Toward Scalable Content Generation For Gamified mHealth Interventions: The Evaluation Of LLM-Generated Goals On User Engagement

Lorenzo J James[1][0009-0003-2505-8057], Laura Genga[1], Barbara Montagne[2], Muriel M Hagenaars[3], and Pieter ME Van Gorp[1]

[1] Eindhoven University of Technology, Eindhoven, The Netherlands
L.J.James@tue.nl
[2] Center for Mental Health Care, Amersfoort, The Netherlands
[3] Utrecht University, Utrecht, The Netherlands

Abstract. Gamified mHealth applications are increasingly used to promote healthy behavior change, yet low participant engagement remains a challenge. Personalized content has shown promise in increasing engagement, however, creating personalized content is a time-consuming task. Generative Artificial Intelligence models, particularly Large Language Models, offer a potential solution with their ability to quickly generate relevant content. In this study, we aim to evaluate the impact of LLM-generated content on participant engagement in health interventions using gamified mHealth applications. A total of 73 students and staff members of a university participated in a health intervention and were assigned into groups receiving either LLM-generated goals or predetermined goals. Engagement and perceived intrinsic motivation were measured and compared between the two groups. The results show no significant difference in engagement or perceived intrinsic motivation between participants receiving LLM-generated goals and those receiving predetermined goals. LLM-generated goals did not significantly negatively impact participant engagement and could therefore possibly offer a time-efficient approach to scalable content generation for mHealth applications.

Keywords: mHealth · health intervention · llm · gamification · goal generation · goal setting · large language models · goals · procedural content generation

1 Introduction

According to data from the World Health Organization, in today's society, Non-Communicable Diseases (NCDs) are responsible for approximately one-third of global fatalities [31]. NCDs such as cardiovascular diseases, cancers, respiratory diseases, and diabetes, are chronic and result from a combination of genetic, physiological, environmental, and behavioral factors [31] [2]. Engaging in unhealthy habits such as smoking, physical inactivity, and unhealthy diets, increases the risk of developing and dying from NCDs. Despite their chronic nature, by making healthier lifestyle choices and retaining healthy habits, some NCD risks can be mitigated and even prevented [27] [3]. However, the process of behavior change toward healthier lifestyle choices is a long, difficult, and non-linear process that happens gradually over time [24].

In recent years, mHealth applications have increasingly been used as a cost-effective and scalable tool to promote healthy behavior change, self-management, and monitoring of daily functioning of patients diagnosed with various NCDs such as diabetes, heart diseases, and asthma [10] [5] [9]. However, despite the increasing evidence of the effectiveness of mHealth applications in fostering behavioral changes, low participant engagement and high dropout rates have commonly been reported [9] [25]. The use of game design elements outside of the context of games also known as gamification, has been reported as a potential factor in increasing participant engagement in mHealth applications during health interventions [4] [6] [32]. An additional potentially effective factor to positively increase engagement in mHealth applications is the use of personalization [21] [9]. Personalization refers to a system tailoring content to the preferences of the user [26].

Although, personalized content has been associated with increased effectiveness of behavior change strategies and adherence [9] [14], creating personalized content for individual participants in health interventions is a time-consuming task [30]. This is especially true when involving caregivers and health coaches in the process of creating personalized content for health interventions [13]. Caregivers and health coaches often have multiple clients to whom they provide services and given their high workload, it may not be an effective use of their time [13]. It is, however, still critical to involve health professionals in the loop during health interventions to take appropriate personalized health goals and milestones into account [12] [30].

Generative Artificial Intelligence models, particularly Large Language models (LLMs) have displayed the capability to generate relevant content for potential use in mHealth applications with the caveat of the generated content being supervised by a healthcare professional [12] [8]. The ability of LLMs to quickly generate relevant content can be leveraged to effectively deliver effective and personalized content for mHealth applications [1]. Previous research exploring the use of LLMs for rapid content creation for health interventions has shown promising results [30]. Through the use of structured prompts including users' context, and behavior change theories, LLMs have the capabilities to possibly generate relevant personalized health goals for users more efficiently than caregivers [12]. To the best of our knowledge, there have been no studies that have empirically evaluated the impact of LLM-generated goals on the engagement of participants in health interventions.

