Towards an AI-Based Therapeutic Assistant to Enhance Well-Being: Preliminary Results from a Design Science Research Project

Research Paper

Katharina-Maria Illgen[1], Enrico Kochon[2], Sergey Krutikov[1], and Oliver Thomas[1,2]

[1] German Research Center for Artificial Intelligence, Smart Enterprise Engineering, Osnabrück, Germany
{katharina.illgen,sergey.krutikov,oliver.thomas}@dfki.de
[2] Osnabrück University, Information Management and Information Systems, Osnabrück, Germany
{enrico.kochon,oliver.thomas}@uos.de

Abstract. Many individuals lack timely access to psychological support, increasing the demand for accessible digital interventions. Artificial intelligence (AI)-based solutions, particularly chatbots, demonstrate potential as scalable complements to therapy. In response to the growing reliance on AI for psychological support and unverified self-diagnoses, this study introduces *ELI*, an AI-based therapeutic assistant, offering a scientifically grounded alternative. ELI aims to complement therapy and enhance well-being by enabling users to apply evidence-based psychological strategies in daily life through real-time, low-threshold support. Adopting a Design Science Research (DSR) approach, we conducted a comprehensive literature review and expert evaluations to derive six design objectives, integrating interdisciplinary insights from information systems, human-computer interaction, and psychology. Our findings contribute design knowledge for AI-based therapeutic solutions, paving the way for broader applications, including those in the workplace. This study marks the initial design cycle, forming the foundation for future evaluation and development within a continuous DSR process.

Keywords: AI Therapeutics, Well-Being, Conversational Assistant, Design Objectives, Design Science Research

1 Introduction

Many individuals lack timely access to psychological support, fueling a growing demand for online solutions and highlighting the need for reliable, effective, and accessible digital tools (WHO 2022). While professional treatment remains essential for diagnosable conditions, Artificial Intelligence (AI)-based solutions, particularly chatbots, offer promising complements to therapy by providing accessible, anonymous, cost-effective, and scalable support, before more severe conditions arise (Hegde 2025; Yonatan-Leus & Brukner 2025).

AI-based cognitive behavioral therapy (CBT) chatbots, for instance, have demonstrated measurable efficacy in reducing depressive symptoms (He et al. 2022). Evidence also suggests that generative AI may reduce self-disclosure barriers, offering less stigmatizing alternatives to human therapists (e.g., narcissistic individuals disclosed more to generative AI than to human therapists; Wang et al. 2024). The widespread reliance on AI (e.g., ChatGPT) for psychological advice (Song et al. 2024) and the growing trend of self-diagnosis on social media (Corzine & Roy 2024) highlight both increasing demand and risks of misinformation and lack of therapeutic grounding. These developments emphasize the need for scientifically validated, evidence-based AI solutions.

This study refers to an AI-based therapeutic assistant as an interactive system that enables users to apply evidence-based psychological strategies in daily life through real-time, low-threshold support (Heron & Smyth, 2010). It focuses on well-being, a multidimensional construct that spans hedonic and eudaimonic perspectives (Gomes et al. 2023; Ryan & Deci 2001; Ryff 1989), overlaps with mental health perspectives (WHO 2025), and encompasses additional well-being-related constructs (Simons & Baldwin 2021).

While existing solutions for AI-based therapeutic support are increasingly adopted and various design challenges are recognized (e.g., Ahmad et al. 2022; Shao 2023; Thunström et al. 2024), design elements remain fragmented across studies rather than being integrated into a comprehensive framework. Moreover, approaches often target specific conditions (Li et al. 2023) rather than broader well-being or hinder interdisciplinary knowledge transfer, as research tends to be either technically or psychologically oriented (Cho et al. 2023). This study addresses these gaps by integrating a broader synthesis of interdisciplinary theory and practical expertise into a systematic, evidence-based design for an AI-based therapeutic assistant to enhance well-being. We investigate the central research question:

RQ: *How can an AI-based therapeutic assistant be designed to enhance well-being?*

Following a Design Science Research (DSR) approach (Hevner et al. 2004), this study develops and evaluates "ELI" (Empathic Listening Intelligence), a scientifically grounded, simulated prototype instantiated as an AI-based therapeutic assistant that facilitates speech interactions. DSR is particularly well-suited for developing AI-based therapeutic solutions due to its iterative cycles of design, evaluation, and refinement, ensuring both continuous improvement and a strong theoretical and practical foundation. While empathic interaction represents a core strength, the name ELI was chosen for its friendly and human-like character to foster familiarity; its design extends beyond empathic responses to address six design objectives. This study contributes (1) six theory-based design objectives informed by an integrative well-being perspective and refined through expert input, (2) a simulated prototype demonstrating their instantiation, and (3) empirical insights, establishing a foundation for future development cycles and the implementation of a working prototype, as the focus of this design cycle is on the theoretical foundation and general design logic. The paper is structured as follows: Section 2 reviews AI-based solutions for well-being; Section 3 presents the DSR methodology; Section 4 details design requirements, meta-requirements, and design objectives; Section 5 describes the instantiated prototype; Section 6 reports evaluation results; Section 7 discusses implications and directions for future research; and Section 8 concludes the paper.

2 Background

Research on AI-based solutions to complement therapy and enhance well-being has expanded across disciplines, especially information systems, human-computer interaction, and psychology (Biswas 2023; D'Alfonso 2020; Shen et al. 2022; Stephanidis et al. 2019). CBT techniques and psychoeducation are widely used to provide evidence-based psychological support, enhance therapeutic adherence, and enable scalable interventions (He et al. 2023; Sadeh-Sharvit et al. 2023). In professional contexts, solutions like "mindline at work" support self-reflection and problem-solving, helping to mitigate stress and burnout (Yoon et al. 2024). Despite their potential, current designs remain fragmented, often target specific conditions (Li et al., 2023), and limit interdisciplinary knowledge sharing (Cho et al. 2023). While some focus on empathy and trust (e.g., Shao 2023), others prioritize personalization and adaptive interventions (e.g., Ahmad et al. 2022). At the same time, ethical challenges such as transparency and communicating AI limitations are often addressed as afterthoughts (Rahsepar Mead et al. 2025).

