Degree Project in Medical Engineering

Second Cycle, 30 credits

Co-Designing a Mobile Health Application to Promote Positive Body Image Among Athletes: A Mixed Method Study

LISA ALMSTEDT

This page intentionally left blank.

Co-Designing a Mobile Health Application to Promote Positive Body Image Among Athletes: A Mixed Method Study

Lisa Almstedt

This page intentionally left blank.

Abstract

Athletes face unique challenges related to body image, shaped by sport-specific demands and sociocultural norms. Mobile health applications have shown potential in promoting positive body image, but existing solutions are often designed for the general population, and sustaining user engagement in these apps remains a challenge. This study aimed to design and evaluate a mobile application to promote a positive body image, specifically tailored for athletes, incorporating a large language model (LLM)-based AI agent, TrueBalance, designed in previous research. A user-centered co-design process was employed, including a literature review, prototype development, a co-design workshop with four athletes, and final testing of the completed application. Participants favored features that support reflection, autonomy, and emotional well-being, and expressed skepticism toward point systems and competitive elements. As a result, features such as journaling, daily challenges, and a simplified interface were prioritized. Final usability testing indicated that the app was perceived as easy to use and supportive during emotionally challenging periods. The study highlights the importance of user-specific and user-centered design in digital interventions. Favoring a non-human appearance for the avatar representing the conversational agent, along with non-competitive and personalized features, users indicated that such design choices may improve engagement, support autonomy, and reduce social pressure in mental health tools for athletes.

Keywords: Body image, athletes, mHealth, conversational AI, co-design, user engagement

This page intentionally left blank.

Sammanfattning

Atleter upplever unika utmaningar relaterade till kroppsbild, kopplade till idrottsspecifika krav och sociokulturella normer. Mobila hälsoappar har potential att främja en positiv kroppsbild, men befintliga lösningar är ofta utformade för den generella befolkningen och att behålla användarengagemang över tid är fortsatt en utmaning. Denna studie syftade till att designa och utvärdera en mobilapplikation för att främja en positiv kroppsbild, specifikt utformad för atleter, med integration av den LLM-baserade AI-agenten TrueBalance. En användarcentrerad co-designprocess tillämpades och inkluderade en litteraturöversikt, prototyputveckling, en co-designworkshop med fyra atleter samt utveckling och slutlig testning av den färdiga applikationen. Deltagarna föredrog funktioner som stödjer reflektion, autonomi och emotionellt välbefinnande, och uttryckte skepsis mot poängsystem och tävlingsinriktade inslag. Därför prioriterades funktioner såsom dagboksskrivande, dagliga utmaningar och ett förenklat gränssnitt. Den avslutande testningen visade att appen upplevdes som lättanvänd och stödjande under emotionellt utmanande perioder. Studien betonar värdet av användarspecifik och användarcentrerad design i digitala interventioner. Preferensen för icke-konkurrensinriktade, personaliserade funktioner samt för ett icke-mänskligt utseende på samtalsagenten tyder på att sådana verktyg kan öka engagemanget, stärka autonomin och minska social press i idrottsrelaterade applikationer för mental hälsa.

Nyckelord: Kroppsbild, atleter, mobil hälsa, konverserande AI, co-design, användarengagemang

This page intentionally left blank.

Acknowledgments

This page intentionally left blank.

CONTENTS

This page intentionally left blank.

1 Introduction

In recent years, body image has gained attention as a key component of mental health, both in the general population and within athletic communities [1], [2]. Athletes face distinctive pressures related to performance, appearance, and sport-specific expectations [3], which may contribute to disordered eating, low self-esteem, and reduced psychological well-being [1]. While mobile health (mHealth) technologies offer promising opportunities for mental health support [4], long-term user engagement remains a challenge [5], and most current applications are not tailored to the specific needs of athletic populations. This thesis explores how digital tools can be designed to promote positive body image and emotional well-being in athletes through a user-centered approach. The work focuses on the design and evaluation of the TrueBalanceApp, a mobile application that incorporates a previously researched large language model (LLM)-based conversational AI agent called TrueBalance [6]. Through a mixed methods and co-design approach involving athletes, the study examines user needs, engagement strategies, and perceptions of trust and emotional support in digital interventions. The findings aim to contribute to athlete-centered digital mental health and inform the future design of inclusive, relevant, and supportive technologies.

2 Background

Body image refers to an individual's perceptions, thoughts, and feelings about their physical appearance. It encompasses both how one sees their body and how one believes others perceive it, and it plays a significant role in psychological functioning and self-worth [7]. Among athletes, body image concerns are shaped by a complex interplay of coaching practices, peer comparison, and sport-specific appearance norms [2], [3].

While athletes may exhibit high body awareness and discipline [2], they are also at elevated risk for body dissatisfaction and disordered eating, especially in sports that emphasize leanness, aesthetics, or weight categories [1], [3], [4]. Distorted body image has been associated with anxiety, depression, and lower psychological well-being, whereas positive body image is linked to greater resilience and self-acceptance [1].

Several digital mental health tools and mobile health (mHealth) applications have emerged as scalable tools to support body image and emotional well-being. These include general-purpose apps that offer, among other things, mindfulness training, cognitive behavioral therapy exercises, or mood tracking [4], [8]. Examples such as Headspace provide mental resilience content, while others focus on fitness or nutrition tracking [9]. However, these tools are not specifically designed for the psychological needs of athletes, including performance pressure, sport-specific appearance standards, and the balance between health and aesthetics [2], [3]. Moreover, sustaining long-term user engagement in mHealth apps remains a challenge [5].

To address this, researchers have explored engagement-enhancing strategies such as gamification [10], behavioral nudging [11], and conversational AI [4], [5]. Although features like gamification and conversational agents are increasingly common in general mHealth interventions [5], their relevance and effectiveness for athlete populations remain insufficiently explored. Open-domain conversational AI, which enables free-text dialogue rather than pre-scripted responses, has shown potential to support reflection and emotional connection, but its use in athlete-focused mental health tools is still limited.

This thesis builds on prior work involving TrueBalance, a large language model (LLM)–based conversational AI agent that integrates biomedical determinants of mental health and principles from cognitive behavioral therapy (CBT) to promote positive body image in young adults [6]. By applying human-centered design (HCD) principles [12] and involving athletes directly in the design process, this study contributes to athlete-centered digital health research and aims to inform the development of more relevant, engaging, and psychologically supportive technologies.

3   Research Methodology

This section outlines the methodological approach used to design, develop, and evaluate the TrueBalanceApp. It presents the design of the research questions, the theoretical foundations guiding the study, the mixed-methods approach adopted, and the procedures for co-design, prototype development, testing, and analysis.

3.1   Design of Research Questions

The research originated from an interest in the intersection of body image, mental health, and digital interventions, particularly within athletic populations. While digital mental health tools have gained traction in the general population, there appeared to be a lack of tailored solutions for athletes facing body image pressures. A preliminary literature review confirmed this gap, especially concerning engagement challenges in existing mHealth applications and the limited exploration of athletes' perspectives on features like gamification, nudging, and AI interaction.

Through this review process, the research focus was refined to emphasize the design, development, and evaluation of a mobile application, the TrueBalanceApp, which integrates a previously developed LLM-based conversational AI agent (TrueBalance) alongside engagement-enhancing features aimed at supporting positive body image in athletes.

While several potential research directions were considered, including those centered on technical implementation and algorithmic performance, the study prioritized questions related to user perception, emotional engagement, and trust in human-like AI elements. This emphasis aligned more closely with the overarching goal of researching how to develop a human-centered, user-driven intervention.

Moreover, by focusing on the design aspects of the application rather than its technical backend alone, the study aimed to generate insights that could inform future athlete-specific digital health applications more broadly. In this way, the research intended to contribute not only to the development of the TrueBalanceApp itself, but also to the wider field of athlete-focused digital mental health design.

This resulted in three research questions:

1. How do athletes perceive and engage with digital interventions aimed at supporting a positive body image?
2. Which features and engagement strategies are perceived to enhance emotional support and sustained use in a body image–focused mHealth app for athletes?
3. How do athletes experience the role of human-like elements (e.g., avatars, voice interaction) in building trust and comfort during digital interactions?

These questions aimed to address both the design and experiential aspects of digital health interventions, with a focus on understanding functional usability as well as user perceptions and acceptability.

3.2 Background Framework

The literature review process was guided by a focus on two intersecting areas: body image concerns in adults and athletes, and digital interventions for mental health, particularly mHealth applications. Structured searches were conducted using the databases PubMed, Web of Science, Google Scholar and Scopus, with keyword combinations including:

- *("adults" OR "athlete*") AND ("body image") AND ("satisfaction")*
- *("mental health") AND ("mobile app*" OR "mHealth") AND ("feature*")*
- *("mHealth") AND ("engagement") AND ("body image")*

Searches were limited to publications from the past five years to ensure relevance to recent advancements in both psychological research and digital health technologies.

In addition to database searches, snowballing was used to identify relevant articles through reference lists of key publications. Targeted searches were also performed for specific concepts such as gamification in mHealth, behavioral nudging, and AI-driven chatbots, particularly when initial database searches produced limited or overly broad results.

While an exact number of initial search results was not recorded, relevance was determined through title and abstract screening, followed by full-text review for conceptual alignment with the study's focus. This approach resulted in a focused selection of sources that informed the study design, including the choice of features evaluated and the methodological emphasis on user experience and engagement in athlete populations.

3.3 Methodology Approach

A mixed-methods, human-centered design (HCD) approach was chosen to ensure that the intervention was theoretically sound and aligned with user expectations and needs. The research process included iterative prototyping, a co-design workshop, development of the application, and final usability testing, reflecting a commitment to participatory design.

3.3.1 Co-Design Workshop

To ground the app in real-world experiences, a co-design workshop was conducted with four athletes representing diverse sport disciplines and competition levels. The workshop included feature prioritization, interactive prototype testing, and visual design feedback. This method was selected over more rigid surveys to allow for rich, qualitative insight and participant-led exploration of ideas.

The co-design session explored athletes' prior use of mental health apps, attitudes toward gamified elements, and trust in AI-based features such as the conversational agent and voice input. A "feature auction" helped identify which features participants valued most, while Miro boards supported collaborative real-time ideation. This method allowed the research to go beyond abstract preferences and examine why certain features mattered to users emotionally and practically.

3.3.2 Usability and Functional Testing

Following app development, the same participants engaged in 24-hour real-world testing. They interacted with the full version of the app and provided feedback through the mHealth App Usability Questionnaire (MAUQ) and semi-structured interviews. This phase focused on capturing usability, technical performance, and subjective user experience.

A survey-only approach was considered, however, it was deemed insufficient for capturing the depth of experiential feedback required to inform both emotional and functional aspects of the app. Given the limited number of participants, interviews were incorporated to complement the quantitative data and provide richer qualitative insights.

3.3.3 Alignment with Research Questions

Each methodological component was selected to address specific aspects of the research questions:

- The co-design workshop was used to explore athletes' experiences and expectations related to digital tools, particularly in relation to body image support (RQ1 & RQ2).
- Discussions on gamification and nudging were included to investigate how these elements were interpreted and their perceived relevance within an mHealth context (RQ2).
- The evaluation of avatar and voice interaction features aimed to examine perceptions of human-like elements and their potential to foster trust and user comfort (RQ3).
- Real-world testing over a 24-hour period, combined with follow-up interviews, provided data on usability, perceived engagement, and anticipated patterns of user interaction over time, based on short-term experience with the app. These insights contributed to all research

questions.

3.3.4   Methodological Limitations

Several limitations should be acknowledged:

- The co-design workshop included a small, non-random sample recruited through convenience sampling. While rich in detail, findings may not be directly generalized to broader athlete populations.
- Usability testing was conducted over a short time frame (24 hours), limiting insights into long-term engagement.
- The evaluation did not include clinical validation of the AI assistant's responses, as the study focused on design and experience.
- While the codebase supports both platforms, the user testing was conducted solely on Android devices.

Despite these limitations, the methodology provided a robust foundation for understanding how digital design choices affect user engagement, emotional trust, and perceived support among athletes.

4 Summary of Article

This section summarizes the main findings and insights from the appended research article. It presents the results of the co-design and user testing phases, followed by a discussion of their implications for athlete-centered mHealth design.

4.1 Results

The article presents three key findings that can inform the design of athlete-centered digital interventions for body image and mental well-being. First, while gamification and nudging are widely used to increase engagement in mHealth apps, the athletes in this study expressed skepticism toward point systems and streak-based features, associating them with pressure and performance anxiety. Instead, participants preferred intrinsically motivating strategies such as open-ended journaling, adaptive challenges, and supportive nudges. This finding contrasts with prior studies on gamification in general populations, suggesting that athletes' competitive mindsets may make certain reward structures counterproductive in mental health contexts.

Second, simplicity and emotional safety were highlighted as essential design elements. Participants described complex or feature-heavy interfaces as overwhelming, especially during periods of stress. This gives an alternative view to engagement strategies that emphasize feature expansion and instead aligns with research suggesting that minimalism can enhance perceived usefulness

and comfort. In this context, simplicity functioned not just as a usability factor, but as a design strategy.

Third, the study showed by the final testing that athletes responded positively to a non-human conversational AI agent, appreciating its lack of judgment and the emotional distance it provided. While the discussion on whether human-like or non-human-like avatars provide comfort and connection continues, this study reinforces emerging evidence that non-human agents can lower barriers to emotional expression.

