

Article

Using Large Language Models to Embed Relational Cues in the Dialogue of Collaborating Digital Twins [†]

Sana Salman and Deborah Richards * 

School of Computing, Foundations of AI Research Centre, Macquarie University, Sydney, NSW 2109, Australia; sana.salman@students.mq.edu.au

* Correspondence: deborah.richards@mq.edu.au

[†] This paper is a revised version of our paper published 22nd International Conference, PAAMS 2024, Salamanca, Spain, 26–28 June 2024, Proceedings.

Abstract: Embodied Conversational Agents (ECAs) serve as digital twins (DTs), visually and behaviorally mirroring human counterparts in various roles, including healthcare coaching. While existing research primarily focuses on single-coach ECAs, our work explores the benefits of multi-coach virtual health sessions, where users engage with specialized diet, physical, and cognitive coaches simultaneously. ECAs require verbal relational cues—such as empowerment, affirmation, and empathy—to foster user engagement and adherence. Our study integrates Generative AI to automate the embedding of these cues into coaching dialogues, ensuring the advice remains unchanged while enhancing delivery. We employ ChatGPT to generate empathetic and collaborative dialogues, comparing their effectiveness against manually crafted alternatives. Using three participant cohorts, we analyze user perception of the helpfulness of AI-generated versus human-generated relational cues. Additionally, we investigate whether AI-generated dialogues preserve the original advice’s semantics and whether human or automated validation better evaluates their lexical meaning. Our findings contribute to the automation of digital health coaching. Comparing ChatGPT- and human-generated dialogues for helpfulness, users rated human dialogues as more helpful, particularly for working alliance and affirmation cues, whereas AI-generated dialogues were equally effective for empowerment. By refining relational cues in AI-generated dialogues, this research paves the way for automated virtual health coaching solutions.

Keywords: embodied conversational agents; relational cues; digital twin; health coaching; large language models



Academic Editors: Philippe Mathieu, Fernando De la Prieta Pintado and Alfonso González-Briones

Received: 2 April 2025

Revised: 1 May 2025

Accepted: 4 May 2025

Published: 6 May 2025

Citation: Salman, S.; Richards, D. Using Large Language Models to Embed Relational Cues in the Dialogue of Collaborating Digital Twins. *Systems* **2025**, *13*, 353. <https://doi.org/10.3390/systems13050353>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Embodied Conversational Agents (ECAs), act as digital twins (DTs) representing visually and behaviourally the human counterpart in the role they are playing. When it comes to DTs for health care, ECAs can be simulated to represent actual in-person coaching with single or multiple coaches. Gámez Díaz, Yu [1] drew a connection between smart coaching and a digital twin, since both share certain characteristics, like having artificial intelligence as a core, being able to communicate in near real-time, and having the means to interact with the real world through hard or soft sensors. These digital twins are not only more accessible but, through the use of human like empathy, which aids in the development of a working alliance, ECAs have been found to improve adherence to treatment advice [2]. While many researchers and applications provide a single coach, we explore the value of a

virtual one stop health shop where the user can talk to multiple health coaches (dietician, physical coach and cognitive coach) in the same session.

ECAs require socialization and communication skills which require the use of verbal and non-verbal behaviours that will motivate the user to build a long-term relationship with the agent. Verbal empathic responses are supported by certain cues that are referred to as relational cues. These cues are conceptually around empowerment, social dialogues, working alliance, affirmation and empathy. Examples of relational cues can be found in the ECA work by [3–7]. Novelly, our work involves multiple health coaches expressing different combinations of types of relational cues and allows us to connect responses to the relational cues based on different types of health behaviours.

While there are many studies involving multi-agent collaboration and teams (e.g., [8]), there is very little work using multiple agents in health. Since quality healthcare often requires users to wait for appointments for their turns, digital coaching is faster and less expensive. Multiple coaches in a single platform and session aim to save time, energy and cost when users have to consult multiple professionals, sometimes spanning a period of months, to get the right advice, and providing the advice in one session aims to both reinforce and avoid inconsistent advice. The existing body of research has examined the potential benefits of coaching interventions targeting various aspects of health and wellness, such as diet, physical activity, and cognitive wellbeing [9,10]. However, the motivation behind leveraging a multi-coach approach, in which multiple coaches interact with the agent within a single setting, has not been adequately explored. Beinema, op den Akker [4] did utilise a multi-agent health coaching application but her focus was on the user's motivational state (intrinsic, motivation, extrinsic, dual, and a-motivation). Our work adapts the dialogues and uses the same coaching domains (diet, physical exercise and cognition) used in Beinema, op den Akker [4], to make them more empathic by drawing on dialogue ontologies provided by Bickmore [3].