For well-being, despite extensive scholarly attention, no universally accepted definition exists (Gomes et al. 2023; Simons & Baldwin 2021). Well-being is widely regarded as a multidimensional construct, distinguishing between hedonic (positive affect, low negative affect, life satisfaction – often referred to as subjective well-being) and eudaimonic (personal growth, purpose, self-realization) perspectives (Gomes et al. 2023; Ryan & Deci 2001; Ryff 1989). The WHO (2025) further highlights the overlap between well-being and mental health, defining mental health *as a state of well-being* that includes functional, emotional, and social capacities. Recent literature increasingly recognizes the interconnectedness of different perspectives on well-being (Martela & Sheldon 2019; VanderWeele 2017); in addition, Simons & Baldwin (2021) illustrate a concept map of terms applied to well-being, showing overlaps with constructs such as (mental) health, resilience, anxiety, depression, burnout, and flourishing.

Against this backdrop, the research question guiding this study is informed by an integrative approach to design for well-being. By integrating interdisciplinary, multidimensional well-being perspectives at the requirement level, this study contributes to extending design knowledge for AI-based therapeutic assistants beyond condition-specific interventions. The system enables evidence-based strategies in daily contexts through real-time, low-threshold conversational support; thus, incorporating principles of "Ecological Momentary Interventions" (Heron & Smyth, 2010). These have been shown to reduce stress, enhance resilience, and sustain behavioral change by delivering adaptive, context-aware support in real-world settings (Amo & Lieder 2025).

3 Design Science Research Approach

To gain design knowledge for implementing an AI-based therapeutic assistant that enhances well-being, this study adopts a DSR approach (Hevner et al. 2004). Our approach builds on the DSR process model by Sonnenberg & Vom Brocke (2012) and subsequent work by Diederich et al. (2020), adapted to our needs. In the first design cycle, as illustrated in Figure 1, we completed the following steps: (1) Justified the research problem; (2) Explored design requirements; (3) Mapping problem-oriented design requirements to abstract design goals (meta-requirements), translating these into operational system behaviors (design objectives); (4) Evaluated the design objectives; (5) Instantiated design objectives in a simulated prototype; (6) Evaluated the simulated prototype; (7) Refined the simulated prototype; and (8) Conducted a summative evaluation of the final simulated prototype through an expert survey. The feedback through the summative evaluation shall validate the *simulated prototype*, laying the groundwork for implementing a *working prototype*. This contribution represents the first design cycle in an ongoing process aimed at (1) providing a solution to the identified problem outlined in the introduction, (2) deriving general design objectives to guide future projects, and (3) presenting a simulated prototype and design foundation for further validation in subsequent iterations. Our approach emphasizes relevant and rigorous artifact development through iterative design and evaluation cycles. Future research will proceed in alignment with the next steps in the DSR process.

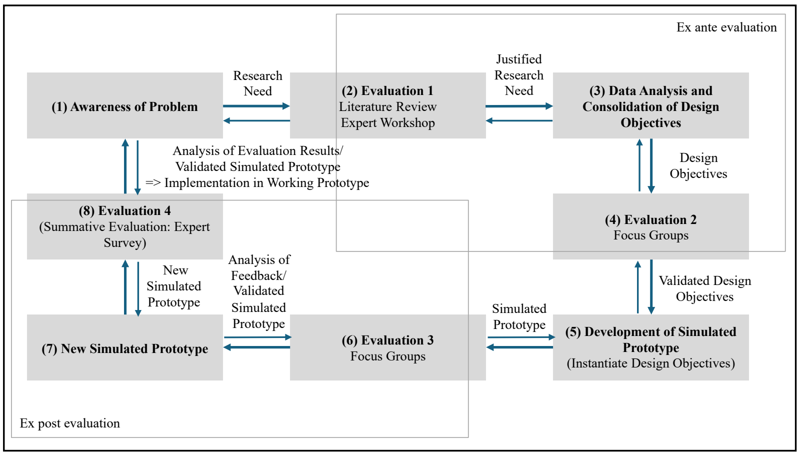

Figure 1. Design Science-based Research Approach

Our process began in December 2024. Guided by Maedche et al. (2019), awareness of the problem became imminent by seeing the need of many people seeking psychological help and not receiving it on time. A systematic literature review and workshop with psychologists confirmed the research need for innovative information systems to enhance well-being and identified relevant requirements for an AI-based therapeutic assistant. Following Vom Brocke et al. (2015), a sequential literature search was conducted to identify relevant articles early in the review process. Bibliographic databases were searched using an iteratively developed keyword-based search string, refined through multiple iterations to ensure relevance. The final search string was:

("therapy" OR "therapist" OR "psychologist" OR "counseling") AND ("conversational AI" OR "ChatGPT" OR "virtual assistant" OR "dialogue system" OR "chatbot" OR "Artificial Intelligence" OR "AI") AND ("mental health" OR "well-being" OR "emotional support" OR "psychological support" OR "psychological resilience").