The study also demonstrated the value of co-design in uncovering user needs that might otherwise go unrecognized. Involving athletes in the design process challenged early assumptions, such as the value of performance-oriented visuals or competitive features, and led to a softer, more reflective interface. These findings contribute to a growing literature on personalized digital health tools and highlight the importance of user-specific design in creating emotionally responsive, sustainable mHealth interventions.

4.2 Discussion

This study highlights the importance of tailoring mHealth tools to the emotional and motivational needs of the user group. While designed with athletes in mind, several findings, such as preferences for simplicity, intrinsic motivation, and emotionally supportive design, may be relevant to broader populations, including adolescents and other individuals who also experience body image pressures. Given the high prevalence of appearance-related concerns among young people, future development could focus on adapting the intervention for this group.

The iterative co-design and user testing process proved essential in surfacing user preferences, such as the clear favoring of a non-human appearance of the conversational agent and minimal, non-demanding interfaces. These results emphasize the value of involving users throughout the whole development process, as needs and perceptions often evolve over time. In particular, the two above stated preferences changed somewhat between the first co-design and the final testing.

At the same time, the integration of LLM-based AI agents in mental health contexts presents important ethical and practical considerations. While users appreciated the non-judgmental and reflective tone of the AI agent, LLMs carry inherent risks, including the potential to generate inaccurate or emotionally inappropriate responses [13]. As this study did not assess clinical accuracy, the app was clearly positioned as a self-reflection tool rather than a therapeutic resource. If future development aims to provide more structured support, collaboration with clinicians will be necessary. Clinician-guided design could enable the integration of goal setting, evidence-based and personalized exercises, or safety checks, as seen in existing platforms offering digital guidance or therapy, such as "Headspace: Sleep & Meditation" [14].

5 Fulfillment of Learning Outcomes

This section reflects on how the thesis meets the formal learning outcomes of the Master thesis course. It addresses the development of scientific and methodological competence, the handling of complex issues, ethical considerations, and the ability to communicate and apply research findings.

5.1 Scientific Knowledge and Methodology

The project demonstrates in-depth understanding of the scientific foundations of body image research, digital mental health interventions, and human-centered design. A literature review was conducted to establish the psychological and social factors influencing body image in athletes, as well as the potential of mHealth interventions. Relevant theories such as Self-Determination Theory (SDT) and techniques from Cognitive Behavioral Therapy (CBT) were integrated into the design process.

Methodologically, the study employed a mixed-methods approach, incorporating a co-design workshop, usability testing, deductive qualitative content analysis, and quantitative feedback gathered through a standardized usability questionnaire. These methods are detailed in the article's methods section under the headings "Co-Design Workshop," "Data Analysis," and "Usability and Functional Testing". The choice of methods reflects current

trends in digital health research and aligns with principles of participatory and user-driven design.

5.2 Systematic Information Gathering

Literature searches were conducted using PubMed, Web of Science, and Scopus, using structured keyword combinations (e.g., "athlete*" AND "body image" AND "satisfaction") and limiting results to the past five years. Additional articles were identified through reference mining and targeted searches on related topics such as gamification and AI-driven chatbots.

This process revealed a gap in the literature concerning athlete-specific digital interventions aimed at addressing body image. This gap helped shape the study's focus and supported the integration of both behavioral and technological engagement strategies, as discussed in the Introduction and further elaborated in the final paragraph of the "Background" section of the article.

5.3 Handling Complex Questions

The project addressed a complex and emerging issue: body image concerns among athletes, particularly within the context of digital self-help tools. It required the integration of psychological, technological, and design perspectives to explore a topic that remains under-researched yet increasingly discussed in both academic and professional settings. Due to the limited availability of prior research on athletes' engagement with AI-based mental health tools, the study relied on qualitative methods, complemented by usability testing of a novel application. The study also faced ethical design questions around voice interaction, avatar design and trust, and personalized nudging, all within a non-clinical yet sensitive health context. These complexities are reflected and discussed in the "Discussion" section, particularly in relation to ethical design decisions, user trust and engagement, and the role of non-human interaction in mental health contexts.

5.4 Planning and Execution

The research was planned and executed over a clearly structured timeline that included phases for literature review, prototyping, co-design, app development, and user testing. While the original plan was adjusted slightly to accommodate technical development, the core phases were completed within the allocated timeframe. The project methodology and development process are described in

the article under the sections "Method" and "Full App Development", and any changes made during the project were clearly reported to maintain transparency.

5.5 Clear Communication and Argumentation

The work was communicated clearly and consistently across multiple formats, including the scientific articles, weekly oral presentations to a research group, co-design facilitation, and documentation. The article presents a logical narrative from problem formulation through design and testing, supported by relevant theory and prior research. Results are contextualized with respect to user needs and engagement strategies, and the rationale behind design choices are discussed.

Throughout the project, findings and ongoing progress were regularly presented to a research group that met weekly, providing opportunities for feedback, discussion, and refinement of the study's direction. Additionally, during the co-design workshop, the researcher introduced the topic, presented the prototype, and facilitated structured feedback sessions with athlete participants. These engagements allowed for continuous improvement of both the intervention and the communication of its purpose.

The article's Discussion section outlines the main arguments, and the overall thesis is structured to be accessible to both academic and non-specialist readers.

Use of AI tools: LocofyAI was used to convert the Figma prototype into a React Native frontend using TypeScript, as described in the "Full app development" section of the "Method" of the article. Additionally, ChatGPT was used during the writing process by submitting individual paragraphs from the thesis and prompting the model to improve language and grammar without altering the original content. All AI-generated content was reviewed, verified, and edited by the researcher to ensure accuracy and appropriateness.

5.6 Scientific, Social, and Ethical Assessments

The study carefully considered the ethical implications of designing app features and applying AI in a mental health context, especially among athletes who may be vulnerable to body image concerns. Potential risks, such as over-reliance on automated support, inaccurate AI responses, and data sensitivity, were addressed both in the app design and in participant information. Social factors, including stigma, accessibility, and inclusivity, were also discussed in relation to

how different types of athletes engage with digital wellness tools. These aspects are addressed in the "Ethical Considerations" section of the article and informed key design decisions, such as the exclusion of competitive gamification features.

5.7   Research and Development Readiness

The project demonstrates readiness for continued work in research and development, particularly within the field of digital mental health technologies. The work involved independently conducting literature reviews, designing, developing and testing a mobile application, and analyzing both qualitative and quantitative data. The iterative design process, along with the incorporation of participant feedback, reflects the ability to operate effectively in evolving, interdisciplinary environments. The Discussion section of the article outlines how the study contributes to future research on athlete-centered mHealth tools and identifies directions for further development and evaluation.

References

[1] J. Jiménez-Morcillo, D. J. Ramos-Campo, S. Rodríguez-Besteiro, and V. J. Clemente-Suárez, "The Association of Body Image Perceptions with Behavioral and Health Outcomes among Young Adults," *Nutrients*, vol. 16, no. 9, p. 1281, Apr. 2024, doi: 10.3390/nu16091281.

[2] Q. Li, H. Li, G. Zhang, Y. Cao, and Y. Li, "Athlete Body Image and Eating Disorders: A Systematic Review of Their Association and Influencing Factors," *Nutrients*, vol. 16, no. 16, p. 2686, Aug. 2024, doi: 10.3390/nu16162686.

[3] J. Borowiec, A. Banio-Krajnik, E. Malchrowicz-Mośko, and A. Kantanista, "Eating disorder risk in adolescent and adult female athletes: the role of body satisfaction, sport type, BMI, level of competition, and training background," *BMC Sports Sci. Med. Rehabil.*, vol. 15, no. 1, p. 91, Jul. 2023, doi: 10.1186/s13102-023-00683-7.

[4] F. Beilharz, S. Sukunesan, S. L. Rossell, J. Kulkarni, and G. Sharp, "Development of a Positive Body Image Chatbot (KIT) With Young People and Parents/Carers: Qualitative Focus Group Study," *J. Med. Internet Res.*, vol. 23, no. 6, p. e27807, Jun. 2021, doi: 10.2196/27807.

[5] A. F. Collier, S. Hagemann, S. B. Trinidad, and M. Vigil-Hayes, "Human-to-Computer Interactivity Features Incorporated Into Behavioral Health mHealth Apps: Systematic Search," *JMIR Form. Res.*, vol. 7, p. e44926, Jun. 2023, doi: 10.2196/44926.

[6] Zhang, X., Zayed, A., Hamrin, J. R., Guneysu, A., & Kuoppamäki, S., "Enhancing Body Image Awareness through Conversations with an AI Agent: Qualitative Study with Young Adults," 2025.

[7] T. L. Tylka and N. L. Wood-Barcalow, "The Body Appreciation Scale-2: Item refinement and psychometric evaluation," *Body Image*, vol. 12, pp. 53–67, Jan. 2015, doi: 10.1016/j.bodyim.2014.09.006.

[8] A. Newton *et al.*, "A Mobile Phone-Based App for Use During Cognitive Behavioral Therapy for Adolescents With Anxiety (MindClimb): User-Centered Design and Usability Study," *JMIR MHealth UHealth*, vol. 8, no. 12, p. e18439, Dec. 2020, doi: 10.2196/18439.

[9] R. Collins *et al.*, "Mobile app use and the mental health of elite athletes," *Sports Psychiatry*, Mar. 2025, Accessed: May 16, 2025. [Online]. Available: https://econtent.hogrefe.com/doi/10.1024/2674-0052/a000104

[10] S. Litvin, R. Saunders, M. A. Maier, and S. Lüttke, "Gamification as an approach to improve resilience and reduce attrition in mobile mental health interventions: A randomized controlled trial," *PLOS ONE*, vol. 15, no. 9, p. e0237220, Sep. 2020, doi: 10.1371/journal.pone.0237220.

[11] H. Auf, J. Dagman, S. Renström, and J. Chaplin, "Gamification and nudging techniques for improving user engagement in mental health and well-being apps," *Proc. Des. Soc.*, vol. 1, pp. 1647–1656, Aug. 2021, doi: 10.1017/pds.2021.426.

[12] T. Farooqui, T. Rana, and F. Jafari, "Impact of Human-Centered Design Process (HCDP) on Software Development Process," in *2019 2nd International Conference on Communication, Computing and Digital systems (C-CODE)*, Mar. 2019, pp. 110–114. doi: 10.1109/C-CODE.2019.8680978.

[13] C. Blease and J. Torous, "ChatGPT and mental healthcare: balancing benefits with risks of harms," *BMJ Ment. Health*, vol. 26, no. 1, Nov. 2023, doi: 10.1136/bmjment-2023-300884.

[14] "Meditation and Sleep Made Simple," Headspace. Accessed: May 19, 2025. [Online]. Available: https://www.headspace.com/

Appendix A

Co-Designing a Mobile Health Application to Promote Positive Body Image Among Athletes: A Mixed Method Study

Author: Lisa Almstedt

Abstract

Background: Athletes face distinct body image pressures stemming from sport-specific demands, performance expectations, and sociocultural influences. Mobile health applications have been proposed as accessible tools to foster body acceptance and reduce the risk of disordered eating. However, sustaining user engagement in these apps remains a challenge. Existing digital interventions are typically designed for general populations, offering limited insight into the psychological and motivational needs of athletes. To improve engagement, strategies such as gamification, behavioral nudging, and AI-driven conversational agents have gained attention. One such intervention is the TrueBalance conversational AI agent, a large language model-based system designed to promote positive body image through open-ended dialogue. While promising, these approaches have yet to be fully evaluated in athlete-specific contexts.

Objective: This study aimed to design and evaluate a mobile application intended to promote a positive body image tailored to the needs of athletes, and integrating the TrueBalance conversational AI agent. This to explore how digital features may support positive body image and sustained engagement among athletes at elite, competitive, and recreational levels.

Methods: A human-centered, co-design approach was used to develop and evaluate the TrueBalanceApp, integrating the TrueBalance AI agent. The design process included four phases: (1) identifying evidence-based engagement features from prior literature; (2) creating an interactive prototype; (3) conducting a co-design workshop with athletes to refine features and interface design; and (4) building the full application and conducting usability and functional testing. Participants (n = 4) trained ≥6 hours/week and represented varying levels of athletic involvement. Workshop data were analyzed using deductive qualitative content analysis. The final testing phase combined qualitative feedback from user interviews with quantitative usability data collected through the mHealth App Usability Questionnaire (MAUQ).

Results: The co-design workshop identified key features supporting reflection and autonomy, including open ended journaling, training logs, daily challenges, and the conversational AI agent. Participants preferred non-competitive engagement strategies and emphasized the importance of a light, intuitive interface. These insights led to removing a gamified point systems, simplifying body image tracking, and including a non-human avatar to reduce performance pressure. The final app included four core features: a chat interface, journaling, a training log, and daily challenges, along with voice interaction, mood tagging, and optional notifications. User testing confirmed the app's ease of use, supportive tone, and perceived usefulness during emotionally challenging periods. MAUQ scores indicated high usability (M = 1.25–2.5 on a 7-point

Likert scale), with participants appreciating the app's simple, non-judgmental design and flexibility for intermittent use.

Conclusions: Co-designing a mobile application with athletes revealed a preference for simple, personalizable, and non-competitive features. The integration of a non-human conversational AI agent was seen as supportive and approachable, suggesting potential to foster engagement and promote positive body image in sport-related mental health contexts.