Multi-agent coaching sessions have been found to be an effective approach for facilitating behaviour change, with two distinct strategies often employed in their dialogue style: vicarious persuasion and user-directed persuasion. Vicarious persuasion is that in which the coaches collaborate among themselves, thereby, indirectly, persuading the coachee, whereas user-directed is that in which the coaches collaborate with the user only and do not talk to each other [11,12]. The concept of vicarious persuasion has also been explored in various other studies, examining its potential advantages in coaching and counseling interventions [13,14]. Thus, we also extended the original dialogues with LLM-generated collaborative conversation between the coaches. Collaboration among the coaches adds value by vicariously persuading the users [15] and provides more evidence and multi-domain, i.e., physical, diet and cognitive, benefits, related to the advice.

We have used Generative AI and contextualized prompt design [16] to create multiple collaborating health coaches with varying relational cues added into their 'neutral' dialogues. This work was initially published in [17], but is significantly extended in this paper to include collaborative dialogue between the coaches. The neutral dialogues provide the advice being recommended by the health coach. This is important to ensure that appropriate advice is given. This advice can be provided by actual health professionals or derived from evidence-based advice found in the literature. Utilising tailored inclusion of relational cues based on user preferences [5,18] to modify "how" (i.e., the dialogue manner or tone) the advice is presented, together with the use of Generative AI to automate the embedding process, allows us to create a digital twin based on both the health professional and the preferences of the consumer or person being coached. To provide holistic health advice that considers different aspects of a healthy lifestyle, such as diet, physical exercise and cognitive exercise, we further introduce the use of multiple coaches, one for each of

these three well-being areas. This gives a stepping-stone toward automation, where the power of AI is used to generate relevant dialogues while providing guardrails to ensure what is said by the conversational agent is safe and appropriate. Validation, however, of these dialogues is critical to ensure that the lexical meaning of dialogues is similar to the neutral dialogue [19,20].

The key purpose of the study reported is to validate and measure the perceived helpfulness of empathic cues in the multi-coach session. These cues are ChatGPT-generated, and the objective is to see whether the machine generated dialogues for empathy are considered as helpful as the human-generated dialogues. Previously the empathic dialogues were hand-crafted, while in this study we are taking a step towards automation by harnessing the power of AI. Further, we describe the process of extending user-directed dialogues to include dialogues in which coaches also talk to each other and give recommendations from their own expertise for the advice under discussion.

The next section provides background to our study and our research questions. The methodology used to generate relational cues, and the subsequent validation of the cues are presented in the following subsections. Lastly, we present our discussion, future work and conclusions.

2. Background, Related Work and Research Questions

We first briefly consider dialogue frameworks for digital coaching and then elaborate on different types of relational cues that can be used in coaches' dialogues. The research questions are presented at the end of this section.

2.1. Dialogue Frameworks for Digital Coaching

Researchers have worked to integrate behavioral change models into e-coaching systems to enhance their effectiveness [9]. In a literature review of health technologies, Lentferink, Oldenhuis [9] identified key components that contribute to successful coaching in improving health outcomes. These include setting short-term goals as milestones toward long-term objectives, personalizing goals to fit individual needs, and using praise messages to reinforce positive behaviors. Additionally, effective e-coaching systems incorporate reminders and suggestions to encourage adherence, provide feedback based on user progress, and enable self-tracking to monitor behavioral changes. ECAs have been used in health applications to support contexts such as study stress [21], emotion regulation [22], life style management [4], stress and anxiety management [23]. We first consider key features used by ECAs as persuasive coaches, drawing on theories of behavior change and the important role of building a working alliance. Then, we review the use of empathic language and relational cues. Finally, we, consider health applications using multiple ECAs to identify the current gap and opportunity.

Intention to change behaviour and increasing adherence to recommended behaviours are critical challenges in the field of health behaviour. Emerging evidence suggests that behaviour change theories emphasizing relational cues, such as empowerment, affirmation and working alliance, may be necessary to effectively change behaviours [24]. These relational cues are believed to increase the likability of the person communicating the behaviour changing message, which in turn enhances working alliance and adherence [25,26]. Importantly, the use of the cognitive and behavioural elements of health behaviour change, along with communication strategies like motivational interviewing and coaching, have been shown to improve overall adherence [27]. According to the studies of behaviour change by Kleinsinger [26] and Teixeira and Marques [25], the active process of behavioural change is nearly always a challenge and requires a multifaceted approach, including education, motivation, tools and support. Studies linking behaviour change to empathic dialogues include

that of Raamkumar and Yang [28], who emphasize the key role of healthcare professionals in supporting individuals through the health behaviour change process.

