.59	1
Attitudes toward reducing Fatigue	1		
BPAE after Experimental condition	.66	1	
BPAE after Control condition	.23	-.43	1

Attitudes toward enhancing Motivation	1		
BPAE after Experimental condition	.62	1	
BPAE after Control condition	.04	.51	1
Attitudes toward enhancing Time Management	1		
BPAE after Experimental condition	.54	1	
BPAE after Control condition	.46	.56	1
Attitudes toward Habit Formation	1		
BPAE after Experimental condition	.68	1	
BPAE after Control condition	.50	.43	1
Attitudes toward resolving USE	1		
BPAE after Experimental condition	.71	1	
BPAE after Control condition	.58	.45	1

5. Discussion

5.1. Main findings

In this empirical study, we aimed to investigate the effectiveness of chatbot-delivered interventions in addressing specific barriers to PA. Our findings provided valuable insights into the potential of chatbot-delivered interventions.

Stress. The major results of this intervention were in line with the hypotheses. Namely, the mindfulness conversation effectively reduced stress in participants and this reduction was associated with increased belief of PA engagement. The findings partly answered the research question that chatbot conversation about mindfulness decreases the likelihood that people experience stress as a barrier to PA. However, participants believed they would engage in more PA after control condition than the mindfulness condition, contrary to our expectations. It would be possible to say while the chatbot conversation about mindfulness can reduce the likelihood that participants experience stress as barrier to PA, the overall effectiveness of intervention needs further exploration.

Fatigue. The intervention utilizing graded exercise therapy [15] to address fatigue was met with positive results, which support the hypothesis that a chatbot-delivered intervention shows potential in decreasing the feeling of fatigue while concurrently increasing the beliefs of PA engagement. This was evidenced by the mean of the participant's attitudes towards the intervention being significantly higher than the neutral value. The finding also supports the hypotheses as the experimental fatigue group showed higher levels of beliefs to engage in PA compared to the control condition. In addition, a stronger positive correlation was found between the mechanism of action and beliefs of PA engagement in the experimental condition than the control condition.

Motivation. The mindfulness-based intervention positively influenced intrinsic motivation in the context of beliefs in PA engagement. This finding aligns with previous research [17] that showed mindfulness-based intervention can increase the intrinsic motivation in PA engagement. The paired samples t-test results suggested that the mindfulness-based intervention used in the experimental condition had a positive influence on the beliefs of PA engagement which is also consistent with previous research [17] and hypothesis. The correlation comparison results suggested that the mindfulness-based intervention had a stronger correlation in establishing the relationship between intrinsic motivation and the beliefs of PA engagement which is in line with the research [18]. This study provided evidence to the effectiveness of a mindfulness-based intervention in enhancing intrinsic motivation for increasing beliefs of PA engagement.

Time management. This intervention aimed to determine whether a chatbot conversation about implementation intentions can decreases the likelihood that people experience time management as a barrier to PA. The results of one sample and paired samples t-tests showed a tendency that a chatbot conversation about implementation intentions better improves time management and beliefs in PA engagement compared to the control condition, and the participants with a high score on time management reported higher beliefs of PA which indicated that better time management predicted higher beliefs of PA engagement. These conclusions do agree with previous research [20] and supported the hypothesis that implementation intentions used to increase time management showing potential effectiveness in increasing the beliefs of PA engagement.

Habit management. This intervention aimed to examine if the chatbot can be an effective tool to help people form new habits to do more PA. Results of the t-tests were in line with previous studies [21][22][23] and the hypothesis that habit formation intervention lead to increased willingness to engage in more PA compared to the control conversation. Results were also in line with the predictions of the habit formation model proposed in [24]. Additionally, the result of the correlational analysis and participants' positive evaluation on the intervention indicated that chatbot-delivered habit formation intervention was highly promising for individuals struggling to establish new PA habits.

USE. The current intervention built on previous research [26], which was one of the first to investigate the mediating effect of USE on PA through a chatbot-delivered MI intervention to the best of our knowledge. This intervention sought to investigate the effectiveness of a chatbot intervention for increasing PA in individuals with an USE. The main results suggested that there was no significant effect of the USE strategy delivered by the chatbot when comparing intervention and control conditions. However, the mean in the intervention group did indicate positive effectiveness of the intervention to a certain degree. Additionally, the correlation analyses did seem to signify that the effectiveness of the USE intervention was predictive of the beliefs in PA engagement.

Overall chatbot experience. Intrigued by previous studies [6][8] using chatbots in psychology and healthcare, this study investigated whether chatbots can be an effective tool for delivering behavioral interventions. In line with our predications the chatbot was evaluated positively by the participants (M=3.43, SD=.78). Participants found the chatbot was easy to understand and gave enough information, which points to the fact that it can be used to mitigate the current need for increased healthcare resources. Despite the positive feedback that was received, some participants reported that they might not utilize the chatbot in their daily lives. This suggested that chatbot-based interventions should not be viewed as a complete replacement for traditional (mental) health interventions but rather as to supplement and enhance existing interventions. It could also be used to bridge the gap for patients who are on the waiting lists to receive treatment.