Keywords: Body image; Athletes; mHealth; Conversational AI; Co-design; User engagement

Introduction

Athletes, particularly young and elite competitors, experience unique body image pressures due to sport-specific demands, sociocultural expectations, and performance-related weight management [1]. These pressures can negatively impact mental health, self-esteem, and overall well-being [2].

The relationship between body image and disordered eating in athletes remains complex [3]. Some studies suggest that athletes, in general, have a more positive body image compared to non-athletes, yet they still engage in unhealthy weight-loss behaviors [3,4]. Others argue that distorted body image is a predictive factor for disordered eating with athletes, particularly in sports where weight and appearance are central [5,6]. Unhealthy behaviours and risk factors include excessive weighing, binge-eating followed by regret, and obsessive calorie tracking [3], highlighting the need for effective interventions to support healthy body image development in athletes. Additionally, research indicates that higher body image positivity is associated with a lower prevalence of eating disorders [2], reinforcing the importance of supportive measures.

Digital mental health interventions, including mobile health (mHealth) applications, offer promising solutions for addressing body image concerns [7]. These apps provide self-guided psychological support, habit-tracking, and personalized interventions [7], yet low engagement remains a challenge [8]. Research suggests that engagement-enhancing strategies such as gamification [9,10], behavioral nudging [11], and AI-driven chatbot interactions [7,8] can improve adherence in mental health applications. Although these features have been studied in mental health applications for the general population, their applicability to athlete-specific body image interventions remains limited. Given the unique psychological and performance-related factors surrounding athletes, further research is needed to explore how these engagement strategies can be adapted to support engagement and effectiveness in this population.

To address this gap, the present study aimed to design the TrueBalanceApp, a mobile application that integrates the TrueBalance conversational AI agent. TrueBalance is a large language model (LLM)–based conversational AI agent developed to promote awareness of positive body image with young adults [12]. The present mixed methods study employed a co-design and user-testing approach to ensure the app is relevant and engaging for its intended users. Participants included athletes engaged in structured sport training, defined as a minimum of six hours of training per week. The sample represented a range of competitive levels, from national-level athletes to recreational individuals focused on personal fitness. Although recreationally active individuals may not compete at elite levels [3], participation in performance-focused fitness culture and exposure to sport-specific ideals may still contribute to body image pressures. By incorporating this diversity, the study intended to capture varied experiences within structured training environments and the performance-driven fitness culture. Specifically, the study addresses three key research questions:

1. How do athletes perceive and engage with digital interventions aimed at supporting a positive body image?
2. Which features and engagement strategies are perceived to enhance emotional support and sustained use in a body image–focused mHealth app for athletes?
3. How do athletes experience the role of human-like elements (e.g., avatars, voice interaction) in building trust and comfort during digital interactions?

A co-design workshop was conducted with athletes to gather feedback on features and engagement mechanisms, with a focus on user preferences, perceptions of avatars, and perspectives on human-like elements such as voice integration. Following this phase, the app was developed and underwent usability and functional testing to evaluate its effectiveness, ease of use, and engagement. By integrating athlete perspectives throughout the design process, the study aimed to ensure the resulting digital intervention was user-driven, contextually relevant, and tailored to the specific needs of athletes.

Figure 1–3 below illustrate parts of the final application interface, including the homepage with core features, the chat interface, and the daily challenge feature. These elements are described in detail in the results section.
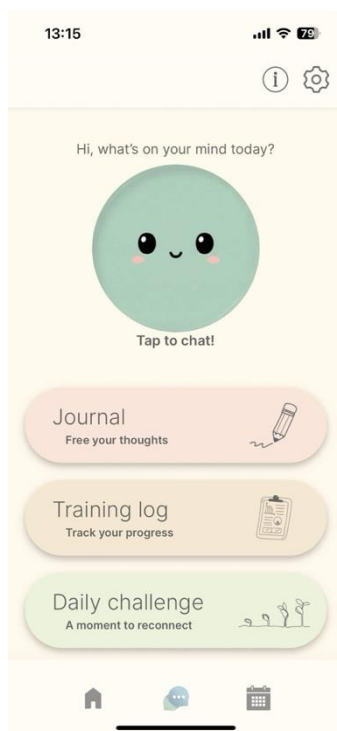

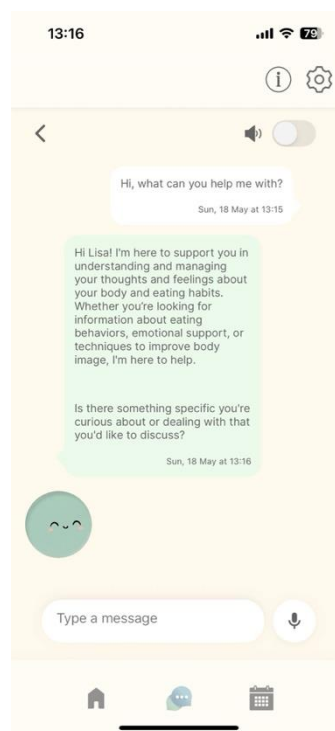*Figure 1. Homepage of the TrueBalanceApp showing core features*
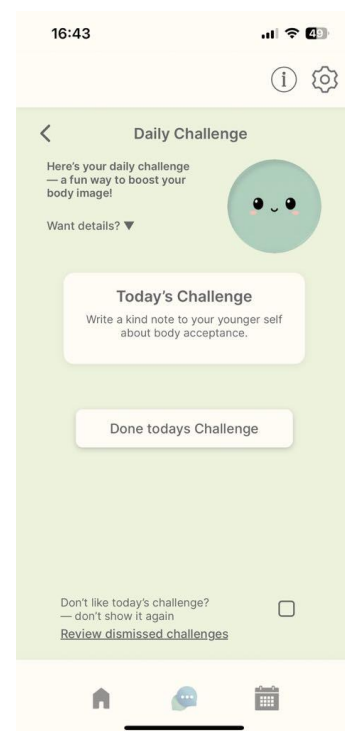
Figure 2. Chat-interface of the TrueBalanceApp

Figure 3. The Daily Challenge interfacace of the TrueBalanceApp

Background

Athletes face pressures regarding body image from multiple sources, including influences from coaches, social media, and peer comparisons, which can contribute to disordered eating behaviors and body dissatisfaction [1,3,13]. Research indicates that young female athletes report higher body dissatisfaction and risk for eating disorders than adults, especially those in lean sports (e.g., endurance running) [1,14], and aesthetic sports (e.g., gymnastics and figure skating) [14]. Male athletes also face body image concerns, particularly in sports that emphasize muscularity and leanness (e.g., team sports, endurance events, weight-category competitions) [13,15]. This is an issue since distorted body image is associated with an increased risk of anxiety, depression, and body dysmorphic disorders, while on the other hand a positive body image is linked to greater psychological resilience and well-being [2].

Mobile applications are widely used by athletes, however, their use is primarily focused on performance tracking, and physical monitoring. While there are some apps focusing on mental resilience and the competitive mindset, mHealth apps remain underutilized in this group, with studies showing inconsistent engagement and limited impact on well-being [16]. This highlights a gap in tools designed specifically to address the psychological needs of athletes, including body image concerns.

Cognitive Behavioural Therapy (CBT) remains a leading approach for addressing body image concerns and eating disorders [7,17]. However, access to therapy is often limited by financial barriers and stigma, particularly among younger individuals who may be reluctant to seek help [7,17,18]. As a result, digital mental health interventions, such as mHealth applications, have gained attention as accessible tools for fostering body acceptance and promoting self-compassion [7,19,20]. However, maintaining long-term engagement in these applications remains a challenge [8].

One emerging solution to this issue is gamification. Gamification is the application of game design principles in non-gaming contexts [21] and has been shown to increase motivation and adherence in mental health applications for general populations. Research indicates that elements such as progress tracking, points, badges, and leaderboards can increase motivation [9–11,20], while features that promote social interaction, such as community challenges and peer support, have been linked to sustained usage [22]. However, gamification must be carefully designed to align with mental health goals, as overuse of competitive elements may lead to negative self-comparisons and increased stress [9].

In this study, the use of gamification refers specifically to motivational elements such as point systems, progress tracking, and feedback mechanisms that are intended to encourage user engagement and reflection. This approach does not include peer competition, leaderboards, or social comparison due to the sensitive nature of mHealth applications. It aligns with definitions of gamification as behavioral reinforcement systems are commonly found in gamification and mHealth literature [9,10].

According to Self-Determination Theory (SDT), people are intrinsically motivated when interventions support autonomy, competence, and relatedness [23]. Research suggests that intrinsic motivation-based gamification (e.g., storytelling, personalization, and interactive challenges) plays a key role in fostering long-term engagement, complementing extrinsic rewards like points and leaderboards [9].

Behavioral nudging is another way to sustain usage and takes a different approach than gamification by subtly guiding users toward certain behaviors [11]. This approach includes various forms of encouragement, such as push notifications that provide motivational reminders, micro-goals that support incremental progress, and prompts that guide users toward healthier behaviors [24]. Unlike gamification, which often incorporates points and progress tracking, nudging relies on environmental or contextual cues to steer behavior in a direction that feels natural rather than externally motivated [11]. Research suggests that subtle, personalized nudges can promote long-term adherence while reducing the risk of app fatigue [11].

AI-driven chatbots have shown promise in mental health interventions [7,20], offering nonjudgmental interactions that may help reduce stigma while providing real-time support around the clock [20]. However, studies caution that over-reliance on AI-driven chatbots may discourage users from seeking human support when necessary [20]. This highlights the need to balance automated assistance with human-guided interventions.

One such AI-driven chatbot is TrueBalance [12], a large language model (LLM)–based conversational AI agent developed using OpenAI's pre-trained language model architecture. TrueBalance was initially deployed on the messaging platform Telegram allowing users to engage with the agent via a chat interface for real-time support. TrueBalance has been evaluated in prior research to assess its potential to promote positive body image in young adults by integrating biomedical determinants of mental health alongside principles from cognitive behavioral therapy (CBT). The agent supports open-domain conversations on topics such as body image and nutrition, offering reflective prompts, cognitive restructuring exercises, and self-guided psychological support [12].

While research on gamification, nudging, and AI-driven chatbots has advanced understanding of user engagement in mHealth applications, their application to body image interventions for athletes remains limited. Only a few studies have explored the use of chatbots for body image perceptions [7], but these did not employ open-domain conversational AI or tailor intervetions to athlete needs. Additionally, the specific contribution of individual features within these applications remain underexamined. Human-Centered Design (HCD) prioritizes user needs to enhance usability and engagement, often through iterative testing and refinement [25]. This study applies HCD principles, specifically co-design and usability testing, to shape the TrueBalanceApp in alignment with athlete preferences. By involving individuals with varying levels of training, the research explores how digital strategies can support body image among physically active populations. Ultimately, this study aims to offer insight into how a mobile application can be designed to promote positive body image in athletes.

Method

Building on the prior development of the TrueBalance conversational AI agent, this study employed a human-centered design (HCD)-inspired process to design a mobile application tailored to the needs of athletes. The goal was to embed the existing conversational agent within a broader app experience that promotes positive body image and fosters long-term user engagement. The AI agent was integrated into the mobile app using the OpenAI API, maintaining its core functionality while offering a more user-friendly and accessible format. The design process included a co-design workshop and usability testing to ensure alignment with athletes' needs, preferences, and psychological engagement factors.

The development process can be seen in Figure 4 and consisted of four phases: (1) identifying evidence-based engagement features from existing literature; (2) creating an interactive prototype; (3) conducting a co-design workshop with athletes to refine features and interface design; and (4) developing the full application followed by usability and functional testing.
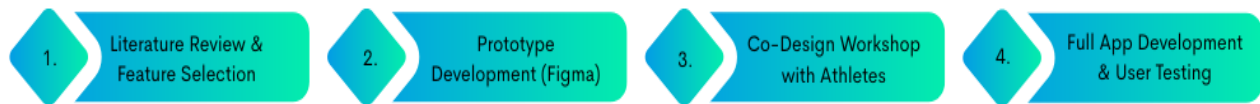

Figure 4. Flow chart of the development process of designing the TrueBalanceApp

TrueBalance conversational AI agent

In this research, the existing TrueBalance conversational AI agent was integrated without altering its core content. The AI agent was slightly personalized for the athletic population by adding a setup questionnaire about sport type and training volume, allowing for more contextually relevant interactions with users engaged in structured physical activity.

The conversational AI agent was re-built by the researcher using OpenAI's Assistants API. It was configured to exactly replicate the functionality and interaction style of the original TrueBalance assistant described in the previous research [12]. The assistant was created using the same instructions and curated content that shaped the behavior of the previously tested version, ensuring continuity in tone, purpose, and psychological framing. The assistant supports real-time, open-domain conversation using OpenAI's large language models.

Initial prototype development

An interactive prototype of the TrueBalanceApp was created using Figma, selected for its efficiency in designing and testing user interfaces. The prototype served as a visual and interactive representation of the app's intended functionality but did not yet include backend capabilities. Instead, it simulated key user interactions to allow for structured evaluation of user experience, interface design, and feature desirability prior to full development. This prototype was used to guide discussions and prioritization activities during the co-design workshop.