In this paper, we will explore the use of scalable goal creation for health interventions by evaluating the impact of LLM-generated goals on the engagement of participants using gamified mHealth applications. In the following section, we explore relevant related work and the behavior change literature considered when generating goals using LLMs for mHealth applications. Next, we outline the method used in the experiment and report the results. We then address the findings, limitations, and future work, followed by the paper's conclusion.

2 Theoretical Background

mHealth applications designed to stimulate behavior change in individuals are more effective when the digital health content delivered is grounded in behavior change theory [30]. To achieve healthier behaviors, individuals need to change their behavior to achieve their health goals. Therefore when generating relevant goals in this study, for the participants using the mHealth applications the prompts used in the LLM integrate behavior change theories. In the following section, we explore the underlying behavior change theories used when generating content for the mHealth application. We also explore using prompt engineering to guide the LLM to generate measurable goals using LLMs. As mHealth applications can track an individual's progression toward measurable goals.

Motivation plays an important role in getting individuals to perform behavior change. Particularly, intrinsic motivation which stems from personal interests [23]. The Self Determination Theory (SDT), states that individuals are intrinsically motivated when three needs are fulfilled: 1) Competence, 2) Autonomy, and 3) Relatedness [23]. Similarly, the Fogg model states that for individuals to be motivated, the necessary ability, and trigger, are needed to foster behavior change [7]. The COM-B model states that individuals need to have the capability, opportunity, and motivation for behavior change [18]. Each behavior change theory highlights the need for individuals to have the right competence, ability, and capability to achieve the behavior change and to have the motivation to change [23] [7] [18]. The SDT states that participants should feel autonomous and related to the goals they are working towards to be intrinsically motivated to work on them [23]. The Fogg model states that a trigger is needed for behavior change, which in this study is the mHealth application [7]. Integrating goal setting based on the Goal Setting Theory (GST) with a behavior change theory proves to be an effective method for fostering behavior change [16]. Goals need to be specific, measurable, attainable, realistic, and time-bound (SMART), and follow the principles of GST [20] [16]. When goals are not measurable, it is more difficult to track the progress toward completing the goal.

The use of LLMs is most promising in domains where humans and AI tools work together as collaborators [29]. To integrate behavior change theories into the goals generated by the LLM, LLMs can be given context through a series of specific instructions called prompts [29]. Effectively using prompts and prompt engineering results in LLMs generating the most relevant content [15]. Prompt engineering involves designing tailored and effective prompts to elicit desired responses from the LLM. For the procedural creation of LLM-generated goals, prompts can be crafted that will leverage behavior change theories, use a robust structure, and include the context of users through user input [12]. Prompt engineering could also include adjusting the parameters of LLMs to generate more accurate, sensitive, and valuable responses related to the health intervention. This process helps diversify the LLM's output to ensure it offers a more variety of content [17]. In this study, we used prompt engineering to generate measurable goals in a SMART structure based on the GST.

3 Method

3.1 Study Design

The methodology used in the study was a mixed-method design, incorporating both qualitative and quantitative approaches. The qualitative aspect involved interviews and surveys to gather subjective perceptions from participants, while the quantitative aspect included an experiment and additional surveys to collect objective behavioral data. This approach was chosen to compare the subjective perceptions of participants with their objective behavioral data.

The experiment was conducted as a two-arm experimental study spanning five weeks. Participants were randomly allocated to either the group receiving LLM-generated goals or the group receiving the predetermined goals. Participants in the group receiving LLM-generated goals were asked to use the GOALS System to generate the goals they will work on during the health intervention. Participants in the group receiving predetermined goals were informed they would receive goals in the lifestyle dimensions of social, physical, and cognitive, to work on during the health intervention. The health intervention was split into two intervention phases, there were two periods of two weeks in which participants were given their respective tasks through the mHealth application, and a one-week break period between the two intervention phases. In both intervention periods, the mHealth application displayed the same goals assigned to them initially.

Before receiving or generating goals using the GOALS System, participants were given an information form and were asked to sign an informed consent form. All ethical procedures, data and privacy protection, and study procedures were approved by the Ethical Review Board of Eindhoven University of Technology under the experiment: ERB2023ID547.