We conducted a comprehensive search across major academic databases, retrieving 640 articles: Scopus (309), ScienceDirect (42), Web of Science (25), PubMed (181), and SpringerLink (46). Additionally, 37 articles were extracted by screening Google Scholar to capture the most recent developments (published since 2024). Inclusion criteria focused on peer-reviewed articles at the intersection of psychology and information systems, published in English between 2022 and January 2025. This timeframe reflects the emergence of large language models (LLMs) like ChatGPT in 2022 and the growing interest in AI-psychotherapy intersections. To ensure objectivity, multiple researchers oversaw the search process. Different database-specific search notations were applied (supplementary data (1) is available for online access via: *https://s.academic-cloud.de/Hqsf2X*), and filters were aligned with the stated inclusion criteria. All 638 articles were available either openly or through institutional access. We removed 16 duplicates, leaving 622 articles for further consideration, and screened the titles and tables of contents, excluding 379 articles. Of the remaining 243, we screened the abstract for relevance, narrowing the selection to 162. Finally, a full-text screening was performed on the remaining articles. To avoid redundancy, we focused on the most relevant findings, ultimately selecting 73 articles from which requirements were extracted.

To complement the literature-derived requirements, additional insights were gathered through an expert workshop. Experts were recruited via an internal psychologist network. Five participants, aged 24 to 33, were selected based on their diverse psychotherapy expertise. The 60-minute workshop began with an introduction to the study's aim, followed by a brainstorming session to identify requirements and a group discussion to refine and evaluate these. Key points were documented collaboratively using Microsoft Whiteboard, and saturation was achieved as no new requirements emerged. While many findings overlapped with the literature, two novel additions complemented the final requirements matrix.

From the final requirements, seven meta-requirements (MRs) and six design objectives (DOs) were derived by the researchers following the anatomy of Gregor et al. (2020) and the structure of Stattkus et al. (2024).

In the iterative evaluation phase, the same workshop participants reviewed the DOs in focus groups, as guided by Myers & Newman (2007), discussing challenges and good practices as well as implementation goals. While providing valuable practical insights, this resulted in only minor adjustments in the MRs and DOs (Section 4; Figure 2). In the subsequent iteration, the six DOs were instantiated into a video-based, simulated prototype. This prototype was presented to the same experts, who provided feedback for modifications, again, in the form of focus groups. The feedback was incorporated into a revised simulated prototype (details on how the DOs were instantiated, including expert feedback, see Section 5).

The updated version underwent a summative evaluation (Section 6) through an expert survey to gather feedback for subsequent design cycles. From the results of the summative evaluation, we derive findings that provide a basis for creating shared understandings of ELI and for implementing a working prototype in the next design cycle. The evaluation was conducted through an expert survey via Microsoft Forms from 18.02.2025 to 06.03.2025, with 18 professionals in the field of psychology recruited from an internal network of psychologists. The survey was preceded by a pre-test with eight researchers in the HCAI field, conducted to ascertain the comprehensibility of the questionnaire as well as its validity, indicating good internal consistency. The pretest resulted in minor adjustments. The final expert survey began with a presentation of the simulated prototype to the experts, followed by demographic questions, an attention check, and a combination of closed and open questions to evaluate ELI. The closed questions utilized well-established scales for the fulfillment of the derived DOs, which variables achieved Cronbach's alpha values of 0.65 to 0.864; thus, reliable results can be assumed (Blanz 2021). In addition to the quantitative measurements, we included open-ended questions to enrich and refine the DOs and to derive further insights concerning positive aspects and areas for improvement.

This first design cycle delivers theoretical grounding, traceable design logic, and an empirically validated simulated prototype for future development.

4 Design Requirements, Meta-Requirements, and Design Objectives

The literature review and expert workshop identified 15 requirements (supplementary data (2) is available for online access via: *https://s.academiccloud.de/Hqsf2X*). These were synthesized into seven MRs and six DOs (Figure 2).

MR1 and DO1 emphasize empathy, trust, and user-centered design. ELI simulates cognitive empathy, including psychotherapeutic dialogue (R14; expert workshop), fostering therapeutic relationships (Rubin et al. 2024). While human-like cues can increase trust (Park et al. 2024), overly anthropomorphic designs may reduce usability (Golubović et al. 2024; Thunström et al. 2024) or lead to a feeling of uncanniness when agents become too human-like (Ahmad et al. 2022; Kim et al. 2019). Thus, ELI prioritizes functional aesthetic consistency, non-stigmatizing, and intuitive interfaces over anthropomorphic features; minimal human-like cues may be used only if functionally justified (this aligns with the feedback from the focus groups in the second evaluation).

MR2 and DO2 address dynamic adaptability to create personalized and engaging therapeutic experiences that enhance well-being. Adaptive conversational styles can improve user engagement and effectiveness (Lee et al. 2024; Ahmad et al. 2022). Immersive environments (e.g., Virtual Reality (VR) or Augmented Reality (AR)) can enhance therapy immersion, creating more compelling and adaptive therapeutic experiences (Cerasa et al. 2022).

MR3, MR6, and DO3 concern ethical standards, crisis intervention, and AI transparency, ensuring that ELI's limitations are delineated. ELI functions as a supportive tool rather than a replacement for therapists, ensuring ethical responsibility (Khawaja

& Bélisle-Pipon 2023). Crisis mechanisms must facilitate human intervention when necessary (Omarov et al. 2023; Heston 2023). Continuous quality assurance is essential (R15; expert workshop).

MR4 and DO4 support seamless integration and collaboration. ELI shall be designed to seamlessly integrate into existing workflows, support human therapists, or incorporate features like mood tracking and wearable technology for real-time feedback (Desage et al. 2024; Ahmed et al. 2023).

MR5 and DO5 highlight evidence-based algorithms and psychotherapeutic dialogue. Interventions must adhere to validated therapeutic methods, such as CBT, and ensure ongoing validation for reliability (Martinengo et al. 2022; Sadeh-Sharvit et al. 2023). CBT provides strategies targeting multiple well-being dimensions (e.g., positive affect, resilience, personal growth, and self-realization).

MR7 and DO6 focus on accessibility and real-time support. LLM integration supports natural language interactions, enhancing the adaptability and accessibility of ELI to enhance well-being in everyday contexts (Guo et al. 2024; Caceres Najarro et al. 2023); also relevant for all MRs and DOs above.