Co-design workshop

The co-design workshop aimed to refine the app's feature selection, engagement mechanisms, and visual design. Co-design was used to ensure the app reflected the real needs, preferences, and lived experiences of the target users, aligning with human-centered design principles. The session was conducted over Zoom to facilitate participation and recording and was organized through KTH Royal Institute of Technology. It lasted approximately two hours. A more detailed description of participant characteristics and workshop activities is provided below.

Participant recruitment and characteristics

The participants were selected through a convenience sampling strategy from the researcher's personal network, based on their interest in body image and mental health, professional background, and diversity in sporting experience. Eligible participants were adults actively engaged in sport at least at the recreational level, training a minimum of six hours per week. This threshold was selected to include individuals with a structured training routine, who are more likely to experience sport-related body image pressures. The participant characteristics can be seen in Table 1.

Table 1. Participants´ characteristics and training background

ID	Age	Gender	Main sport	h/w* training	Classification**	Profession
P1	30	Male	Triathlon	16	National level	Engineer
P2	28	Male	Running	8	Recreational	Engineering student
P3	30	Female	Cross-country skiing	6	Recreational	Dietitian
P4	31	Female	Running	10	Competitive	Psychologist

*h/w = hours per week
**The classification is based on a combination of training volume and competitive experience, with national-level athletes participating in elite competitions, competitive athletes training regularly and competing at club level [26], and recreational athletes engaged in non-competitive or self-directed fitness.

Workshop structure

The workshop was designed to explore participants' experiences with mental health apps, evaluate the usability of the prototype, and gather input for feature refinement. An agenda was developed in advance, including discussion prompts, interactive tasks, and group activities. To begin, participants were asked to reflect on and discuss their prior use of digital wellness or mental health tools, focusing on perceived strengths, limitations, and engagement patterns. This was followed by a guided prototype walkthrough, during which participants were instructed to perform specific tasks such as logging training sessions or make a journal entry. The researcher planned to observe and document usability issues, user expectations, and general ease of navigation.

A feature prioritization activity was prepared using a "feature auction" format, where participants received a virtual budget to allocate toward the features they considered most valuable. This activity, along with other interactive components, was conducted using a shared Miro board, allowing participants to collaborate in real time and enabling

structured capture of feedback. The session also included a facilitated discussion on gamification and behavioral nudging. Participants were prompted to consider how elements such as rewards, progress tracking, reminders, and positive reinforcement might enhance engagement. Finally, the workshop concluded with a visual design and customization segment, where participants were asked to give feedback on UI elements including color schemes, personalization options, avatar design, and interaction preferences (e.g., voice vs. text input, human likeness of avatars). The entire session was video- and audio-recorded for transcription and later analysis.

Data analysis

The workshop was video recorded, and the audio was transcribed verbatim. Deductive qualitative content analysis was chosen for its suitability in exploring user experiences and perceptions in a structured yet flexible way, following the principles outlined by Bengtsson [27]. A deductive coding strategy was used, based on five themes identified in advance from the study's aims: Preferred Features, Gamification Preferences, Nudging Strategies, UI/Design, and Avatar Preferences. Transcripts were reviewed multiple times to identify meaning units relevant to each theme. These were then coded and organized into sub-themes to capture participants' views in more detail. The analysis primarily focused on manifest content, what was clearly stated, while allowing for interpretation of latent meanings where appropriate. To improve trustworthiness, the coding process was discussed within the research team to ensure consistency and reduce bias. A full summary of the coding framework, including themes, sub-themes, and quotes, is available in Appendix B.

Full app development

Based on insights from the co-design workshop, the Figma prototype was revised and prepared for full development. The interface was converted into front-end code using LocofyAI, a plugin that translates user interface designs into React Native components. Locofy interprets layout structures from Figma and exports JSX code in TypeScript, which was then reviewed and manually refined to ensure consistent styling, accessibility, and responsiveness across screen sizes.

The mobile application was developed using React Native in combination with the Expo framework, which enabled cross-platform development. Expo was chosen due to its integrated testing environment, simplified build tools, and support for native device features. The app was distributed to participants as an Android APK generated via Expo's EAS Build service. While testing was limited to Android, due to participant devices, the codebase remains compatible with iOS.

Firebase Authentication and Firestore database were used for user management and real-time data storage. Each user's data were stored in secure, user-specific document collections. This allowed for persistent access to personal data across sessions and devices, while maintaining data separation between users.

The TrueBalance conversational AI agent was integrated using the OpenAI Assistants API, which allowed the app to interact with the TrueBalance AI assistant. When a user sent a message through the chat interface, the app transmitted it to the API. The assistant

responded with structured messages, which were rendered in real time in the app's chat interface.

To support voice-based interaction, the app implemented both input and output features. Voice input was transcribed using OpenAI's Whisper speech-to-text model and displayed in the chat as user textmessage. Voice output was generated from the assistant's messages using OpenAI's text-to-speech (TTS) model, and playback was handled via Expo's Audio API. These features were implemented to support multimodal interaction and accessibility. The app followed a modular architecture to enable future updates or additions without extensive restructuring.

Usability and functional testing

After the app was fully developed, user testing was conducted to evaluate interface usability, user engagement, and technical performance. The same group of athletes who participated in the initial co-design workshop was recruited for testing. Participants were instructed to use the app on their personal smartphones in real-life settings for a minimum of 24 hours, during which they were encouraged to explore all available functionalities and document any comments or issues.

After this period, participants completed the mHealth App Usability Questionnaire (MAUQ) [28]. Subsequently, each participant took part in a short, semi-structured interview that incorporated both structured questions informed by the MAUQ and open-ended prompts to elicit qualitative feedback. These interviews explored user experience (e.g., navigation clarity, ease of use), technical performance (e.g., responsiveness, functionality), and overall impressions, including perceived usability, satisfaction with app features, and likelihood of continued use. Error handling was also assessed by identifying bugs, crashes, or unexpected behaviors reported during use.

Ethical considerations

All participants provided informed consent prior to participation in this study. The study followed established ethical guidelines for human-centered research [29], including the principles outlined in the Declaration of Helsinki [30]. Participation was voluntary, and participants were informed of their right to withdraw at any time without consequence. All participant data was anonymized and stored securely using encrypted Firestore database collections accessible only to the researcher. Audio recordings and interview transcripts were saved locally on a password-protected device and deleted after transcription. No identifying information was included in the final analysis.

Given the sensitivity of topics related to mental health and body image, additional ethical considerations were addressed beyond standard informed consent. Participants were informed that the app is not intended to diagnose or treat clinical conditions, and while in the testing stage it was not suitable for individuals with active eating disorders or severe psychological distress. This was clearly communicated during recruitment and reinforced through disclaimers within the app interface.

The conversational AI agent was developed using a widely used large language model (LLM), which can support reflective dialogue and nonjudgmental interaction. However, the use of LLMs in mental health contexts involves certain risks, including

the potential for inaccurate or unexpected responses, accurate but inappropriate responses, and the possibility of users encountering emotionally triggering content. This study did not evaluate the chatbot's clinical accuracy, as the focus was on user experience and design considerations for integrating such technology into athlete-centered wellness tools.

Result

The results are presented in four main sections. (1) The first section outlines the initial prototype, including its design considerations and proposed features. (2) The second section summarizes findings from the co-design workshop, highlighting user input on feature preferences, emotional support, and interface expectations. (3) The third section describes the final version of the app developed for testing, reflecting adjustments based on workshop feedback. (4) The final section presents user testing results. Feedback addressed aspects such as feature utility, interface design, emotional tone, and perceived potential for continued use, which helped inform future development directions.

Prototype overview

Two design variations of the same prototype were developed, drawing on findings from literature on digital mental health interventions, body image support tools, gamification, and self-monitoring strategies. Its aim was to support athletes' body image perception and mental well-being by promoting intrinsic motivation, self-reflection, and long-term engagement. The prototype included several features intended to promote a holistic self-monitoring experience and reinforce positive body image, see an overview of the features in the homepage shown in Figure 5. The features are further described below.

Journaling, Training & Body Image Logging

The prototype included a journaling feature, see Figure 6, that was implemented to allow open-ended self-reflection, providing users with a space to express thoughts on body image, athletic performance, or overall well-being. Unlike structured assessments, this feature aimed to offer a flexible space for personal insights.

It also included a self-monitoring feature, see Figure 7, to track changes in body appreciation over time, integrating the validated Body Appreciation Scale-2 (BAS-2), a widely used measure assessing positive body image [31], for repeated assessments. This approach was intended to help users observe fluctuations in body image and potentially encourage long-term awareness.

Additionally, a training log was included to document physiological responses, session intensity, injury status, illness, and fatigue levels, potentially offering a broader view of physical and psychological well-being, this feature can be seen in Figure 8. By combining quantitative tracking (BAS-2) with qualitative self-expression (journaling), the prototype sought to create a holistic self-monitoring experience.
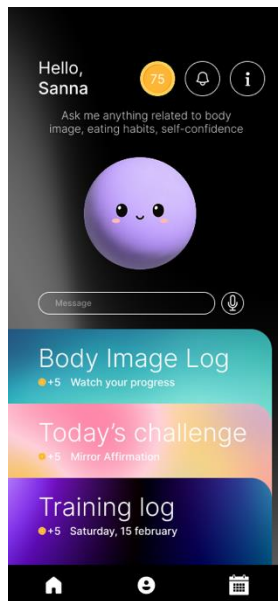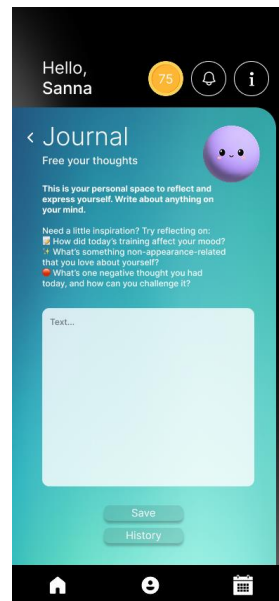
Figure 5. Home page of the initial prototype

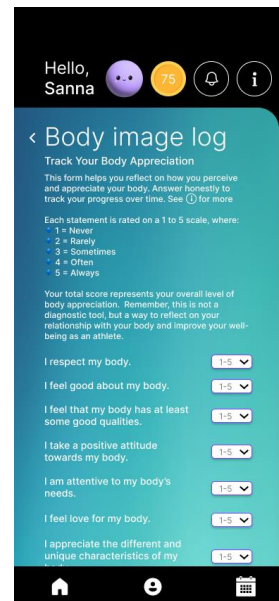*Figure 6. Journaling feature of the initial prototype*

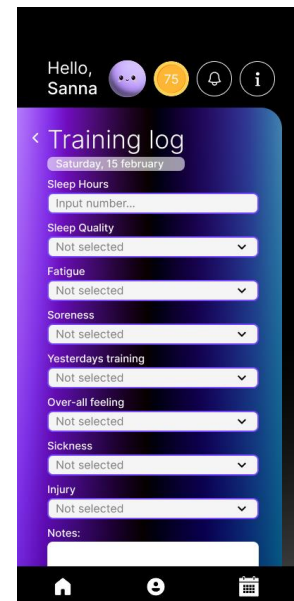*Figure 7. Body image tracking feature with BAS-2 of the initial prototype*

Figure 8. Training log feature of the initial prototype

Daily challenges for positive body image

The prototype also included a daily challenges feature aimed at reinforcing positive body image through small, proactive, structured tasks, see Figure 9. These challenges were intended to be automatically generated by the TrueBalance AI agent through a prompt that asked the assistant to provide a challenge designed to promote positive body image, while also ensuring variety and minimizing repetition. Since the assistant is grounded in cognitive-behavioral theory and self-compassion principles, the challenges could involve activities such as identifying three things one appreciates about their body, practicing mindful movement, or reflecting on body functionality. While the previous features focuses on reflection the challenges were intended to promote a positive body image through meaningful hands-on activities.

Gamification element

A points-based system was added to encourage user engagement by reinforcing consistent interaction with the app. Users earned points for completing activities such as logging body image reflections, journaling, completing daily challenges, and tracking their training data. The system was designed to emphasize individual progress rather than competition, aligning with the app's non-comparative, self-focused approach to well-being.
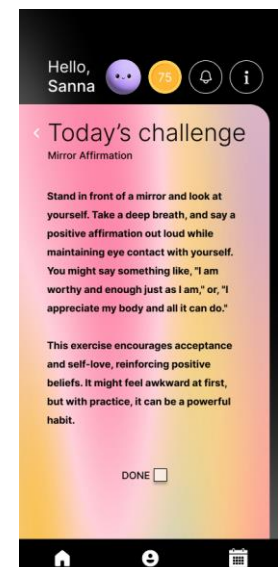

Figure 9. Daily challenge feature in the initial prototype

Co-Design workshop

The co-design workshop allowed participating athletes to provide feedback on the initial prototype of the TrueBalanceApp. Their responses offered insights into which features were considered most relevant, which design elements influenced engagement, and how the app might better support body image and emotional well-being. The following section summarizes the main themes that emerged from the feedback, including preferred features, design-related considerations, and user expectations. The predefined key themes and the sub-themes identified during analysis are presented in Table 2.

Table 2. Key themes and sub-themes identified from the Co-design workshop.