5.2. Strengths and limitations

This is one of the first studies exploring the use of theory-grounded chatbot to help people overcome the specific perceived barriers to PA as tailored intervention, thereby promoting the level of PA. The feasibility and effectiveness of this novel approach was investigated in this study. All pre-scripted conversational data and study design enabled replication and provided training examples for future chatbot development in this direction. Moreover, the findings of this study provide insights and possibility to develop our envisioned chatbots integrating automatic barrier identification model and provide subsequent user-tailored intervention, which build theoretical basis for future work.

The present study sheds some light on the use of chatbots for promoting PA, however, there are several limitations. Firstly, a self-constructed questionnaire was used to assess the beliefs of PA engagement, potentially leading to biased responses as people tend to respond in a way that others view as favorable. Future research should employ more objective measurements. Secondly, the missing comprehensive demographic information of participants made it impossible to draw conclusions about the representativeness of the tested sample. However, participants were recruited through the university which can therefore be assumed that a large part of the sample consisted of students. Thirdly, there was no pre-assessment to compare the attitudes towards interventions and beliefs of PA engagement. The use of pre-assessment measurement is important in experimental research as they establish a baseline to assess the change caused by the intervention. However, the decision not to use a pre-assessment was made to focus on the effects of different conversations on overcoming barriers and to keep the study design streamlined, reducing participant burden, and maintaining engagement. Baseline differences were assumed to be equally distributed across experimental conditions. The merits of pre- and post-test measurements are acknowledged, and appropriate pre-assessments will be considered for future studies with more complex designs. Fourthly, the study population may not fully represent individuals facing PA barriers and willing to engage in interventions. Current study aimed to initially explore the effectiveness of chatbot-delivered interventions in a broader population. Future studies will target more diverse and representative samples, including individuals facing greater PA challenges and intervention engagement. The low study completion rate may be attributed to factors such as strict criteria for excluding participants who did not meet all requirements, research credit-driven participation, curiosity clicks, and technical difficulties. To mitigate this limitation, future studies will encourage completion through clearer instructions, incentives, and more stable technical configurations.

6. Conclusion and Future work

Previous studies have demonstrated the gravity of finding innovative strategies to improve people's PA. This study provides valuable insights into the potential effectiveness of using chatbot-delivered interventions to affect the likelihood that people experience specific barriers to PA, thereby promoting the level of PA. Substantial evidence has been found proving the relevance of overcoming specific barriers for increasing PA and demonstrating the positive perception of chatbots in the context of PA promotion. The present study has enhanced the understanding of the interventions delivered by chatbots that target the barriers to PA. The exploratory nature of this study allows us to draw a few implications for future research. Firstly, although chatbots can provide human-like interaction, they should not substitute for human healthcare providers. However, adding chatbots can enhance healthcare services by improving anonymity, accessibility, and personalization. Secondly, the utilization of AI and generative models in chatbots can enhance the efficacy of behavioral change therapy. For instance, integrating a text classification model to detect potential user barriers and provide tailored interventions may help people overcome the perceived barriers more efficiently. Thirdly, a long-term exploratory study on more representative population is planned to investigate the effectiveness of chatbots incorporating additional AI models and targeting a wider range of barriers to behavioral change, such as diet and sleep, leading to more comprehensive and reliable results and evidence.

References

[1] Dishman, Dishman, R. K., Berthoud, H.-R., Booth, F. W., Cotman, C. W., Edgerton, V. R., Fleshner, M. R., & Zigmond, M. J. (2006). Neurobiology of exercise. Obesity, 14(3), 345-356.

[2] Veldman, M. H. J., van der Aa, H. P. A., Bode, C., Knoop, H., Hulshof, C. T. J., Koopmanschap, M., Stavleu, E., van Rens, G. H. M. B., & van Nispen, R. M. A. (2021). E-nergEYEze, a vision-specific eHealth intervention based on cognitive behavioral therapy and self-management to reduce fatigue in adults with visual impairment: study protocol for a randomized controlled trial. Trials, 22(1). https://doi.org/10.1186/s13063-021-05935-w

[3] Chodzko-Zajko, W. J., Proctor, D. N., Fiatarone Singh, M. A., Minson, C. T., Nigg, C. R., Salem, G. J., & Skinner, J. S. (2009). Exercise and Physical Activity for Older Adults. Medicine &Amp; Science in Sports &Amp; Exercise, 41(7), 1510–1530. https://doi.org/10.1249/mss.0b013e3181a0c95c