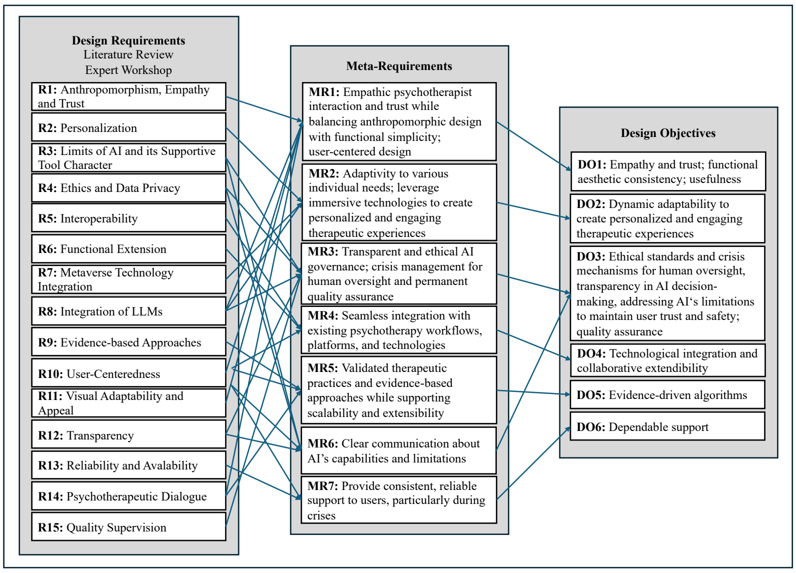

Figure 2. Requirements, Meta-Requirements, and Design Objectives for ELI

5 Development of ELI: Instantiating Design Objectives in a Simulated Prototype

To fulfill **DO1**, ELI is designed as an animated bubble, driven by Babylon.js[1], avoiding inconsistent anthropomorphic cues. Instead, minimal human-likeness is conveyed

[1] https://www.babylonjs.com/, last accessed on 21.02.2025.

through speech-to-speech interaction and empathic language. Speech-to-speech enhances engagement, while Python-based models ("Hugging Face"[2]) and Llama's[3] conversational capabilities ensure nuanced responses. A calming color and voice scheme fosters user comfort. Dynamic emotional adaptation was not considered in the revised simulated prototype, as experts noted in the focus groups (third evaluation) that it poses challenges, as it might not always align perfectly with the user's emotional state, potentially causing confusion or discomfort, particularly significant for vulnerable users. ELI's non-intrusive design prevents overstimulation, balancing empathy and functionality without overwhelming the user. This resulted in the visual representation of ELI (Figure 3). The name ELI was chosen as a friendly and human-like name.

Figure 3. Interaction with ELI (left); Screenshot of ELI during interaction (right)

DO2 was instantiated via adaptive design. Focus groups in the third evaluation highlighted user engagement through customizable environments (e.g., forest, ocean). AR via WebXR[4] and Babylon.js was chosen as it provides immersive 360-degree settings for a deeper sense of presence and a certain degree of immersion. Further feasibility options will be evaluated during the summative evaluation. The initial simulated prototype was developed in German to match the language of the evaluation participants.

For **DO3**, ELI includes features such as hotline support, emergency detection, and explainable AI to clarify limitations. User feedback informs quality assurance, with anonymous metadata collected for hyperparameter optimization (Ouyang et al. 2022). Crisis response mechanisms ensure transparency and security, demonstrated in the simulated prototype and planned for rule-based integration in the working prototype.

DO4 supports encrypted session exports for therapists and collaboration options. Future iterations may connect users to self-help groups. ELI is designed as part of a broader therapeutic ecosystem, whose feasibility options will be evaluated in the summative evaluation.

DO5 ensures evidence-based support. CBT scripts dynamically adjust to user input. Future iterations will employ "Retrieval-Augmented Generation" (RAG), verified response grounding, and improved reliability. Leveraging Llama's conversational strengths, ELI maintains structured yet adaptive support.

[2] https://huggingface.co/, last accessed on 10.02.2025.
[3] https://www.llama.com/, last accessed on 21.02.2025.
[4] https://immersiveweb.dev/, last accessed on 19.02.2025.

For **DO6**, ELI provides continuous support via smartphones due to their easy accessibility and wide reach (Amo & Lieder 2025). The feasibility of more specific devices will be assessed through feedback collected during the summative evaluation.

6 Evaluation

The final sample of the summative evaluation consisted of 18 experts: two identified as male, 15 as female, and one did not specify. Participants' ages ranged from 18 to 29 (*n* = 5), 30 to 39 (*n* = 11), and 40 to 49 (*n* = 2). Their psychological professions included clinical psychology and psychotherapy (*n* = 9), work, organizational and business psychology (*n* = 6), legal psychology (*n* =1), educational psychology (*n* = 1), research psychology (*n* = 1), neuropsychology (*n* = 1) and other, specified as "prevention (mental health) in organizations" (*n* = 1). Some were active in several areas (*n* = 2). Their work experience ranged from less than one year (*n* = 2), one to two years (*n* = 3), three to five (*n* = 7), six to 10 (*n* = 5), and 11 to 15 years (*n* = 1).

To address the research question, we first quantitatively surveyed the fulfillment of the derived DOs using descriptive statistics. Empathy, trust, and attractiveness (DO1) were measured using the scales by Charrier et al. (2019), Borsci et al. (2022), and Kemper et al. (2012), while perceived usefulness was assessed by Davis et al. (1989). An open-ended question further explored expert perspectives on ELI's visual design. Personalized therapeutic interventions (DO2) were evaluated using the "recognition and facilitation of users' goal and intent" subscale (Borsci et al. 2022), along with qualitative feedback on the feasibility of implementing immersive technology. Ethical standards and crisis mechanisms (DO3) were assessed using the "users' privacy and security" subscale (Borsci et al. 2022). Integration (DO4) was examined via experts' agreement towards the current feasibility option of therapist protocol forwarding, complemented by a question concerning the feasibility of peer-to-peer support. Evidence-based algorithms (DO5) were evaluated with the "psychotherapeutic competencies (CBT-based)" scale (Weck et al. 2010), ensuring validated therapeutic approaches. Finally, dependable support and accessibility (DO6) were measured using the "accessibility" subscale (Borsci et al. 2022), complemented by a question regarding device preferences. Finally, we asked about overall impression, recommendation likelihood, and potential target groups.