3.2 Intervention Context

mHealth app The mHealth application used for this study was a configuration of the GameBus gamification platform. GameBus facilitates researchers to create various configurations of applications for use in health interventions [28]. We created, a cross-platform web application variant for the intervention phase of this study. The custom configuration of the application was titled "New Year's Resolution Campaign", and was designed to promote participant engagement with lifestyle goals. The application encouraged participants to complete lifestyle goals through gamification elements such as points, leaderboards, and a level system.

On the registration page of the application as seen in 3.2, participants can self-report the completion of goals. When self-reporting an activity, participants can optionally add a description of how the challenge was completed and optionally provide photographic or video evidence of them achieving their goal.

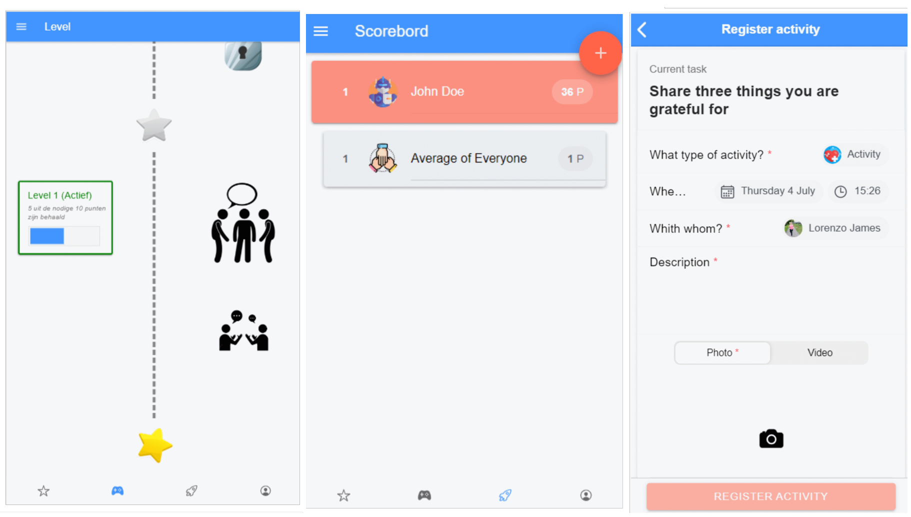

Fig. 1. The figure on the left shows the level system page displaying players' progress. The middle figure showcases the leaderboard page where players can compare how they scored against the average of all participants. The figure on the right displays the registration page used to self-report the completion of goals.

The lifestyle-related goals in the mHealth application were structured in a level system, where participants needed to earn 10 points per level to proceed to the next level. Participants earned points by either completing the LLM-generated or predetermined goals shown in the application, depending on which study group they were in. Once all the available goals within a level were achieved, participants moved on to the next level with new goals. The points assigned to each goal were based on the difficulty of the goal. The points participants earned were also displayed on the leaderboard page which showed the average points of all participants, for participants to compare how well they are performing compared to the average participant. Both the leaderboard and level pages can be seen in Figure 3.2.

Procedural Goal Generation The Procedural generation of goals was done using the GOALS system [12]. The GOALS system is an LLM-powered goal augmentation system created to facilitate the procedural generation of measurable goals for use in mHealth applications [12]. In the system, users are given three plain text fields in a form to insert the goals they want to work on in the following lifestyle dimensions: mental, social, and physical. Once the user has inserted and submitted their goals through the form, the GOALS system uses LLMs to generate three measurable goals of increasing difficulty for the user to work on. The user then has the option to regenerate or edit the generated goals before submitting the final measurable goals to an mHealth application. The GOALS system instructs the LLMs to generate goals in a structure based on the GST and is instructed to keep the SDT in mind.

Predetermined Goal Creation To compare LLM-generated goals to non-generated goals, predetermined goals were created for the participants partaking in the intervention. The predetermined goals were created by surveying (n=18) students on their social, physical, and cognitive, new year's resolution goals. Students were instructed to create three goals per dimension, each goal with an increasing level of difficulty. The goals created by the students were then aggregated by the researchers, into three increasingly difficult goals per dimension, totaling nine goals. The students involved in taking this survey were not part of the intervention phase of this study. The list of predetermined goals given to participants can be found on Figshare [11].