[4] Lim, S. M., Shiau, C. W. C., Cheng, L. J., & Lau, Y. (2022). Chatbot-Delivered Psychotherapy for Adults With Depressive and Anxiety Symptoms: A Systematic Review and Meta-Regression. Behavior Therapy, 53(2), 334–347. https://doi.org/10.1016/j.beth.2021.09.007

[5] Ahmed, A., Ali, N., Aziz, S., Abd-alrazaq, A. A., Hassan, A., Khalifa, M., Elhusein, B., Ahmed, M., Ahmed, M. a. S., & Househ, M. (2021). A review of mobile chatbot apps for anxiety and depression and their self-care features. Computer Methods and Programs in Biomedicine Update, 1, 100012. https://doi.org/10.1016/j.cmpbup.2021.100012

[6] Almusharraf, F., Rose, J., & Selby, P. (2020). Engaging Unmotivated Smokers to Move Toward Quitting: Design of Motivational Interviewing–Based Chatbot Through Iterative Interactions. Journal of Medical Internet Research, 22(11), e20251. https://doi.org/10.2196/20251

[7] Bendig, E., Erb, B., Schulze-Thuesing, L., & Baumeister, H. (2019). The Next Generation: Chatbots in Clinical Psychology and Psychotherapy to Foster Mental Health – A Scoping Review. Verhaltenstherapie, 1–13. https://doi.org/10.1159/000501812

[8] D'Alfonso, S. (2020). AI in mental health. In Current Opinion in Psychology (Vol. 36, pp. 112–117). Elsevier B.V. https://doi.org/10.1016/j.copsyc.2020.04.005

[9] Rollnick, S., & Miller, W. (1995). What is Motivational Interviewing? Behavioural and Cognitive Psychotherapy, 23(4), 325-334. Doi:10.1017/S135246580001643X

[10] Borsci, S., Malizia, A., Schmettow, M. et al. The Chatbot Usability Scale: the Design and Pilot of a Usability Scale for Interaction with AI-Based Conversational Agents. Pers Ubiquit Comput 26, 95–119 (2022). https://doi.org/10.1007/s00779-021-01582-9

[11] Stults-Kolehmainen, M. A., & Sinha, R. (2013). The Effects of Stress on Physical Activity and Exercise. Sports Medicine, 44(1), 81–121. https://doi.org/10.1007/s40279-013-0090-5

[12] Schachter, J., Ajayi, A. A., & Nguyen, P. L. (2022). The moderating and mediating roles of mindfulness and rumination on COVID-19 stress and depression: A longitudinal study of young adults. Journal of Counseling Psychology, 69(5), 732–744. https://doi.org/10.1037/cou0000626

[13] Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A Global Measure of Perceived Stress. Journal of Health and Social Behavior, 24(4), 385. https://doi.org/10.2307/2136404

[14] Hegarty, D. (2022, May 20). Fatigue Assessment Scale (FAS). NovoPsych. https://novopsych.com.au/assessments/health/fatigue-assessment-scale-fas/

[15] White, P., Goldsmith, K., Johnson, A., Potts, L., Walwyn, R., DeCesare, J., Baber, H., Burgess, M., Clark, L., Cox, D., Bavinton, J., Angus, B., Murphy, G., Murphy, M., O'Dowd, H., Wilks, D., McCrone, P., Chalder, T., & Sharpe, M. (2011). Comparison of adaptive pacing therapy, cognitive behaviour therapy, graded exercise therapy, and specialist medical care for chronic fatigue syndrome (PACE): a 13andomized trial. The Lancet, 377(9768), 823–836. https://doi.org/10.1016/s0140-6736(11)60096-2

[16] Brière, N. M., Vallerand, R. J., Blais, M. R., & Pelletier, L. G. (1995). Developing and validating a measure of intrinsic and extrinsic motivation in sports: A scale of motivation for Sports (EMS). International Journal of Sport Psychology, 26(4), 465–489.

[17] Robin, N., Toussaint, L., Sinnapah, S., Hue, O., & Coudevylle, G. R. (2020). Beneficial Influence of Mindfulness Training Promoted by Text Messages on Self-Reported Aerobic Physical Activity in Older Adults: A Randomized Controlled Study. Journal of Aging and Physical Activity, 28(3), 406–414. https://doi.org/10.1123/japa.2019-0002

[18] Dacey, M. (2008). Older Adults' Intrinsic and Extrinsic Motivation Toward Physical Activity. American Journal of Health Behavior, 32(6). https://doi.org/10.5993/AJHB.32.6.2

[19] Suzanne M. White, Anne Riley & Peter Flom (2013) Assessment of Time Management Skills (ATMS): A Practice-Based Outcome Questionnaire, Occupational Therapy in Mental Health, 29:3, 215-231, DOI: 10.1080/0164212X.2013.819481