The average fulfillment is shown in Figure 4. Experts rated their responses on a five-point Likert scale ("1 – strongly disagree" to "5 = strongly agree"). On average, they agreed with the measured DOs: Empathy scored highest at 4.32 (*SD* = 0.62), followed by accessibility at 4.39 (*SD* = 0.73), recognition and facilitation of users' goal and intent at 4.11 (*SD* = 0.66), psychotherapeutic competencies (CBT-based) at 4.03 (*SD* = 0.65), perceived usefulness at 4.02 (*SD* = 0.74), and user privacy/security at 4.02 (*SD* = 0.76). Lower agreement was found for conversational credibility at 3.74 (*SD* = 0.71), integration feasibility of protocol forwarding at 3.44 (*SD* = 0.86), and attractiveness at 3.39 (*SD* = 1.09).

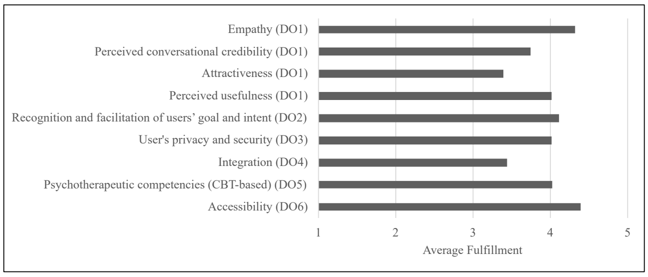

Figure 4. Average fulfillment with the evaluated constructs

Experts provided insights for refining the DOs. For attractiveness (DO1), they suggested improvements in visual design, including customizable avatars (n = 1), a calmer background (n = 1), and a color change for ELI (n = 1). Regarding recognition and facilitation of users' goal and intent (DO2), experts assessed immersive technology options: 28% (n = 11) supported AR, while an equal proportion preferred no immersive technology. Additionally, 23% (n = 9) favored VR (e.g., VR glasses), and 21% (n = 8) suggested using a virtual environment as a screen background. For further integration (DO4), peer-to-peer support received higher agreement (M = 3.89, SD = 0.76) than the current integration option. Regarding accessibility (DO6), smartphones were the preferred device (n = 17), though one expert recommended multi-device access (n = 1). When identifying target groups for ELI, 25% (n = 12) suggested adults, followed by employees (21%, n = 10), young people (17%, n = 8), and individuals with specific needs (13%, n = 6). Another 13% (n = 6) recommended general use, while older adults (6%, n = 3) and niche groups (6%, n = 3) were mentioned less frequently. These niche groups included “persons with specific problems but no serious mental illnesses” (n = 1), “persons with sufficient reflective skills” (n = 1), and “people with social anxiety” (n = 1). Overall, experts rated ELI positively on a scale from “1 – negative” to “5 – positive” (M = 4.17, SD = 0.62). However, when asked if they would recommend ELI to patients, experts were more cautious. The willingness to recommend had an average rating of 3.56 (SD = 0.62), from “1 – strongly disagree” to “5 – strongly agree”.

Second, a qualitative analysis of expert responses identified strengths and areas for further development regarding ELI. Open-ended answers were categorized into two clusters, with two codes representing positive aspects and three addressing potential improvements.

Cluster 1 addresses the strengths and positive aspects of ELI. Participants frequently highlighted ELI's usability and accessibility as significant strengths (Cluster 1.1; n = 6). Experts valued the intuitive, speech-based interaction and immediate feedback, features often lacking in traditional therapy. ELI was perceived as user-friendly, always available, and easily accessible, allowing users to seek support anonymously and without waiting times. Experts generally agreed that ELI is most beneficial for addressing less severe psychological issues. Additionally, ELI's psychological and therapeutic support was positively evaluated (Cluster 1.2; n = 5). Respondents noted that ELI demonstrates empathic listening skills, fostering an environment in which users feel genuinely understood. The ability to offer concrete, tailored problem-solving strategies for mild psychological issues was praised. Several experts viewed ELI as a valuable complement to traditional therapy, providing personalized support between professional sessions.

Cluster 2 addresses areas for improvement of ELI. One frequently mentioned concern was related to data privacy (Cluster 2.1; $n = 4$). Experts emphasized the necessity for greater transparency regarding how user data, particularly sensitive mental health information, is handled and protected. Clearer communication of data security measures was strongly recommended to enhance user trust. Another significant area identified for development was crisis handling capabilities (Cluster 2.2; $n = 5$). Respondents recommended that ELI include features for detecting and addressing acute psychological crises, such as suicidal ideation. Incorporating additional therapeutic approaches, particularly systemic therapy, was suggested to enable a more comprehensive psychological intervention. Experts further advised that ELI should recognize and integrate users' pre-existing conditions, thereby providing more targeted recommendations. Finally, some experts commented on the conversational interactivity of ELI (Cluster 2.3; $n = 2$). Although the speech-based interaction was appreciated for its naturalness, improving the balance between user input and system responses was recommended. Adjusting this balance could enhance the dynamic nature of dialogues, making interactions feel more engaging and responsive.

In summary, experts positively evaluated ELI's accessibility and psychological support but identified essential areas for improvement, particularly concerning data privacy, crisis response capabilities, therapeutic comprehensiveness, and conversational interaction quality.