Goal Structure Both LLM-generated and pre-determined goals were structured the same, including a description, frequency, points, and a step-by-step plan for achieving the goal. An example of a goal used in the intervention is the following: "Description: Plan an activity with your family/friends, Frequency: 1x within five days, Points: 5, Step-by-step plan: 1. Identify a friend to call, 2. Set a date and time for the call, 3. Prepare topics to discuss, 4. Call the friend.". In each of the three levels in the mHealth application, participants in both groups received three goals related to the three lifestyle dimensions.

3.3 Recruitment

Participants were recruited through convenience sampling, wherein the researchers recruited participants in person on the campus at a university through flyers, banners, posters, and directly inquiring individuals to join the health intervention. Each individual addressed, was given an information flyer with a QR code they could use to schedule an appointment with the researchers for additional information about the study. The inclusion criteria used for recruitment were that participants needed to be older than eighteen and were either students or staff members at a university in the Netherlands. Participants were recruited over a period of two weeks in December 2023.

3.4 Measurements

System data To measure the engagement of participants taking part in the health intervention the mHealth application measured 1) the active engagement of the user through user self-reported activities, and 2) the passive engagement through user navigation through the application.

To analyze the impact of the LLM-generated goals on the engagement level of participants. A statistical comparison between the engagement of both groups of participants was conducted. Analysis was performed on active, passive, and total user engagement (i.e., the combination of passive and active engagement). First, the user engagement data was cleaned: 1) outliers were removed, 2) dropouts were removed, and 3) mean values of the activities of each user were calculated for each phase of the intervention. Second, both the Shapiro-Wilk and Anderson-Darling tests were conducted on the data to determine if the data fit a normal distribution. Lastly, two independent t-tests were conducted to measure the statistical significance between the groups. A p-value of 0.05 was used to determine significance.

Survey data To measure the perceived intrinsic motivation of the users, participants were asked to complete the Intrinsic Motivation Inventory (IMI) survey. The IMI survey assesses participants' subjective experience through self-reporting using a 5-point Likert scale [22]. Participants were asked to complete the IMI three times: 1) Before the start of the intervention, 2) During the break period, and 3) After the end of the intervention. The IMI survey was used to measure the perceived intrinsic motivation of the participants. The elements measured during the study were: 1) enjoyment, 2) competence, 3) relatedness, 4) choice.

To analyze the perceived intrinsic motivation of participants. A statistical analysis was conducted comparing the two groups across the three measurement periods. The means of the groups were calculated for each IMI measurement period (i.e., before the intervention, during the break period, and after the intervention). A one-way ANOVA test was performed to determine the statistical significance between the two groups. A p-value of 0.05 was used to determine statistical significance.

Interview data To gather additional feedback and to capture user perceptions of the mHealth application, and more particularly the goals given to them for the intervention, interviews were conducted at the end of the intervention period.

To analyze the interview data, each interview was recorded and transcribed. Thematic analysis was then conducted on the transcription data to analyze the feedback given by the participants.

4 Results

4.1 User data

A total of 73 participants were recruited to take part in the study. All participants signed informed consent, were randomly allocated to one of the research groups, received their credentials to log into the mHealth application, and received their goals for the intervention. Participants are considered active participants when they have performed at least one active activity in both phases of the intervention. There were a total of 36 participants in the group given LLM-generated goals of which 19 participants were active, and a total of 37 participants in the group given predetermined goals of which 20 participants were active. After the intervention phase, 15 interviews were conducted, with 8 interviewees from the group given LLM-generated goals, and 7 interviewees from the group given predetermined goals.

4.2 System data

The active participants in the group given predetermined goals performed a total of 1216 activities within the mHealth application, of which 172 of those were active, and 1044 were passive activities. During the first phase of the intervention, this group of participants performed a total of 836 activities, whereby 109 of which were active and 727 of them were passive activities. In the second phase of the intervention, the participants performed a total of 380 activities, whereby 63 of the activities were active and 317 of them were passive.

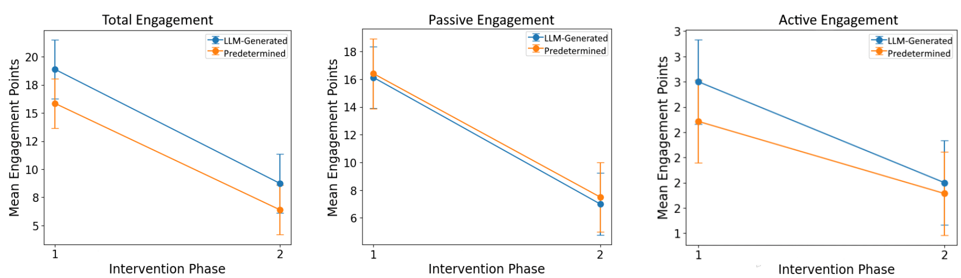

Fig. 2. The left figure displays the mean total engagement points measured for both groups of participants. The middle figure shows the mean passive engagement measured for both groups. The right figure showcases the mean active engagement measured for both groups.