Theme	Sub-theme
Preferred features	Conversational AI agent for emotional support
	Body Image Tracking
	Adaptive Daily Challenges
	Journaling and Training Log for Holistic Awareness
	Goal Setting
	Daily Affirmations
	Integrated Features
Gamification preferences	Preference for Intrinsic over Extrinsic Motivation
	Progress Tracking
	Celebratory and Visual Feedback as Motivation
Nudging Strategies	Gentle, Personalized Nudging
	Exploratory Notifications
	Adaptive Notifications
UI / Design	Calm Aesthetics
	Customizable Themes
Avatar Preferences	Human vs. Non-Human Avatars

Preferred features

The co-design workshop revealed that participants valued several features of the prototype, while also identifying areas for improvement to enhance engagement and usability. The result from the feature prioritization activity or "feature auction" can be seen in Table 3.

Table 3. Feature prioritization activity* result from the Co-design workshop

Feature	Coins awarded					Reasoning behind choices
	P1	P2	P3	P4	Total	
Chatbot	1	4	3	2	10	"Can help you to get out of negative cycle". "Saying it out loud/writing it can help you reflect/identify thoughts".
Body Image Tracking	2	1	1	2	6	"Mixed feelings. Good to track body image compared to training etc". "Risky to put so much focus on good or bad feelings about your body". "Could lead to negative reinforcement".
Journal	0	2	2	3	7	"Nice that you can track your mood". "Writing helps you put words to your feelings". "Free journaling is better than tracking".
Daily Challenge	0	1	1	3	5	"Should not be random, should be personalized challenges". "Prompts you to actively seek positive body image". "Focused task to promote positivity". "Important feature".
Training Log	5	2	1	2	10	"Can show a trend between positive image and performance, which is interesting".
Point system (gamification)	2	0	0	0	2	"Coins respresent a risk of performance anxiety". "Do not think they are necessary".

*Participants were given 10 coins each and asked to "bid" on the features they considered most valuable. They were then asked to explain the reasoning behind why they allocated or chose not to allocate coins to each feature.

Conversational AI agent for emotional support

Participants valued the AI agent for emotional support but emphasized that its responses should be adaptive and responsive to user input.

> *I think it [the AI agent] is something that can help you to get out of your negative circle. (P2, male, 28)*

During discussions on voice integration, opinions varied. Some participants felt expressing thoughts aloud could be therapeutic and inclusive, while others argued it was impractical, as users might not always have privacy to speak openly.

> *… maybe it's like an intervention itself to say stuff out loud, just not keeping it inside. Allow yourself to think what you think about yourself and say what you say about yourself and hear it. (P4, female, 31)*

> *… So my parents are downstairs, I want to talk to the assistant, but like I'm really nervous, I'm really self conscious, and I don't want to risk my parents to hear uh this kind of stuff. So, and that's why am I saying I think it [voice intergration] is a bad idea. (P1, male, 30)*

Ultimately, participants agreed that having the option for the AI agent to audibly read its messages to the user while still seeing them in text was beneficial, providing a therapeutic experience similar to interacting with a therapist. Integration with journaling and daily challenges was also suggested to enhance engagement and interactivity.

> *It would be nice if you can implement that you have still the avatar talking to you while you text back. (P1, male, 30)*

Body image tracking

Participants had mixed opinions on the body image log feature. While some found it useful for self-reflection and monitoring progress, several others expressed concerns that repeatedly logging negative feelings about their body could reinforce those thoughts over time. There was also apprehension about engaging with the feature on days when body image perception was particularly low, as it could feel overwhelming or discouraging.

> *If someone tracks their body image and keeps seeing negativity, it might reinforce the negative thoughts. (P4, female, 31)*

To address these concerns, participants suggested making the feature optional and simplifying the assessment process by reducing the number of questions per session, making it more accessible without adding unnecessary pressure.

> *When I see this screen, I feel like, 'Oh no, it's so much. I don't have time or I don't have the energy to do it.' But if it was fewer questions at a time, it wouldn't feel so overwhelming. (P3, female, 30)*

Adaptive daily challenges

Daily challenges were seen as valuable in promoting positive body image, but participants stressed the need for adaptability. Rather than being randomly assigned, challenges should align with the user's progress and comfort level, ensuring they foster confidence while also gently pushing users out of their comfort zones.

> *...it has to be adaptive. It can't be like you have 100 challenges and then you randomly choose one. Like you both want to build confidence and put the person out of their comfort zone at the same time but gradually. (P1, male, 30)*

> *I think its good ... it's like the only feature in the app that prompts you to do things that could have positive effects on your feelings about yourself. The other ones are only like you are talking out about it. (P4, female, 31)*

Journaling and training log for holistic awareness

Journaling was unanimously seen as a valuable feature, particularly for tracking emotional and psychological patterns over time. Overall, open journaling was more appreciated than body image logging as a method of tracking progress. Participants noted that reviewing past journal entries would help them recognize personal growth and improvements in their body image perceptions.

I'm journaling quite a lot recently and I think that sometimes, even if you have bad thoughts, but then you start to feel a bit better and then you see that the week before you were very low, but now it's going better. Maybe it helps you to say, okay, but last week I was very shitty. Now I'm feel a bit better, so it can help you a bit to see the improvements because sometimes you cannot see the improvements day by day, but just if you look from a more distant perspective. (P2, male, 28)

Similarly, The training log was praised for revealing insights linking body image with athletic performance, supporting holistic self-awareness.

"I like the idea that this can show a trend between positive body image and performance. Or a trend with a lack of injury in my sport, haha. (P1, male, 30)

Goal-setting

While goal setting was proposed as an additional feature, participants cautioned that it should be accompanied by structured guidance. Without support, setting goals could become discouraging if progress was unclear. There was a concern that goal setting might lead to frustration if users did not see immediate improvements, emphasizing the need for adaptive goal-tracking mechanisms that provide encouragement rather than pressure.

I totally agree that it's good to have a goal, but since the app is not an app that guides you towards the goal, it's also a risk to have a goal if someone doesn't, like, it can also be a goal that you see and, okay, I'm not moving forward. (P4, female, 31)

Daily affirmations

Participants suggested incorporating affirmations that adaptively align with users' emotional states and recent experiences, providing positive reinforcement.

What are three good things with yourself today? Or one good thing. That you have to focus on something good today.... And it can be anything. It could be that I slept well or that I did work well or that I did a good workout. (P4, female, 31)

Integrated features

Participants favored seamless interaction between features (e.g., AI agent prompting journaling or challenges).

If the chatbot is more interactive with the rest of the app maybe, so as we said previously, you write something and then it says, oh, do you want to write this in your journal? Or maybe it prompts you directly to do the daily challenge. (P2, male, 28)

They cautioned against feature overload, recommending a clearer focus on core functionalities. They suggested refining the app's core functions to ensure that it

remains specialized and easy to navigate. While they acknowledged the importance of balancing features that actively promote positive body image with those that provide passive self-reflection, they recommended narrowing the focus to maintain clarity and effectiveness.

> *I'm really curious about the thought behind the app, because as I understand it, the goal is to help you build a positive self-image. But how is it actually supposed to do that? Because it seems really different depending on who uses it. If someone mostly has positive thoughts, that's what they'll see when they open the app. But if someone is struggling and having lots of negative thoughts, the app might just fill up with negativity, which could make things worse. As the psychologist that I am, that worries me a bit, because even though you might feel validated by the chatbot, you're not really getting the tools or strategies you need to actually deal with those negative thoughts. (P4, female, 31)*

Gamification preferences

In this study, as stated in the background, gamification refers to non-competitive design elements, such as earning points or receiving rewards for completing tasks, intended to promote motivation. The prototype did not include competition, leaderboards, or direct comparison with others, however competitive gamification was also discussed during the workshop.

Preference for intrinsic over extrinsic motivation

Participants were generally skeptical about the point-based gamification feature, fearing that it could create unnecessary pressure rather than motivation. Many were hesitant about gamification overall, especially competitive forms that involve comparison with others. Instead, they preferred elements that support self-reflection and personal growth over features that foster competition or chasing points. Some participants saw value in tracking progress but emphasized that earned points should not be lost due to inactivity, as this could lead to feelings of failure rather than encouragement.

> *I think the coin thing will add a on a 'prestationskrav' [performance requirement]. (P3, female, 30)*

> *Duolingo is quite toxic if you lose the streak, for example, but like here you don't. I mean, you can log it, you can skip a day. If you feel like crap and you still have like those points that you've earned, so that's nice, I think, as a positive feedback. (P1, male, 30)*

Progress tracking

Progress tracking, in general, was seen as useful, but only if it remained optional. Participants were wary that seeing negative trends in their body image scores could reinforce self-doubt rather than encourage improvement. Instead, they suggested that users should have the choice to engage with progress tracking at their discretion.

> *I think it can be good for people who want to track their progress, but only if it's not something that the app pushes on you. (P4, female, 31)*

Celebratory and visual feedback as motivation

They suggested incorporating encouraging messages or visual reinforcements, such as animations or positive feedback, to acknowledge progress without using points or rewards.

> *Instead of earning points, what if the app just gave encouraging messages when you complete something? (P1, male, 30)*

Nudging strategies

Participants favored nudging strategies that were gentle, emotionally aware, and responsive to their individual usage patterns and moods. Rather than relying on rigid or performance-based elements, they preferred subtle prompts that acknowledged achievements, such as encouraging messages or simple visual celebrations, supporting motivation without pressure.

Gentle, exploratory, personalized nudging

Participants expressed a clear preference for gentle, personalized nudges rather than directive reminders, highlighting the importance of notifications feeling supportive, not demanding.

> *If I get a notification every day based on what I filled in my journal or what I did or didn't do the day before, I think that would be positive. Just reminding me gently that the app is there, nothing too demanding. (P4, female, 31)*

Exploratory notifications

Notifications should invite reflection rather than explicitly instruct users to complete tasks. Participants preferred open-ended, reflective questions that felt like genuine check-ins.

> *It's better if a notification just asks how I'm doing instead of telling me to do something. (P3, female, 30)*

Adaptive notification frequency and content

Participants suggested that notifications should adapt dynamically based on prior app usage, mood entries, and frequency preferences. For instance, increased check-ins following negative journal entries, or encouraging re-engagement after prolonged app inactivity, were seen positively.

> *It can connect to something you wrote previously, if two days ago I wrote 'I feel very bad,' it can later ask, 'How do you feel today? Better?' (P2, male, 28)*

> *The more you've been away, the more the app could encourage you back. If you haven't journaled in a month, maybe the next week it prompts journaling more often, instead of random notifications. (P1, male, 30)*

Participants also emphasized the importance of customizable notification frequencies.

> *Can notifications be optional? Maybe let me choose if I want them daily, weekly, or only when I'm inactive, so I can choose myself. (P3, female, 30)*

UI and design

Calm asthetics

Participants preferred lighter, "calming" colors, associating them with journaling and well-being. They found the colors in the second design of the initial prototype more appealing and in line with the reflective nature of the app, see Figure 10.

> *I agree with the color, um, the color in light version are more like paper, more calming. Which I associate with journaling. (P1, male, 30)*

Customizable themes

Participants desired theme customization to enhance personal comfort and engagement, allowing users to select themes that resonate personally.

> *Maybe, like in other apps, you could switch between light and dark mode, or even between different avatar types. (P2, male, 28)*
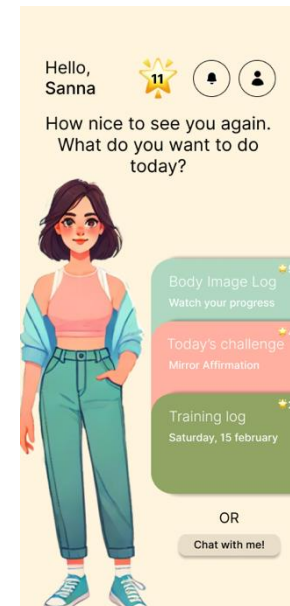

Figure 10. The second design of the initial prototype with a lighter color-scheme but the same functionality.

Avatar preferences

Human vs non-human avatars

Opinions regarding the avatar varied. Some participants preferred the idea of human-like avatars, believing they added a more personal touch, while others felt that a less humanized avatar would help set more realistic expectations for the AI agent's capabilities. Given the diverse perspectives, participants discussed whether the actual appearance of the avatar was truly important. They suggested that if a human-like avatar were included, users should have the ability to customize its appearance to better reflect their personal identity, and to be able to identify with the avatar.

> *When you talk with ChatGPT, you don't see something that reminds you of a person or an animal, but you use it anyway. So maybe it's not so important what face you have there. (P2, male, 28)*

Summary of key findings

The co-design workshop provided critical insights that informed the refinement of the TrueBalanceApp prototype. Participants emphasized the importance of adaptable AI agent responses, optional body image tracking, seamless feature integration, and a preference for nudging over gamification. They also highlighted the value of

personalized, non-directive nudging and a streamlined interface. These findings guided subsequent stages of app development, helping to ensure a user-centered, engaging, and supportive experience for athletes navigating body image challenges.

Final application design

The final version of the *TrueBalanceApp*, developed for usability testing, featured a simplified and calming user interface designed with a light color palette to support reflection and emotional ease. The app included four core features: the TrueBalance conversational AI agent though a chat interface, journaling, training log, and daily challenges. Additional secondary functions included a profile page, a information screen, a simplified tracking system through mood tagging and a calendar view, see Figure 11.