7 Discussion and Outlook

This study presents the initial design and evaluation of ELI, an AI-based therapeutic assistant, using a DSR approach. It contributes (1) six systematically derived DOs informed by an integrative well-being perspective and expert input; (2) a simulated prototype instantiating these DOs; and (3) empirical insights from an initial evaluation, laying the groundwork for future research. This study *contributes to research* by providing a structured, evidence-based design framework that links design requirements, abstract goals (MRs), and instantiable system behaviors (DOs), enhancing design traceability and addressing fragmentation in existing literature. Unlike approaches focused solely on psychological theory or technical development, our work integrates knowledge from psychology, information systems, and human-computer interaction, fostering interdisciplinary knowledge transfer (Cho et al. 2023). By incorporating multidimensional perspectives on well-being at the design requirement level, we extend design knowledge for AI-based therapeutic assistants beyond condition-specific interventions. ELI promotes everyday well-being through real-time, low-threshold interventions grounded in evidence-based strategies, offering *practical guidance*.

Evaluation findings indicate that ELI provides empathic feedback, simulating therapeutic aspects while ensuring accessible support. Experts generally supported the proposed DOs and evaluated ELI positively ($M = 4.17$; $SD = 0.62$), particularly for its empathic response capability ($M = 4.32$; $SD = 0.62$). However, lower ratings for conversational credibility ($M = 3.74$; $SD = 0.71$) and recommendation likelihood ($M = 3.56$; $SD = 0.62$) highlight areas for improvement. Smartphones were the preferred device, while VR-based implementation emerged as a feasible option, particularly targeting adults, employees, and young people. Experts emphasized ELI's accessibility, user-friendly design, and immediate availability as key strengths, alongside its ability to deliver tailored, low-threshold, and anonymous support. Nonetheless, concerns were raised regarding data privacy, underscoring the need for enhanced transparency and robust crisis intervention mechanisms. Experts also recommended integrating additional therapeutic frameworks (e.g., systemic therapy), refining interaction flow, and expanding customization options for ELI's visual design.

Several limitations should be acknowledged. The small, specialized expert sample limits generalizability, and perspectives from end-users remain unexplored. Thus, results are exploratory and require broader validation to ensure generalizability and mitigate selection bias.

The next DSR cycle will focus on (1) implementing a working prototype, (2) strengthening crisis response and privacy safeguards, (3) refining personalization and interaction quality, and (4) exploring peer support features. VR implementations and workplace scenarios will also be examined based on expert feedback. To deepen theoretical grounding, kernel theory (Möller et al. 2022) may be used to link design elements to well-being outcomes. Future evaluations may apply inferential statistics and validated well-being scales to evaluate the robustness of the DOs across user groups and systematically assess their impact on enhancing well-being. Broader evaluations with end-users are planned to ensure therapeutic relevance and empirical validation. Through iterative refinement, this research aims to establish a validated AI-based therapeutic assistant that complements therapy while meeting ethical and technical standards.

8 Conclusion

This contribution outlines our initial efforts in an ongoing DSR project to develop ELI, an AI-based therapeutic assistant to enhance well-being. Grounded in interdisciplinary research and practical expertise in the field of psychology, we derived seven MRs and six DOs. Since the derivation process from MRs to DOs was based on the researchers' perspective, the next step was to evaluate the DOs in focus groups ("Evaluation 2" in the DSR process depicted in Figure 1), leading to only minor adjustments. Based on this, a simulated prototype was presented and subsequently evaluated by these experts, again, in focus groups ("Evaluation 3"), which led to a few refinements and to the final simulated prototype that was ready for the "Evaluation 4", the summative evaluation. From this, we derive findings that provide a basis for creating shared meanings of ELI as well as a basis for implementing a working prototype. As this project aims to progress through DSR, the findings presented in this study serve as an initial foundation for designing AI-based therapeutic assistants to enhance well-being, as well as a starting point for the future development and evaluation of ELI.

References

Ahmad, R., Siemon, D., Gnewuch, U. & Robra-Bissantz, S. (2022), Designing personality-adaptive conversational agents for mental health care. Information Systems Frontiers, 24(3), 923–943.

Ahmed, A., Hassan, A., Aziz, S., Abd-Alrazaq, A. A., Ali, N., Alzubaidi, M., Al-Thani, D., Elhusein, B., Siddig, M. A., Ahmed, M. & Househ, M. (2023), Chatbot features for anxiety and depression: A scoping review. Health Informatics Journal, 29(1), 14604582221146720. https://doi.org/10.1177/14604582221146719.

Amo, V., & Lieder, F. (2025), Evaluating the Effectiveness of InsightApp for Anxiety, Valued Action, and Psychological Resilience: Longitudinal Randomized Controlled Trial. JMIR Mental Health, 12(1), e57201.

Biswas, S. S. (2023), Role of chat gpt in public health. *Annals of biomedical engineering*, *51*(5), 868-869.

Blanz, M. (2021), Forschungsmethoden und Statistik für die soziale Arbeit: Grundlagen und Anwendungen. Kohlhammer Verlag.

Borsci, S., Malizia, A., Schmettow, M., Van Der Velde, F., Tariverdiyeva, G., Balaji, D. & Chamberlain, A. (2022), The chatbot usability scale: the design and pilot of a usability scale for interaction with AI-based conversational agents. Personal and Ubiquitous Computing, 26, 95–119.

Caceres Najarro, L. A., Lee, Y., Toshnazarov, K. E., Jang, Y., Kim, H. & Noh, Y. (2023), WMGPT: Towards 24/7 Online Prime Counseling with ChatGPT. Adjunct Proceedings of the 2023 ACM International Joint Conference on Pervasive and Ubiquitous Computing & the 2023 ACM International Symposium on Wearable Computing, 142–145.

Cerasa, A., Gaggioli, A., Marino, F., Riva, G. & Pioggia, G. (2022), The promise of the metaverse in mental health: the new era of MEDverse. Heliyon, 8(11).