Participants in the group given LLM-generated goals performed a total of 1161 activities within the mHealth application. Which 184 were active, and 977 were passive activities. In the first phase of the intervention, this group performed a total of 765 activities, whereby 113 of them were active and 652 were passive. In the second phase of the intervention, the participants did a total of 396 activities, whereby 71 were active activities and 325 were passive activities

Overall, the total engagement dropped between the two phases for both groups. As indicated in Figure 4.2, the total mean engagement of participants in the group given LLM-generated goals was higher in both phases of the intervention. The same is observed when comparing the mean active engagement.

However, also displayed in Figure 4.2, the mean passive engagement was slightly higher in the group given predetermined goals.

When conducting statistical comparisons of the overall engagement between both groups during phase one of the intervention, the p-value was found to be 0.38. Specifically, for active engagement, the p-value was also 0.38, while for passive engagement, it was 0.8. During the second phase of the intervention, the p-value of the total engagement was 0.36, for active engagement, the p-value was 0.2, and for passive engagement, the p-value was 0.17. According to the statistical comparison, there was no significant difference found in the total, active, or passive engagement between the two groups of participants. Indicating that despite, the overall engagement being higher in the group receiving LLM-generated goals, the difference in engagement between the groups is minimal.

4.3 Survey data

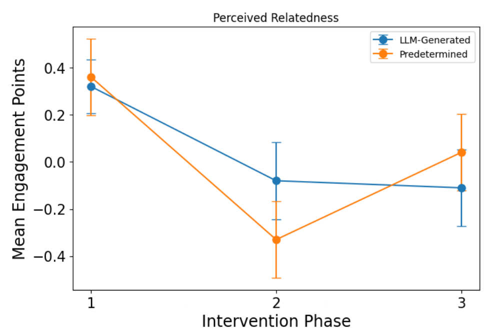

Fig. 3. The figure below displays the mean perceived relatedness measured of participants in both groups, at three points of the intervention.

All 73 participants completed the initial IMI survey before the start of the intervention. Only 22 participants filled in the second survey, given during the break period of the intervention. At the end of the intervention, 19 participants completed the last IMI survey. The perceived enjoyment was initially, slightly higher in the group receiving predetermined goals, however after using the application, the group with LLM-generated goals scored higher both during the break period and after the intervention. The same trend is observed in perceived competence and perceived choice. However, as displayed in Figure 4.3 participants in the group with predetermined goals measured relatedness higher in the phases before and after the intervention than those in the group with LLM-generated goals. The results of the statistical analysis conducted revealed the following p-values: Enjoyment = 0.77, Competence = 0.65, Relatedness = 0.82, and Choice = 0.94. These findings suggest that there was no statistical significance observed when comparing the self-reported intrinsic motivation of participants from both groups based on the survey data.

4.4 Thematic analysis

The results of the thematic analysis of the interviews identified three main themes: 1) goal structure, 2) goal relevance, and 3) goal refinement.

Goal structure Both groups of participants expressed satisfaction with the structure of the goals. Participant 8 noted that "[The GOALS System] gave me ideas on how to structure my goals," while Participant 9 appreciated the diverse range of goal categories provided by the application, indicating, "It's not just focused on one aspect; it was a wide range of categories, including sports, work, and social interactions, which strongly relates to my real-life activities." Similarly, Participant 7 remarked, "[The dimensions were] nice it covered different areas of activities, being physically active and socially and cognitively active"

Goal relevance Both groups of participants were overall mixed about how relevant the goals provided in the mHealth application were to their personal new year's resolution goals they wanted to work on.