[20] Hagger, M., & Luszczynska, A. (2014). Implementation Intention and Action Planning Interventions in Health Contexts: State of the Research and Proposals for the Way Forward. https://doi.org/10.1111/aphw.12017

[21] Feil, K., Allion, S., Weyland, S., & Jekauc, D. (2021). A Systematic Review Examining the Relationship Between Habit and Physical Activity Behavior in Longitudinal Studies. In Frontiers in Psychology (Vol. 12). Frontiers Media S.A. https://doi.org/10.3389/fpsyg.2021.626750
[22] Kaushal, N., Rhodes, R. E., Spence, J. C., & Meldrum, J. T. (2017). Increasing Physical Activity Through Principles of Habit Formation in New Gym Members: a Randomized Controlled Trial. Annals of Behavioral Medicine, 51(4), 578–586. https://doi.org/10.1007/s12160-017-9881-5
[23] Lally, P., Wardle, J., & Gardner, B. (2011). Experiences of habit formation: A qualitative study. Psychology, Health and Medicine, 16(4), 484–489. https://doi.org/10.1080/13548506.2011.555774
[24] Verplanken, B. (2018). The psychology of habit: Theory, mechanisms, change, and contexts. In The Psychology of Habit: Theory, Mechanisms, Change, and Contexts. Springer International Publishing. https://doi.org/10.1007/978-3-319-97529-0
[25] Choi, J., Cho, J., Shin, N. M., & Tsoh, J. (2020). Exploring Barriers to and Facilitators of Physical Activity among Korean American Women. Western Journal of Nursing Research, 43(9), 817–827. https://doi.org/10.1177/0193945920980453
[26] Zhang, J., Oh, Y. J., Lange, P., Yu, Z., & Fukuoka, Y. (2020). Artificial Intelligence Chatbot Behavior Change Model for Designing Artificial Intelligence Chatbots to Promote Physical Activity and a Healthy Diet: Viewpoint. Journal of Medical Internet Research, 22(9), e22845. https://doi.org/10.2196/22845
[27] Oh, Y.J., Zhang, J., Fang, ML. et al. A systematic review of artificial intelligence chatbots for promoting physical activity, healthy diet, and weight loss. Int J Behav Nutr Phys Act 18, 160 (2021). https://doi.org/10.1186/s12966-021-01224-6
[28] Mikkelsen, K., Stojanovska, L., Polenakovic, M., Bosevski, M., & Apostolopoulos, V. (2017). Exercise and mental health. Maturitas, 106, 48–56. https://doi.org/10.1016/j.maturitas.2017.09.003
[29] Warburton, D. E. R., Nicol, C. W., & Bredin, S. S. D. (2006). Health benefits of physical activity: The evidence. Canadian Medical Association Journal, 174(6), 801-809.
[30] Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol. 2008 May;27(3):379-87. Doi: 10.1037/0278-6133.27.3.379. PMID: 18624603.
[31] Behavior Change Interventions Delivered by Mobile Telephone Short-Message Service. Fjeldsoe, Brianna S. et al. American Journal of Preventive Medicine, Volume 36, Issue 2, 165 – 173.
[32] Mindfulness-Based Stress Reduction for Stress Management in Healthy People: A Review and Meta-Analysis. Alberto Chiesa and Alessandro Serretti. The Journal of Alternative and Complementary Medicine 2009 15.
[33] Telegram FZ LLC & Telegram Messenger Inc. (2013). Telegram (9.3.3) [Software]. https://telegram.org.
[34] CORDIS | European Commission. (2020). https://cordis.europa.eu/project/id/101017424.
[35] Rothbaum, B. O., Meadows, E. A., Resick, P., & Foy, D. W. (2000). Cognitive-behavioral therapy. In E. B. Foa, T. M. Keane, & M. J. Friedman (Eds.), Effective treatments for PTSD: Practice guidelines from the International Society for Traumatic Stress Studies (pp. 320–325). The Guilford Press.
[36] Xin Sun, Boris Schmitz, and Jos A. Bosch. 2023. TIMELY: Providing In-Time and Intelligent Support for Cardiovascular Rehabilitation with 'Patients and Practitioners in the Loop' Interaction. In Companion Proceedings of the 28th International Conference on Intelligent User Interfaces (IUI '23 Companion). Association for Computing Machinery, New York, NY, USA, 155–157. https://doi.org/10.1145/3581754.3584158
[37] Schmitz, Boris & Wirtz, Svenja & Sestayo-Fernández, Manuela & Schäfer, Hendrik & Douma, Emma & Alonso, M & González-Salvado, Violeta & Habibovic, Mirela & Kop, Willem & Peña, Carlos & Mooren, F. (2022). Defining patients needs and expectations for eHealth-based cardiac rehabilitation in Germany and Spain: living lab data from the TIMELY study. European Heart Journal. 43. 10.1093/eurheartj/ehac544.2828.