Charrier, L., Rieger, A., Galdeano, A., Cordier, A., Lefort, M. & Hassas, S. (2019), The rope scale: a measure of how empathic a robot is perceived. 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 656–657.

Cho, Y. M., Rai, S., Ungar, L., Sedoc, J. & Guntuku, S. C. (2023), An integrative survey on mental health conversational agents to bridge computer science and medical perspectives. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing* (Vol. 2023, p. 11346).

Corzine, A., & Roy, A. (2024), Inside the black mirror: current perspectives on the role of social media in mental illness self-diagnosis. Discover Psychology, 4(1), 40.

D'Alfonso, S. (2020), AI in mental health. *Current opinion in psychology*, *36*, 112-117.

Davis, F. D., Bagozzi, R. P. & Warshaw, P. R. (1989), User acceptance of computer technology: A comparison of two theoretical models. Management Science, 35(8), 982–1003.

Desage, C., Bunge, B. & Bunge, E. L. (2024), A Revised Framework for Evaluating the Quality of Mental Health Artificial Intelligence-Based Chatbots. Procedia Computer Science, 248, 3–7. https://doi.org/https://doi.org/10.1016/j.procs.2024.10.356.

Diederich, S., Brendel, A. B. & Kolbe, L. M. (2020), Designing anthropomorphic enterprise conversational agents. Business & Information Systems Engineering, 62, 193–209.

Golubović, G., Đurđević, S., Kerac, J., Juretić, I., & Cvetojević, S. (2024). Use Of Ai In Mhealth Mobile Applications: The Influence of Virtual Therapists'visual Characteristics on User Perception. Proceedings-The Twelfth International Symposium Grid 2024.

Gomes, R., Mathew, J., Nair, S., Mulasi, A. & Yadav, P. (2023), Design and validation of the digital well-being scale. Journal of Theories and Research in Education, 18(1), pp. 239-251.

Gregor, S., Chandra Kruse, L., & Seidel, S. (2020), Research perspectives: the anatomy of a design principle. Journal of the Association for Information Systems, 21(6), 2.

Guo, Z., Lai, A., Thygesen, J. H., Farrington, J., Keen, T., & Li, K. (2024), Large Language Models for Mental Health Applications: Systematic Review. JMIR Mental Health, 11, e57400. https://doi.org/10.2196/57400.

Hegde, A. (2025), AI-Assisted Mental Health Interventions through Chatbot Therapies.

He, Y., Yang, L., Qian, C., Li, T., Su, Z., Zhang, Q. & Hou, X. (2023), Conversational Agent Interventions for Mental Health Problems: Systematic Review and Meta-analysis of Randomized Controlled Trials. Journal of Medical Internet Research, 25, e43862. https://doi.org/10.2196/43862.

He, Y., Yang, L., Zhu, X., Wu, B., Zhang, S., Qian, C. & Tian, T. (2022), Mental Health Chatbot for Young Adults With Depressive Symptoms During the COVID-19 Pandemic: Single-Blind, Three-Arm Randomized Controlled Trial. Journal of Medical Internet Research, 24(11), e40719. https://doi.org/10.2196/40719.

Heron, K. E. & Smyth, J. M. (2010), Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. British Journal of Health Psychology, 15(1), 1–39.

Heston, T. F. (2023), Safety of large language models in addressing depression. Cureus, 15(12).

Hevner, A. R., March, S. T., Park, J. & Ram, S. (2004). Design science in information systems research. MIS Quarterly, 75–105.

Kemper, C., Lutz, J., Margraf-Stiksrud, J., Beierlein, C., Kovaleva, A. & Rammstedt, B. (2012), Eine Ein-Item-Skala zur Einschätzung von Attraktivität: Das Attraktivitätsrating (AR1).

Kim, S. Y., Schmitt, B. H., & Thalmann, N. M. (2019). Eliza in the uncanny valley: Anthropomorphizing consumer robots increases their perceived warmth but decreases liking. *Marketing letters*, *30*, 1-12.

Khawaja, Z. & Bélisle-Pipon, J.-C. (2023), Your robot therapist is not your therapist: understanding the role of AI-powered mental health chatbots. Frontiers in Digital Health, 5, 1278186.

Lee, J., You, J. S. & Lee, D. (2024), User Preferences on AI Psychotherapy Based on Moderating Effects of Individual Personality Traits: Employing a Clustering Analysis. International Journal of Human–Computer Interaction, 1–18.

Li, H., Zhang, R., Lee, Y.-C., Kraut, R. E. & Mohr, D. C. (2023), Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. NPJ Digital Medicine, 6(1), 236.

Maedche, A., Gregor, S., Morana, S. & Feine, J. (2019), Conceptualization of the problem space in design science research. Extending the Boundaries of Design Science Theory and Practice: 14th International Conference on Design Science Research in Information Systems and Technology, DESRIST 2019, Worcester, MA, USA, June 4–6, 2019, Proceedings 14, 18–31.

Martela, F. & Sheldon, K. M. (2019), Clarifying the concept of well-being: Psychological need satisfaction as the common core connecting eudaimonic and subjective well-being. *Review of General Psychology*, *23*(4), 458-474.

Martinengo, L., Jabir, A. I., Goh, W. W. T., Lo, N. Y. W., Ho, M.-H. R., Kowatsch, T., Atun, R., Michie, S. & Tudor Car, L. (2022), Conversational Agents in Health Care: Scoping Review of Their Behavior Change Techniques and Underpinning Theory. Journal of Medical Internet Research, 24(10), e39243. https://doi.org/10.2196/39243.

Möller, F., Schoormann, T., Strobel, G. & Hansen, M. R. P. (2022), Unveiling the Cloak: Kernel Theory Use in Design Science Research. ICIS 2022 Proceedings.