Highlighting the positives, participants in the group receiving predetermined goals, Participant 11 regarded the goals positively, viewing them as "stepping stones for achieving bigger goals", while Participant 7 found the goals "pretty useful.". Likewise, in the group receiving LLM-generated goals, Participant 15 found the activities to be "good and relevant,", Participant 12 considered the goals to be "respectable [and] genuinely nice things to do," and Participant 9 echoed that sentiment by stating that "the goals were relevant and interesting, based on personal interest."

Nevertheless, in both groups, certain participants failed to relate to the goals provided by the mHealth applications. Within the group receiving LLM-generated goals, Participant 14 perceived that the "goals generated were a bit too ambitious," and Participant 8 felt the "Generated goals were a bit distant from what I achieved." Similarly, participants in the group receiving predetermined goals Participant 1 expressed that "some tasks were pretty generic" and suggested they "could be personalized a little bit.", Participant 4 stated that the goals "did not resonate with me" while Participant 5 remarked that "The personalization was missing" this was echoed by Participant 10, who indicated that "[predetermined] goals were not personalized enough"

Goal refinement Participants in both groups mentioned some aspects of goals that need refinement over time. The participants in the group receiving LLM-generated goals enjoyed the goals however as Participant 8 states, the generated goals "needed quite some refinement." and are "either too broad or they're all in a kind of similar direction." the latter statement being echoed by Participant 2 stating "I felt like I always saw quite similar goals or maybe even the same goals". Additionally, Participant 14 mentioned "I would have preferred it if you have daily goals that you can kind of tick the box" In the group receiving predetermined goals, Participant 4 noted, "Some of the goals felt a bit too broad to me," while Participant 12 stated, "I couldn't do some of the goals because they were too difficult." Participant 15 found fault with the duration of certain goals, mentioning, "The duration of some goals wasn't suitable for me." Participant 5 emphasized the importance of the ability to "adjust elements of goals" over time, which was echoed by Participant 9, who suggested, "Maybe you can change to another format or switch to another activity."

5 Discussion

5.1 Principle findings

In this study, we evaluated the impact of LLM-generated goals on the engagement of participants in health interventions using gamified mHealth applications. We evaluated no significant difference in participant engagement with the mHealth application when participants were given LLM-generated goals compared to predetermined goals. There was also no statistically significant impact evaluated on the perceived intrinsic motivation of participants between the two groups. The results indicate that LLMs may be a viable option for researchers for time-efficient scalable personalized goal generation for participants in gamified mHealth applications. Participants in both groups were evenly mixed on the goals they received for the intervention. Participants stated that this is due to the goals not being relevant, personal, or refined enough. Overall participants indicated that the goals and several elements of the goals were too static or not specific enough. These results indicate that LLM-generated goals are not particularly better than predetermined goals, however, it does indicate that due to the time efficiency of generating goals, LLMs could possibly be used for efficient goal creation for health interventions. They could also possibly be leveraged for quick and dynamic goal iteration and over-time goal refinement.

5.2 Limitations

This study was conducted with students and staff members from universities in the Netherlands thus, the findings may not be generalizable beyond this specific population. We compared pre-determined, one-size-fits-all goals to personalized goals created by users using LLMs as a tool. The pre-determined goals lacked personalization. It is therefore important to note that this study did not compare LLM-generated personalized goals with handcrafted personalized goals. Although the literature suggests healthcare professional supervision for LLM-generated content in health interventions, it was not implemented in this study, nor for the pre-determined goals. Moreover, participants were not asked for their opinions on sharing personal data with LLM tools for highly personalized goals during interviews [30]. While user input is incorporated into the prompt of the current GOALS system, it utilizes one prompt engineered for generating personalized goals, potentially limiting the personalization of goals for participants. Lastly, the scope of the study was to evaluate the user engagement with the application, not the effectiveness of the health intervention in preventing NCDs.

5.3 Future work

This study must be replicated with a different study population to strengthen the finding that LLM-generated goals have no negative impact on user engagement and are more efficient for health interventions when compared to manually crafted goals. This would broaden our understanding of the generalizability of the findings. Further research should also compare hand-crafted personalized goals to LLM-generated personalized goals on a larger scale. It should also research the effectiveness of LLM-generated goals in the prevention of NCDs, by conducting long-term studies incorporating LLM-generated goals. Another promising direction to explore is to leverage the quick regeneration of goals, and prompt engineering techniques, to dynamically refine goals and the elements of the goals. The elements of goals that are recommended to be refined are 1) difficulty, possibly through altering the frequency of tasks individuals need to do within a goal, 2) variety, possibly through changing the type of goals individuals need to do avoiding the goals being too similar to each other, and 3) daily goals, by providing a new daily goal an individual can work toward [19].