Core features

TrueBalance conversational AI agent

The *TrueBalance* conversational AI agent is accessible via the home page and at all times in the bottom navigation bar, serving as a key feature, supporting open-domain conversations about body image, training, and emotional well-being. Optional two-way voice interaction is available, allowing users to speak and receive spoken responses. The AI agent visualized through a non-human avatar, a round ball with a friendly animated face, which provide supportive feedback based on mood and interaction context. The final design of the chat interface can be seen in Figure 12.

Journaling

The journaling feature was kept, where users can write open-ended journal entries about their emotional and mental state, see Figure 13. Each entry is analyzed with a back-end sentiment analysis which gives information to generate a small feedback from the avatar, reflecting the tone of the entry. Users were also asked to rate their mental state/body image using a custom emoji scale ("Confident", "Calm", "Insecure", and "Unhappy"), see Figure 14.

Training log

The training log enable users to document physical activity, including session intensity, fatigue, or recovery notes. Similar to the journal, users can select an emoji reflecting their physical state ("Strong & Energized", "Balanced & Neutral", "Fatigued or Sore" and "In Pain or Unwell"), the design for these screens can be seen in figure Figure 15-16.
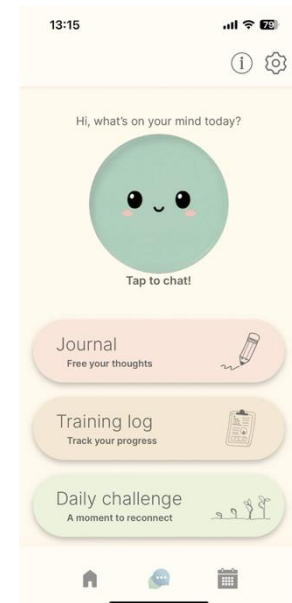

Figure 11. Final design of the homepage of the TrueBalanceApp
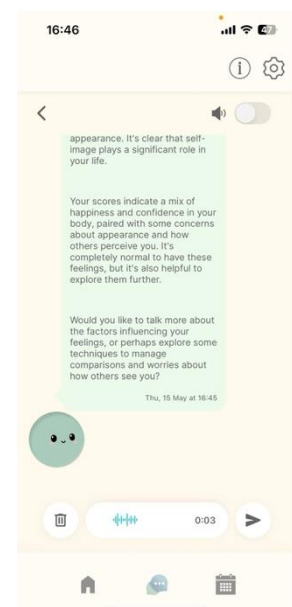

Figure 12. Chat interface of the final design of the TrueBalanceApp – voice recording activated.

Daily challenges

In contrast to the earlier concept where daily challenges were intended to be generated by the AI assistant, the final app used a predefined set of 50 body-positivity exercises curated from reputable mental health and body image literature and online forums. Each day, users were presented with one randomly selected challenge from this database. To support autonomy and comfort, users could dismiss any challenge they felt was not relevant or that they were uncomfortable completing, see design in Figure 17. Completing a challenge triggered a small celebratory animation from the avatar and screen, reinforcing positive engagement, see Figure 18.

Tracking and review

There were no gamified features (e.g., points) in the final version. A calendar view allowed users to revisit past entries, training logs, and challenges. Emoji-based mood ratings were displayed within the calendar, allowing users to identify trends over time, for example, noticing days marked as "strong" physically or "confident" mentally, hence providing an optional tracking functionality, see Figure 19-20.
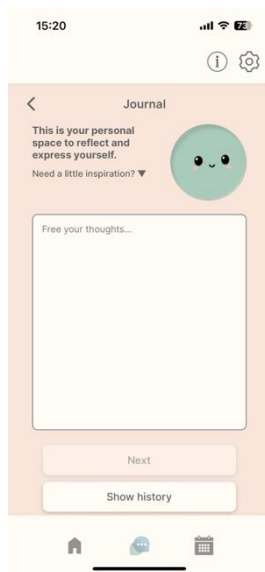
Figure 13. First screen of the journal feature
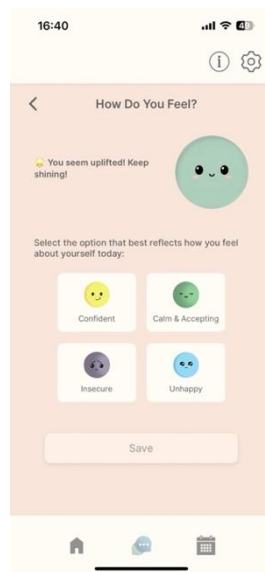
Figure 14. Second screen of the journal feature

Figure 15. First screen of the training log feature

Figure 16. Second screen of the training log feature

Figure 17. Daily challenge feature interface

Figure 18. Figure showing the celebratory interface when the user completes a daily challenge

Figure 19. The calendar function in the final design

Figure 20. Expanded calendar view for a date with a completed challenge, journal entry, and training log.

Other features

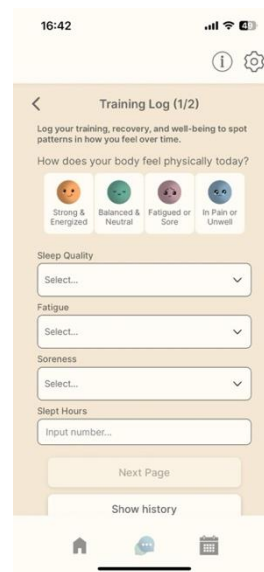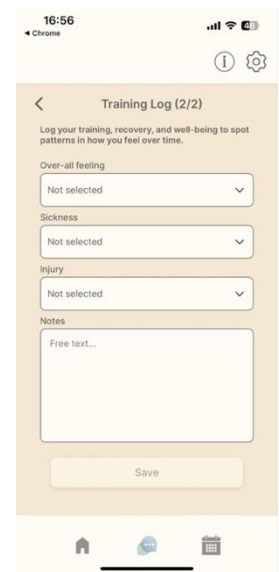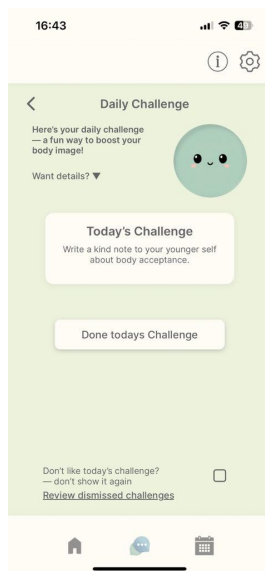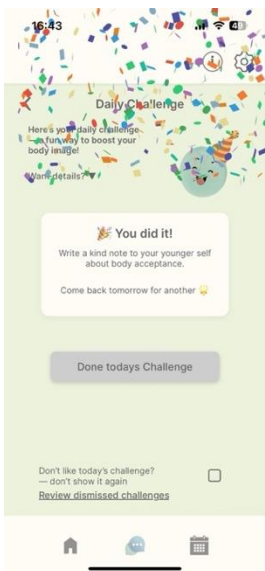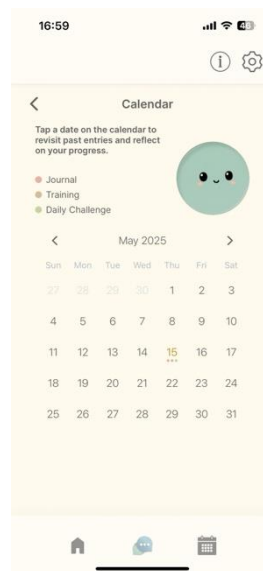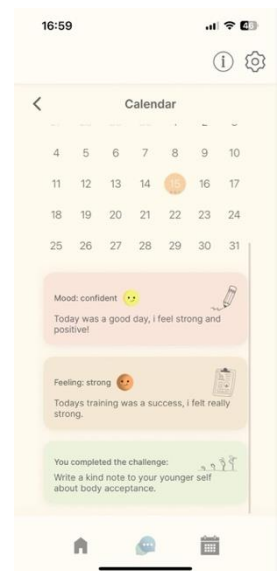The app also includes a profile page where users can change their nickname and other personal information, change the AI agents' voice, and manage notification preferences, see Figure 23. An information screen is provided to explain the purpose of each feature and prominently displays the following disclaimer: *"TrueBalance is not a medical or therapeutic tool. It is for self-guided reflection and does not diagnose or treat conditions. If you need professional support, please seek expert guidance"* (Figure 22).

Both the profile and information screens are accessible from any page via a top bar. A bottom navigation bar provided quick access to the conversational AI agent, the calendar view, and the app's home page, supporting smooth and intuitive navigation throughout the app. The log in page with the app logo can be seen in Figure 21.
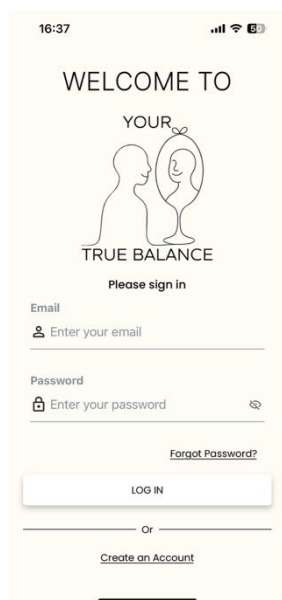
Figure 21. Log in page of the final design
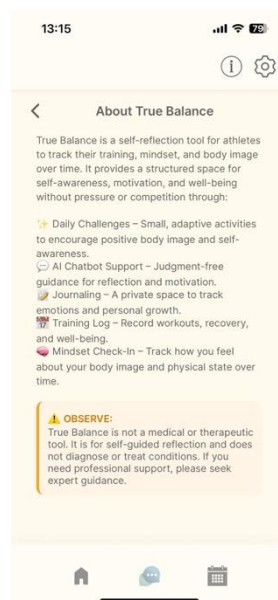
Figure 22. Information screen design of the TrueBalanceApp
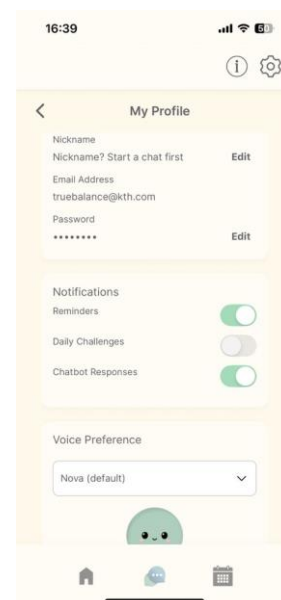
Figure 23. Profile screen of the final design of the TrueBalanceApp

Notifications

Based on participant feedback from the co-design workshop, the final version of the app included a gentle notification system to support ongoing engagement without creating pressure. Users receive a daily reminder about the new challenge at 9:00 AM and a reflective "check-in" prompt at 7:00 PM. These evening notifications are drawn from a curated list of encouraging, open-ended messages designed to invite self-reflection and emotional awareness without being directive. Examples include: *"A little check-in can go a long way. Wanna reflect?"*, *"Even small steps count. How are you feeling today?"*, and *"Body and mind check — want to log today's thoughts or training?"*. In addition, if a user initiates a message in the chat and leaves the app before receiving a response from the AI agent, the reply is delivered via a push notification. All notification types are optional and can be enabled or disabled individually through

the profile settings, ensuring users maintain full control over how and when they engage with the app.

User testing

User testing of the final applicaion gave insight into user experience and perceptions of the app´s functionality, design, and tone. Feedback was gathered through the MAUQ, see Table 4, and through individual interviews with guided questions related to feature usefulness, design elements and engagement potential.

Feature usefulness and engagement

Participants responded positively to the final app's limited number of features, noting that its simplicity contributed to a clearer and more focused user experience. The available features were generally described as helpful for self-reflection and mood tracking. Journaling was appreciated for offering a space to process thoughts, while some participants valued the training log for promoting awareness of the connection between mood and physical activity. One participant stated:

> Its the type of app that can be there for you when you need it. (P3, female, 30)

Several participants also noted that the app's automatic display of recent entries after logging new input prompted unintentional reflection. This design element encouraged users to pause and notice patterns, such as changes in mood or body image over time, without feeling pressured. The most appreciated features varied between participants, with the daily challenge, journal, training log, and chat interface all mentioned as useful depending on individual preferences.

Reactions to human-like design elements

Participants generally responded positively to the human-like design elements of the app, including the movement of the avatar, voice integration, and the tone of messages. However, several emphasized that the non-human appearance of the avatar enhanced their comfort with the app. As one participant put it,

> I loved being able to avoid that [a human figure]… I don't think I would have opened up as much if it had looked like a real person. (P4, female, 31)

The avatar was described as soft, friendly, and calming. Participants appreciated that there was no sense of being judged or observed by another person, which helped them open up more freely. Subtle human-like movements and voice features were seen as pleasant. This balance contributed to a sense of psychological distance that felt comforting and appropriate for self-reflection. One participant noted that the app was *"very gentle… and soft. It doesn't feel like someone is watching you"*. One participant, who initially expressed concern that the app might discourage professional help-seeking, later reflected that its validating tone and clear messaging made it more likely to encourage help-seeking by raising awareness that support is available and by normalizing emotional reflection.

Sustained use potential

Although participants did not imagine using the app daily, most said they could see themselves using it during periods of need. The app was described as something they would return to when experiencing stress or wanting to reflect on their emotional state or routines. The daily challenge feature and journaling were considered useful for this kind of occasional engagement, and also the chat function for emotional support. One participant shared,

> *I wouldn't use it every day, but I would use it when I needed to reflect or felt overwhelmed. (P1, male, 30)*

Others expressed interest in using the app over a longer period to track emotional patterns or body image perceptions across weeks or months. The option to use features as needed, without pressure, was considered a key factor supporting potential long-term use.