Myers, M. D. & Newman, M. (2007), The qualitative interview in IS research: Examining the craft. Information and Organization, 17(1), 2–26.

Omarov, B., Narynov, S. & Zhumanov, Z. (2023), Artificial Intelligence-Enabled Chatbots in Mental Health: A Systematic Review. Computers, Materials & Continua, 74(3).

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K. & Ray, A. (2022), Training language models to follow instructions with human feedback. Advances in Neural Information Processing Systems, 35, 27730–27744.

Park, G., Chung, J., & Lee, S. (2024). Human vs. machine-like representation in chatbot mental health counseling: the serial mediation of psychological distance and trust on compliance intention. Current Psychology, 43(5), 4352-4363.

Rahsepar Meadi, M., Sillekens, T., Metselaar, S., van Balkom, A., Bernstein, J. & Batelaan, N. (2025), Exploring the ethical challenges of conversational AI in mental health care: scoping review. *JMIR mental health*, *12*, e60432.

Rubin, M., Arnon, H., Huppert, J. D. & Perry, A. (2024), Considering the Role of Human Empathy in AI-Driven Therapy. JMIR Mental Health, 11(1), e56529.

Ryan, R. M. & Deci, E. L. (2001), On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual review of psychology*, *52*(1), 141-166.

Ryff, C. D. (1989), Happiness is everything, or is it? Explorations on the meaning of psychological well-being. Journal of personality and social psychology, 57(6), 1069.

Sadeh-Sharvit, S., Camp, T. Del, Horton, S. E., Hefner, J. D., Berry, J. M., Grossman, E. & Hollon, S. D. (2023), Effects of an Artificial Intelligence Platform for Behavioral Interventions on Depression and Anxiety Symptoms: Randomized Clinical Trial. Journal of Medical Internet Research, 25, e46781. https://doi.org/10.2196/46781.

Shao, R. (2023), An Empathetic AI for Mental Health Intervention: Conceptualizing and Examining Artificial Empathy. Proceedings of the 2nd Empathy-Centric Design Workshop, 1–6.

Shen, J., Iandoli, L. & Aguirre-Urreta, M. (2022), Human-Centered Design for Individual and Social Well-being: Editorial Preface. AIS Transactions on Human-Computer Interaction, 14(4), 446–460. https://doi.org/https://doi.org/10.17705/1thci.00175.

Simons, G. & Baldwin, D. S. (2021), A critical review of the definition of 'wellbeing'for doctors and their patients in a post Covid-19 era. *International Journal of Social Psychiatry*, *67*(8), 984-991.

Song, I., Pendse, S. R., Kumar, N. & De Choudhury, M. (2024), The typing cure: Experiences with large language model chatbots for mental health support. ArXiv Preprint ArXiv:2401.14362.

Sonnenberg, C. & Vom Brocke, J. (2012), Evaluations in the science of the artificial–reconsidering the build-evaluate pattern in design science research. Design Science Research in Information Systems. Advances in Theory and Practice: 7th International Conference, DESRIST 2012, Las Vegas, NV, USA, May 14-15, 2012. Proceedings 7, 381–397.

Stattkus, D., Göritz, L., Illgen, K.-M., Beinke, J. H. & Thomas, O. (2024), A Blueprint for Breaking Stereotypes: Design Principles for Conversational Agents to Empower Girls in Computer Science Education. PACIS 2024 Proceedings. 2.

Stephanidis, C., Salvendy, G., Antona, M., Chen, J. Y. C., Dong, J., Duffy, V. G., Fang, X., Fidopiastis, C., Fragomeni, G. & Fu, L. P. (2019), Seven HCI grand challenges. International Journal of Human–Computer Interaction, 35(14), 1229–1269. https://doi.org/https://doi.org/10.1080/10447318.2019.1619259.

Thunström, A. O., Carlsen, H. K., Ali, L., Larson, T., Hellström, A. & Steingrimsson, S. (2024), Usability Comparison Among Healthy Participants of an Anthropomorphic Digital Human and a Text-Based Chatbot as a Responder to Questions on Mental Health: Randomized Controlled Trial. JMIR Human Factors, 11. https://doi.org/https://doi.org/10.2196/54581.

VanderWeele, T. J. (2017), On the promotion of human flourishing. Proceedings of the National Academy of Sciences, 114(31), 8148-8156.

Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R. & Cleven, A. (2015), Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. Communications of the Association for Information Systems, 37(1), 9.

Wang, P., Yue, Y., Ouyang, M., Yin, L., Yin, Y. & Li, B. (2024), Do Narcissistic People Exhibit More Authentic Self-Disclosure to Generative AI? The Roles of Short-Form Video Addiction, Loneliness, and Usage Intention. Social Science Computer Review, 08944393241308511.

Weck, F., Hautzinger, M., Heidenreich, T. & Stangier, U. (2010), Erfassung psychotherapeutischer Kompetenzen. Zeitschrift Für Klinische Psychologie Und Psychotherapie.

World Health Organization (2022), COVID-19 pandemic triggers 25% increase in prevalence of anxiety and depression worldwide. https://www.who.int/news/item/02-03-2022-covid-19-pandemic-triggers-25-increase-in-prevalence-of-anxiety-and-depression-worldwide. Accessed: 15.01.2025.

World Health Organization (2025), Health and Well-Being. https://www.who.int/data/gho/data/major-themes/health-and-well-being. Accessed: 04.05.2025.

Yonatan-Leus, R. & Brukner, H. (2025), Comparing perceived empathy and intervention strategies of an AI chatbot and human psychotherapists in online mental health support. Counselling and Psychotherapy Research, 25(1), e12832.

Yoon, S., Goh, H., Low, X. C., Weng, J. H. & Heaukulani, C. (2024), User perceptions and utilisation of features of an AI-enabled workplace digital mental wellness platform ‘mindline at work.’ BMJ Health & Care Informatics, 31(1).