6 Conclusion

In this study, we explored the use of LLM-generated content for scalable goal creation in health interventions by evaluating the impact of LLM-generated goals on the engagement of participants in health interventions, using mHealth applications. The results indicate no significant impact on engagement in health interventions when participants are given LLM-generated goals, compared to predetermined goals. Due to there being no statistical significance in the engagement of participants in the health intervention, LLM-generated content, particularly LLM-generated goals may be an efficient way of creating scalable content for mHealth applications with minimal to no loss of participant engagement. Using LLM tools may also be a time saver when involving health professionals or coaches in the process of creating personalized content for their clients involved in health intervention studies. However, it is important to note that the results of this study are not generalizable to other populations in different contexts. Further research should be done to determine the effectiveness of LLM-generated goals in the prevention of NCDs.

Acknowledgments The authors would like to thank the participants, and the GameBus team with a special thanks to Alireza Khanshan, Andree Vela, Mohammed Irfan Khalid, Suzan Evers, Catarina de Oliveira, and Mengyan Guo for bringing this paper together.

Disclosure of Interests. The authors have no competing interests to declare relevant to this article's content.

References

1. Abd-Alrazaq, A., AlSaad, R., Alhuwail, D., Ahmed, A., Healy, P.M., Latifi, S., Aziz, S., Damseh, R., Alrazak, S.A., Sheikh, J., et al.: Large language models in medical education: opportunities, challenges, and future directions. JMIR Medical Education **9**(1), e48291 (2023)
2. Balwan, W.K., Kour, S., et al.: Lifestyle diseases: The link between modern lifestyle and threat to public health. Saudi J Med Pharm Sci **7**(4), 179–84 (2021)
3. Budreviciute, A., Damiati, S., Sabir, D.K., Kodzius, R.: Management and prevention strategies for non-communicable diseases (ncds) and their risk factors. Frontiers in public health **8**, 574111 (2020)
4. Deterding, S., Sicart, M., Nacke, L., O'Hara, K., Dixon, D.: Gamification. using game-design elements in non-gaming contexts. In: CHI'11 extended abstracts on human factors in computing systems, pp. 2425–2428 (2011)
5. Eberle, C., Löhnert, M., Stichling, S., et al.: Effectiveness of disease-specific mhealth apps in patients with diabetes mellitus: scoping review. JMIR mHealth and uHealth **9**(2), e23477 (2021)
6. El-Hilly, A.A., Iqbal, S.S., Ahmed, M., Sherwani, Y., Muntasir, M., Siddiqui, S., Al-Fagih, Z., Usmani, O., Eisingerich, A.B.: Game on? smoking cessation through the gamification of mhealth: A longitudinal qualitative study. JMIR serious games **4**(2), e5678 (2016)
7. Fogg, B.J.: A behavior model for persuasive design. In: Proceedings of the 4th international Conference on Persuasive Technology. pp. 1–7 (2009)
8. Hueso, M., Álvarez, R., Marí, D., Ribas-Ripoll, V., Lekadir, K., Vellido, A.: Is generative artificial intelligence the next step toward a personalized hemodialysis? Revista de investigación clínica **75**(6), 309–317 (2023)
9. Jakob, R., Harperink, S., Rudolf, A.M., Fleisch, E., Haug, S., Mair, J.L., Salamanca-Sanabria, A., Kowatsch, T.: Factors influencing adherence to mhealth apps for prevention or management of noncommunicable diseases: systematic review. Journal of Medical Internet Research **24**(5), e35371 (2022)
10. Jameel, L., Valmaggia, L., Barnes, G., Cella, M.: mhealth technology to assess, monitor and treat daily functioning difficulties in people with severe mental illness: A systematic review. Journal of psychiatric research **145**, 35–49 (2022)
11. James, L.: List of pre-determined goals (7 2024). https://doi.org/10.6084/m9.figshare.26186609.v1
12. James, L., Genga, L., Montagne, B., Hagenaars, M., Van Gorp, P.: Caregiver's evaluation of llm-generated treatment goals for patients with severe mental llnesses. In: International Conference on Pervasive Technologies Related to Assistive Environments. Association for Computing Machinery, Inc (2024)
13. James, L., Maessen, M., Genga, L., Hagenaars, M., Montagne, B., Van Gorp, P.: Towards augmenting mental health personnel with llm technology to provide more personalized and measurable treatment goals for smi patients. In: Pervasive Health conference (2023)
14. Kankanhalli, A., Xia, Q., Ai, P., Zhao, X.: Understanding personalization for health behavior change applications: A review and future directions. AIS Transactions on Human-Computer Interaction **13**(3), 316–349 (2021)
15. Lo, L.S.: The art and science of prompt engineering: a new literacy in the information age. Internet Reference Services Quarterly **27**(4), 203–210 (2023)
16. Locke, E.A., Latham, G.P.: Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. American psychologist **57**(9), 705 (2002)