Technical feedback

Participants noted a few minor issues affecting responsiveness and clarity, including occasional delays in chatbot replies and inconsistent notification delivery. However, the style and tone of both the AI-agents responses and the notifications were appreciated by all. Some users suggested improving navigation within the calendar view to make it possible to add new journal or training log entries from this screen. Features such as confetti animations and chat pop-ups were described as motivating. Additional suggestions included enabling the option to save chatbot messages to the journal and renaming the training log to better reflect holistic well-being.

Table 4. Participant responses to the mHealth App Usability Questionnaire (MAUQ), with mean scores across four participants. Rated on a 7-point Likert scale (1 = Strongly Agree, 7 = Strongly Disagree).

mHealth App Usability Questionnaire (MAUQ)	R 1	R 2	R 3	R 4	Mean
1. The app was easy to use.	1	3	2	1	1.75
2. It was easy for me to learn to use the app.	1	3	1	1	1.5
3. The navigation was consistent when moving between screens.	1	4	2	1	2
4. The interface of the app allowed me to use all the functions (e.g., entering information, responding to reminders, viewing information) offered by the app.	1	2	1	1	1.25
5. Whenever I made a mistake using the app, I could recover easily and quickly.	1	1	1	1	1
6. I like the interface of the app.	1	1	1	1	1
7. The information in the app was well organized, so I could easily find the information I needed.	1	1	1	1	1
8. The app adequately acknowledged and provided information to let me know the progress of my action.	1	2	1	1	1.25
9. I feel comfortable using this app in social settings.	1	3	5	1	2.5
10. The amount of time involved in using this app has been fitting for me.	1	1	2	1	1.25
11. I would use this app again.	1	3	1	1	1.25
12. Overall, I am satisfied with this app.	1	2	2	1	1.5
13. The app would be useful for my health and well-being.	1	2	2	1	1.5
14. The app improved my access to health care services.	n/a*	n/a*	n/a*	n/a*	n/a*
15. The app helped me manage my health effectively.	2	5	4	1	3
16. This app has all the functions and capabilities I expected it to have.	1	3	2	1	1.75
17. This mHealth app provided an acceptable way to receive health care services (e.g., accessing educational materials, tracking my activities, performing self-assessment).	1	2	3	1	1.75

*n/a = Not applicable, as the app does not provide or connect to health care services.
R1–R4 = Individual anonymous participant responses, they do not correspond to P1-P4.

Discussion

This study aimed to explore how a mobile application integrating a conversational AI agent could be designed, together with athletes, to support positive body image within structured sport and fitness contexts. This section discusses the principal results in relation to the study's research questions, followed by a comparison with prior work, reflections on limitations, and suggestions for future research.

Principal results

Three principal findings emerged, each offering insight into how digital interventions might more effectively meet the needs of this population.

1. Athlete responses to gamification and engagement features

While gamification is widely used to boost engagement in mental health apps, the athletes in this study expressed skepticism toward elements such as point systems and streaks. These features were often linked to increased pressure, performance anxiety, and a reliance on external validation. In contrast, participants favored strategies that fostered intrinsic motivation, emotional self-reflection, and gentle, non-directive feedback. Features like free-form journaling, small adaptive challenges, and personalized nudges were perceived as more meaningful and emotionally supportive. This finding directly addresses RQ2, suggesting that non-competitive engagement strategies are better suited to promoting sustained use among athletes for promoting a positive body image.

2. Simple and non-demanding design preferences

Participants emphasized that mental health apps should be easy to use, calming, and low in pressure. Interfaces that felt complex or overloaded with features were often described as overwhelming, particularly during times of stress or periods of emotional strain. Clear onboarding, intuitive navigation, and a gentle, supportive tone were seen as essential to creating a positive user experience. While much of the existing research focuses on adding features to enhance engagement, these findings underscore the importance of simplicity and focus. This directly informs RQ1, indicating that perceived usefulness and emotional safety are central to how athletes interact with body image interventions. It also contributes to RQ2, suggesting that simplicity itself can serve as an effective engagement strategy.

3. Design evolution through Co-design and user feedback

The initial prototype of the TrueBalanceApp incorporated a darker, more performance-oriented visual design and a structured interface similar to traditional fitness and training tools. However, the co-design and testing phases revealed a clear preference for a lighter and more minimal interface. This feedback led to a significant visual redesign, moving away from competitive or performance-driven aesthetics toward a more emotionally safe and reflective environment. Participants described the final design as "calming", "non-demanding" and "inviting", particularly appreciating the limited number of features and the app's non-judgmental tone. Similarly, early assumptions that users might prefer a relatable, human-like avatar shifted as the final

testing revealed the opposite: participants found comfort in the avatar's non-human nature. They described it as "soft", "pleasant", and "easier to open up to", specifically because it avoided the pressure of human interaction. Rather than reducing connection, the lack of a human presence allowed users to express themselves more freely without fear of being judged. These findings underscore the value of co-design in uncovering user needs that challenge initial design assumptions.

Together, these findings suggest that mHealth interventions for athletes benefit from a design approach that prioritizes user comfort, simplicity, and flexibility. Actively involving users in the development process through co-design proved essential for uncovering meaningful preferences, challenging assumptions, and ensuring that the app aligns with the real-world needs of its intended audience.

Usability assessment

The MAUQ scores, seen in Table 4, supported the qualitative findings, with consistently low (positive) ratings across ease of use, interface design, and satisfaction. All participants rated the app as easy to navigate and visually clear, reinforcing the value of the simplified, low-pressure interface developed through co-design. One item related to access to health care services was marked as not applicable, as the app is intended for self-reflection rather than clinical care.

Comparison with prior work

Previous research has highlighted the potential of gamification and nudging to improve engagement in digital mental health tools [9,11,20]. However, much of this work focuses on general populations, with limited attention to athletes and body image–specific interventions. This study adds nuance to that literature by showing that athletes may not respond positively to traditional gamified elements. A reason for this might be the competitive and performance oriented nature of athletes and the feeling of needing to preform or score the highest, which might impact negatively or add stress in a mental health focused app. Instead, participants preferred engagement strategies aligned with intrinsic motivation, echoing principles from SDT [23], but placing greater emphasis on comfort and autonomy over performance metrics.

Similarly, while existing mHealth design research sometimes focuses on feature expansion to boost engagement [8], the current findings suggest that minimalism and emotional accessibility may be more effective for this population. The preference for simple, non-demanding interfaces aligns with studies showing that feature overload can reduce engagement, especially during periods of psychological strain [8,11]. However, this study expands on that by linking simplicity directly to perceived usefulness and emotional comfort in the context of body image support.

The role of conversational AI in mental health apps has received growing attention [7,20], particularly for its potential to reduce stigma and offer real-time support. This study builds on prior work by showing that athletes valued the AI agent not for its human-likeness in appearance, but precisely because it was not human. These findings are consistent with previous research suggesting that non-human avatars can reduce the pressure to self-present [20]. In this study, athletes reported feeling more at ease and more likely to open up because the avatar was clearly not human.

While co-design has been broadly recommended in mHealth app development [25], its specific application in athlete-focused mental health tools has been less researched. This study shows how involving athletes directly in the design process helped surface user needs that might not have emerged otherwise. For example, early thoughts about the benefits of performance-oriented visuals, competitive features, or human-like avatars were challenged during testing. Participant input guided fundamental shifts in visual design, tone, and feature integration, contributing to a more soft and non-demanding user experience. These findings reinforce the value of co-design in developing contextually relevant

While this study focused on athletes, the identified preferences, such as a non-human avatar, non-competitive features, and simple interfaces, may also be relevant for other populations experiencing body image pressures including, as an example, adolescents. Future research could explore how these design choices resonate across broader user groups.digital interventions.

Limitations

This study involved a small, convenience-based sample of athletes (n = 4), which limits generalizability. While participants provided rich insights, a larger and more diverse sample would strengthen future conclusions. The 24-hour final testing period also limits understanding of long-term engagement or behavioral outcomes.

As the focus was on user experience and design, the app's clinical effectiveness and the psychological accuracy of the AI agent's responses were not evaluated. Although the agent avoids diagnostic content, its outputs were not systematically assessed.

Testing was limited to Android devices, potentially excluding iOS-specific usability feedback. Some technical issues were present during the final testing but did not prevent design-focused feedback. Lastly, the app is intended for self-guided reflection and might not be suitable for individuals with active eating disorders or severe distress until further extensive testing is done.

Future work

Future research could benefit from a larger and more diverse testing group, along with extended testing periods, to evaluate long-term engagement, emotional impact, and potential feature fatigue. The role of conversational AI also warrants further investigation, particularly regarding how users interpret responses and perceive the emotional tone of interactions over time.

It may be valuable, as brefly mentioned earlier, to test the application with adolescent athletes or a broader physically active population to explore whether its benefits extend to groups that may also experience body image concerns. Involving clinicians in co-design processes could support the development of adaptive features, such as personalized guidance, journaling prompts, or goal-setting tools.

Integrating the app with existing training platforms to automatically import physical activity data could further enhance usability. This would allow the app to focus more directly on the mental aspects of athlete well-being and reduce redundant data entry.

Finally, future iterations should consider testing the iOS compatibility and expanding personalization options, which were strongly desired by participants, while preserving simplicity and accessibility in the design.

Conclusions

This study examined how a mobile application incorporating a conversational AI agent could be co-designed with athletes to support positive body image in sport and fitness contexts. Using a human-centered design approach, the study identified preferences for non-competitive engagement strategies, supportive features, and a minimal, intuitive interface. The findings underscore the importance of simplicity, personalization, and user autonomy in digital interventions for athletes. Additionally, the use of non-human design in conversational agents was found to potentially promote openness and reduce perceived social pressure. These insights may guide the development of future mental health–oriented mHealth tools tailored to the specific needs of athletic populations, with an ongoing emphasis on user-led design.

Conflicts of interest
None declared.

Abbreviations
AI: artificial intelligence
API: application programming interface
BAS-2: body appreciation scale-2
CBT: cognitive behavioral therapy
HCD: human-centered design
LLM: large language model
MAUQ: mHealth App Usability Questionnaire
mHealth: mobile health
SDT: self-determination theory
TTS: text-to-speech
UI: user interface

References

1. Borowiec J, Banio-Krajnik A, Malchrowicz-Mośko E, Kantanista A. Eating disorder risk in adolescent and adult female athletes: the role of body satisfaction, sport type, BMI, level of competition, and training background. BMC Sports Sci Med Rehabil 2023 Jul 25;15(1):91. doi: 10.1186/s13102-023-00683-7

2. Jiménez-Morcillo J, Ramos-Campo DJ, Rodríguez-Besteiro S, Clemente-Suárez VJ. The Association of Body Image Perceptions with Behavioral and Health Outcomes among Young Adults. Nutrients 2024 Apr 25;16(9):1281. doi: 10.3390/nu16091281

3. Li Q, Li H, Zhang G, Cao Y, Li Y. Athlete Body Image and Eating Disorders: A Systematic Review of Their Association and Influencing Factors. Nutrients 2024 Aug 13;16(16):2686. doi: 10.3390/nu16162686

4. Satterfield NA, Stutts LA. Pinning down the problems and influences: Disordered eating and body satisfaction in male wrestlers. Psychol Sport Exerc 2021 May;54:101884. doi: 10.1016/j.psychsport.2021.101884

5. Cusack KW, Petrie TA, Moore EWG. Self-compassion, body satisfaction, and disordered eating symptoms in male collegiate athletes: A longitudinal analysis. Body Image 2022 Dec;43:134–142. doi: 10.1016/j.bodyim.2022.08.016

6. Voelker DK, Gould D, Reel JJ. Prevalence and correlates of disordered eating in female figure skaters. Psychol Sport Exerc 2014 Nov;15(6):696–704. doi: 10.1016/j.psychsport.2013.12.002

7. Beilharz F, Sukunesan S, Rossell SL, Kulkarni J, Sharp G. Development of a Positive Body Image Chatbot (KIT) With Young People and Parents/Carers: Qualitative Focus Group Study. J Med Internet Res 2021 Jun 16;23(6):e27807. doi: 10.2196/27807

8. Collier AF, Hagemann S, Trinidad SB, Vigil-Hayes M. Human-to-Computer Interactivity Features Incorporated Into Behavioral Health mHealth Apps: Systematic Search. JMIR Form Res 2023 Jun 30;7:e44926. PMID:37389916

9. Cheng VWS, Davenport T, Johnson D, Vella K, Hickie IB. Gamification in Apps and Technologies for Improving Mental Health and Well-Being: Systematic Review. JMIR Ment Health 2019 Jun 26;6(6):e13717. doi: 10.2196/13717

10. Litvin S, Saunders R, Maier MA, Lüttke S. Gamification as an approach to improve resilience and reduce attrition in mobile mental health interventions: A randomized controlled trial. Matsuoka YJ, editor. PLOS ONE 2020 Sep 2;15(9):e0237220. doi: 10.1371/journal.pone.0237220

11. Auf H, Dagman J, Renström S, Chaplin J. Gamification and nudging techniques for improving user engagement in mental health and well-being apps. Proc Des Soc 2021 Aug;1:1647–1656. doi: 10.1017/pds.2021.426

12. Zhang X, Zayed A, Hamrin JR, Guneysu A, & Kuoppamäki S. Enhancing Body Image Awareness through Conversations with an AI Agent: Qualitative Study with Young Adults. 2025 May. Manuscript in preparation

13. Baldó Vela D, Bonfanti N, Villarino Marín LA. Eating disorders risk assessment in men who practice team sports. Front Nutr 2022 Sep 29;9:957075. doi: 10.3389/fnut.2022.957075

14. Baceviciene M, Jankauskiene R, Rutkauskaite R. The Comparison of Disordered Eating, Body Image, Sociocultural and Coach-Related Pressures in Athletes across Age Groups and Groups of Different Weight Sensitivity in Sports. Nutrients 2023 Jun 12;15(12):2724. doi: 10.3390/nu15122724

15. Pinto AJ, Dolan E, Baldissera G, Gonçalves LDS, Jardim FGDA, Mazzolani BC, Smaira FI, Sato PDM, Scagliusi FB, Gualano B. "Despite being an athlete, I am also a human-being": Male elite gymnasts' reflections on food and body image. Eur J Sport Sci 2020 Aug;20(7):964–972. doi: 10.1080/17461391.2019.1682059

16. Collins R, Yang R, D'Souza BN, Gore A, Bay C, Erondu I, Menon S. Mobile app use and the mental health of elite athletes. Sports Psychiatry Hogrefe AG; 2025 Mar 7; Available from: https://econtent.hogrefe.com/doi/10.1024/2674-0052/a000104 [accessed May 16, 2025]

17. Murphy R, Straebler S, Cooper Z, Fairburn CG. Cognitive Behavioral Therapy for Eating Disorders. Psychiatr Clin North Am 2010 Sep;33(3):611–627. doi: 10.1016/j.psc.2010.04.004

18. Pretorius C, Chambers D, Coyle D. Young People's Online Help-Seeking and Mental Health Difficulties: Systematic Narrative Review. J Med Internet Res 2019 Nov 19;21(11):e13873. doi: 10.2196/13873

19. Newton A, Bagnell A, Rosychuk R, Duguay J, Wozney L, Huguet A, Henderson J, Curran J. A Mobile Phone-Based App for Use During Cognitive Behavioral Therapy for Adolescents With Anxiety (MindClimb): User-Centered Design and Usability Study. JMIR MHealth UHealth 2020 Dec 8;8(12):e18439. PMID:33289671

20. Haque MDR, Rubya S. An Overview of Chatbot-Based Mobile Mental Health Apps: Insights From App Description and User Reviews. JMIR MHealth UHealth 2023 May 22;11:e44838. PMID:37213181

21. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification." Proc 15th Int Acad MindTrek Conf Envisioning Future Media Environ Tampere Finland: ACM; 2011. p. 9–15. doi: 10.1145/2181037.2181040

22. Baumel A, Muench F, Edan S, Kane JM. Objective User Engagement With Mental Health Apps: Systematic Search and Panel-Based Usage Analysis. J Med Internet Res 2019 Sep 25;21(9):e14567. doi: 10.2196/14567

23. Deci EL, Ryan RM. The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. Psychol Inq Routledge; 2000 Oct 1;11(4):227–268. doi: 10.1207/S15327965PLI1104_01

24. Mamede A, Noordzij G, Jongerling J, Snijders M, Schop-Etman A, Denktas S. Combining Web-Based Gamification and Physical Nudges With an App (MoveMore) to Promote Walking Breaks and Reduce Sedentary Behavior of Office Workers: Field Study. J Med Internet Res 2021 Apr 12;23(4):e19875. doi: 10.2196/19875

25. Farooqui T, Rana T, Jafari F. Impact of Human-Centered Design Process (HCDP) on Software Development Process. 2019 2nd Int Conf Commun Comput Digit Syst C-CODE 2019. p. 110–114. doi: 10.1109/C-CODE.2019.8680978

26. Jonker L, Elferink-Gemser MT, and Visscher C. Talented athletes and academic achievements: a comparison over 14 years. High Abil Stud Routledge; 2009 Jun 1;20(1):55–64. doi: 10.1080/13598130902863691

27. Bengtsson M. How to plan and perform a qualitative study using content analysis. NursingPlus Open 2016;2:8–14. doi: 10.1016/j.npls.2016.01.001

28. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): Development and Validation Study. JMIR MHealth UHealth 2019 Apr 11;7(4):e11500. doi: 10.2196/11500

29. Iio J, Hasegawa A, Iizuka S, Hayakawa S, Tsujioka H. Ethics in Human-Centered Design. Hum-Comput Interact Theory Methods Tools Themat Area HCI 2021 Held Part 23rd HCI Int Conf HCII 2021 Virtual Event July 24–29 2021 Proc Part I Berlin, Heidelberg: Springer-Verlag; 2021. p. 161–170. doi: 10.1007/978-3-030-78462-1_12

30. World Medical Association. WMA - The World Medical Association-WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Participants. Available from: https://www.wma.net/policies-post/wma-declaration-of-helsinki/ [accessed May 14, 2025]

31. Tylka TL, Wood-Barcalow NL. The Body Appreciation Scale-2: Item refinement and psychometric evaluation. Body Image 2015 Jan;12:53–67. doi: 10.1016/j.bodyim.2014.09.006

Appendix B

Coding framework from co-design workshop including themes, sub-themes, and participant quotes.

Themes:	Sub-theme	Explanation	Participant Quote 1:	Participant Quote 2:	Participant Quote 3:	Participant Quote 4:
Preferred Features	**Chatbot for emotional support**	Participants valued the chatbot for emotional support and journaling but emphasized the need for adaptive responses.	P2: I think it's (the chatbot) something that can help you to get out of your uh negative circle.	P4: And maybe it's like an intervention itself to say stuff out loud, just not keeping it inside. Uh, allow yourself to think what you think about yourself and say what you say about yourself and hear it.		
Preferred Features	**Body Image Tracking**	Body image tracking was seen as a double-edged sword, helpful for some but potentially reinforcing negative thoughts for others.	P4: If someone tracks their body image and keeps seeing negativity, it might reinforce the negative thoughts.	P1: If I feel shit about my body, then I probably don't want to assess that. So how do you involve me to assess my body when I'm feeling shit about it? And then I'm going to see like a 1.5 out of 5.	P3: When I see this screen, I feel like, 'Oh no, it's so much. I don't have time or I don't have the energy to do it.' But if it was fewer questions at a time, it wouldn't feel so overwhelming.	P4: I think it can be good for people who want to track their progress, but only if it's not something that the app pushes on you. It should be optional, so if you want to check back and see how things have changed, you can.

Preferred Features	Adaptive Daily Challenges	Daily challenges were appreciated for actively guiding behavior toward positive body image, but they should be adaptive.	P1: Like it has to be adaptive. It can't be like you uh you have 100 challenges and then you random choose one. Like you always you you both want to build confidence and put the person out of their comfort zone at the same time but gradually. Otherwise it's uh well, it's the same thing that you say was was saying uh because I see a little bit in this app a risk of something that it I think in psychology is called negative reinforcement.	P4: I think a good thing with that feature, it's like the only feature in the app that prompts you to do things that could have positive effects on your feelings about yourself. The other ones are only like you are talking out about it, but you don't have like any eh jag tänker den här med spegelgrejen, it's like one way to help you.		
Preferred Features	Journaling and Training Log for Holistic Awareness	The training log was seen as something positive in combination with the journaling feature. Open journaling was seen as a good feature by all participants.	P1: I like the idea that this can show a trend between positive body image and performance. Or a trend with a lack of injury in my sport, haha.	P2: I think it is nice that you can track your mood across all the days, so you can note if something nice happen or if something bad happens.	P2: I'm journaling quite a lot uh recently and I think that sometimes, even if you have bad thoughts, but then uh you start to feel a bit better and then you see that the week before you were very low, but now it's going better. Maybe it helps you to say, okay, but last week I was very shitty. Now I'm feel a bit better, so it can help you a bit, I don't know, to see the improvements because sometimes you cannot see the improvements day by day, but just if you look uh from a more distant uh perspective.	

Preferred Features	Goal Setting	Goalsetting was proposed as an additional feature, but there was a potential risk if the app does not actively help you toward your goal.	P1: Say, okay the app ask you what's your goal? Yeah, my goal is to be able to dive in a bikini instead of the whole swimming costume. And then, like, those challenges could be adaptive.	P4: I totally agree that it's good to have a goal, but since the app is not an app that guides you towards the goal… it's also a risk to have a goal if someone doesn't… like, it can also be a goal that you see that, okay, I'm not moving anything forward.		
Preferred Features	Daily Affirmations	A new proposed feature was a pop-up with a prompt to write a daily positive affirmation.	P4: What are three good things with you with you with yourself today? Or there, or one good thing. Alltså, you help that you have to focus on something good today... And it can be anything. It could be that I slept well or that I did do work well or that I did a good workout, but I uh something else.	P2: Ja, but maybe not something that you have to do, otherwise it can be ordering and I can say, okay, I don't want to do this and I Yeah. Mm.		
Preferred Features	Integrated Features	Integration of chatbot with with the other features was suggested for a more seamless experience.	P2: if the chatbot is more interactive with the rest of the app maybe, so like as we said previously, you write something and then you say, oh, do you want to write this in your journal to say this or maybe it prompts you directly the daily challenge.			

Gamification Preferences	**Preference for Intrinsic over Extrinsic Motivation**	Gamification should focus on self-reflection and reinforcement rather than competition. Participants were skeptical about point-based gamification, fearing it could create pressure rather than motivation.	P3: Well I thought the function of the app or the purpose of the app is to think about body image and get like perspective and stuff and I think the coin thing will add a on a "prestationskrav".	P1: Yeah, this is a good thing, like gamifying. Uh, and like it's nice that you don't have to because if you if tomorrow you still, you still see this 5 points, then it's it's nice because Duolingo is quite toxic if you lose the streak, for example, but like here you don't I mean, you can log it, you can you can skip a day. Mhm. if you feel like crap and you still have like those points that you've earned, so that's nice, I think, as a positive feedback.		
Gamification Preferences	**Progress tracking**	Progress tracking could be available but must be optional, as seeing negative trends could reinforce negativity.	P4: (Rewritten due to mixed languages, without changing the context): I think it can be good for people who want to track their progress, but only if it's not something that the app pushes on you. It should be optional, so if you want to check back and see how things have changed, you can"	P2: sometimes, even if you have bad thoughts, but then uh you start to feel a bit better and then you see that the week before you were very low, but now it's going better. Maybe it helps you to say, okay, but last week I was very shitty. Now I'm feel a bit better, so it can help you a bit, I don't know, to see the improvements		
Gamification Preferences	**Celebratory and Visual Feedback as Motivation**	Celebrations for achieving self-set goals was suggested instead of rigid scoring.	P1: Instead of earning points, what if the app just gave encouraging messages when you complete something?	P2: Maybe instead of coins, the app could celebrate when you accomplish a goal like, show something on screen, maybe the Avatar doing a party.		

Nudging Strategies	Gentle, Personalized Nudging	Participants preferred gentle, personalized nudging rather than forced reminders.	P4: If I get a notification every day based on what I filled in my journal or what I did not do or do the day before, I think that would be a positive thing. Just reminding me of having the app on and not… on a non-demanding way	P3: It's better if a notification just asks how I'm doing instead of telling me to do something.		
Nudging Strategies	Exploratory Notifications	Notifications should be exploratory rather than directive (e.g., 'How did today's workout go?' instead of 'Fill in your training log').	P4: Not like, 'Oh, remember to fill in your daily blah blah,' more like, 'Hey, how are you feeling today?' or 'How did today's workout go?' Just like… a check-in.			
Nudging Strategies	Adaptive Notifications	Some suggested that notifications adapt based on app usage patterns (e.g., checking in more if a user logs a negative entry). Having adjustable notification frequency was seen as a good customization feature.	P2: It can be connected to something that you wrote previously. So as you said, maybe two days ago, I wrote, oh, I feel very bad and then, how do you feel today? Is it better?	P1: The more you've been away, the more you sort of stay back in the app. Like, if you haven't been journaling, maybe after a month, you do another week where you journal a little more. Something like that instead of just notifications.	P3: Can it be optional? Maybe I can click in, I want to get it once a day or once a week or only if I don't use the app, so I can choose myself?	
UI / Design	Calm Aesthetics	Simple UI and Lighter, calming colors were preferred over darker ones, associating them with journaling and well-being.	P4: I like this one best (pointing to the lighter colorscheme of the first prototype). I think the colors are more like harmoniska to me.	P3: I agree, both with the colors. Um, I think it's more tilltalande.	P1: I agree with the color, um, the color in the, um, uh, in the light are more paper, more calming. Which I associate with journaling	
UI / Design	Customizable Themes	Customizability (e.g., allowing users to switch between avatars or themes) was	P2: Maybe, like in other apps, you could switch between light and dark mode, or even between different avatar types			

		suggested to cater to different preferences.			
Avatar Preferences	**Human vs. Non-Human Avatars**	Human-like avatars were appealing to some and to others non-humanlike was prefered. The humanlike carried a risk of over-expectation from the chatbot.	P2: When you talk with ChatGPT, you don't see something that reminds you of a person or an animal, but you use it anyway. So maybe it's not so important what face you have there.	P4: It feels a bit more personal with a human being.	P1: I don't identify with this person, so I don't know.

Participants answered yes to the question if a humanlike avatar would make them expect more human-like answers.