17. Memmert, L., Cvetkovic, I., Bittner, E.: The more is not the merrier: Effects of prompt engineering on the quality of ideas generated by gpt-3 (2024)
18. Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., Eccles, M.P., Cane, J., Wood, C.E.: The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Annals of behavioral medicine **46**(1), 81–95 (2013)
19. Nuijten, R., Van Gorp, P., Khanshan, A., Le Blanc, P., van den Berg, P., Kemperman, A., Simons, M., et al.: Evaluating the impact of adaptive personalized goal setting on engagement levels of government staff with a gamified mhealth tool: results from a 2-month randomized controlled trial. JMIR mHealth and uHealth **10**(3), e28801 (2022)
20. Ogbeiwi, O.: Why written objectives need to be really smart. British Journal of Healthcare Management **23**(7), 324–336 (2017)
21. Orji, R., Nacke, L.E., Di Marco, C.: Towards personality-driven persuasive health games and gamified systems. In: Proceedings of the 2017 CHI conference on human factors in computing systems. pp. 1015–1027 (2017)
22. Reynolds, J.L.: Intrinsic Motivation Inventory (IMI) Scale (DUTCH). Handbook of Research on Electronic Surveys and Measurements pp. 1–6 (2006)
23. Ryan, R.M., Deci, E.L.: Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American psychologist **55**(1), 68 (2000)
24. Schwarzer, R.: Modeling Health Behavior Change: How to Predict and Modify the Adoption and Maintenance of Health Behaviors. Applied Psychology **57**(1), 1–29 (2008). https://doi.org/10.1111/j.1464-0597.2007.00325.x, https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1464-0597.2007.00325.x
25. Short, C.E., DeSmet, A., Woods, C., Williams, S.L., Maher, C., Middelweerd, A., Müller, A.M., Wark, P.A., Vandelanotte, C., Poppe, L., et al.: Measuring engagement in ehealth and mhealth behavior change interventions: viewpoint of methodologies. Journal of medical Internet research **20**(11), e292 (2018)
26. Sundar, S.S., Marathe, S.S.: Personalization versus customization: The importance of agency, privacy, and power usage. Human communication research **36**(3), 298–322 (2010)
27. Tabish, S., et al.: Lifestyle diseases: consequences, characteristics, causes and control. J Cardiol Curr Res **9**(3), 00326 (2017)
28. Van Gorp, P., Nuijten, R.: 8-year evaluation of gamebus: Status quo in aiming for an open access platform to prototype and test digital health apps. Proceedings of the ACM on Human-Computer Interaction **7**(EICS), 1–24 (2023)
29. White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., Schmidt, D.C.: A prompt pattern catalog to enhance prompt engineering with chatgpt. arXiv preprint arXiv:2302.11382 (2023)
30. Willms, A., Liu, S.: Exploring the feasibility of using chatgpt to create just-in-time adaptive physical activity mhealth intervention content: Case study. JMIR Medical Education **10**, e51426 (2024)
31. World Health Organization: Noncommunicable diseases (ncds) - fact sheet. https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases (2023), accessed on April 22, 2024
32. Xu, L., Shi, H., Shen, M., Ni, Y., Zhang, X., Pang, Y., Yu, T., Lian, X., Yu, T., Yang, X., et al.: The effects of mhealth-based gamification interventions on participation in physical activity: systematic review. JMIR mHealth and uHealth **10**(2), e27794 (2022)