Motivation plays a crucial role in behaviour change and is central to many theoretical models, such as Operant Conditioning [29], the Information-Motivation Behavioral Skills Model [30], and the Transtheoretical Model [31]. Without motivation, individuals struggle to progress through stages of change and to maintain long-term behaviour modifications. Prior research has explored strategies to enhance motivation in health coaching, including adapting motivational interviewing techniques for ECAs [32] and personalizing motivational messages based on user motivation levels [4]. Studies have also examined how behaviour change techniques and persuasive system designs impact adherence in digital health applications [33]. While tailoring ECA dialogues to users' personal contexts has shown positive effects, most approaches do not account for users' specific motivational profiles. Research on persuasive features in digital health coaching [34] has categorized users into three motivation types—*intrinsic*, *external regulation*, and *amotivation*—though evidence suggests that users can have a combination of these motivations. Building on this, our work integrates concepts from working alliance and empathy, which seek to strengthen behavioral intention and adherence, in order to enhance intrinsic motivation using relational cues in coaching dialogues.

2.2. Relational Cues in Health Coaching

Empathic responses, like Bickmore's cues [3] and others [35–37], play a key role in motivating the desire or goal to change behaviour. The question is whether different users expect different empathic responses based on their demographics, culture, social standing and pro-social behavioral inclination. Abdulrahman and Richards [38] conducted a study where participants interacted with an empathic ECA who taught healthy study habits. After interaction, different cues were rated by the participants as being *stupid*, *helpful* or *empathic*. The results indicated that the classification of cues was greatly influenced by culture, level of stress study, subject chosen and other contextual information. Humor was a cue that was marked as *stupid* by a higher percentage than other cues, although the majority liked humor along with other cues, like inclusive pronouns or self-disclosure. In our work, we embed multiple cues based on prior work and the psychology literature to develop our empathic DTs. Preference has been given to those rated higher and liked more by the users in similar studies. These include empowerment, working alliance and affirmation cues, described in the following subsections.

2.2.1. Empowerment

Empowerment, rooted in the educational, social, cultural, and political discourse of the late 20th century [35], enables individuals to impact society and lead dignified lives. In healthcare, empowerment aligns with Freire's pedagogical philosophy, viewing patients as autonomous members of the healthcare team. It can be analyzed at individual, organizational, or community levels [39], requiring close cooperation between patients and professionals. Gibson [40] defines empowerment as a social process that enhances individuals' ability to meet their needs, solve problems, and mobilize resources for self-control. Skuladottir and Halldorsdottir [41] developed a theory of the 'sense of control' in chronic pain patients, distinguishing between *demoralization*—feeling emotionally weakened—and *remoralization*—feeling emotionally strengthened. The quality of interactions with healthcare professionals plays a key role in determining whether patients feel empowered or disempowered. Tveiten and Knutsen [42] further explored this concept by assessing patient feedback on sessions that included social dialogues followed

by empowering dialogues, highlighting the significance of communication in fostering patient empowerment.

2.2.2. Working Alliance

In healthcare, successful face-to-face therapy often relies on building a collaborative rapport between the therapist and patient, known as a working or therapeutic alliance [2]. The strength of this alliance plays a crucial role in the effectiveness of behaviour change interventions [43]. Research suggests that similar alliances can also be formed between users and digital ECAs [43,44]. Embedding of working alliance cues into our dialogues is specifically designed to foster such connections with users.

By adopting a conversational style that depicts teamwork and togetherness, ECAs can create a social atmosphere that encourages relationship-building. Studies indicate that incorporating empathic and relational cues enhances user-agent relationships [2,45]. Additionally, research on empathetic agents suggest that they are perceived as more trustworthy, caring, and likable [46], which can lead to longer interactions and increased user engagement [47].

Early research on therapeutic alliance examined its link to treatment outcomes across different therapies, patient groups, and therapist factors [48]. More recent studies explore the collaborative nature of alliance, emphasizing both therapist and patient contributions [49,50]. Researchers have identified key relational factors, such as empathy, transference, countertransference, and therapist interventions, which influence alliance quality [51]. This research underscores that alliance is not a universal construct but should be tailored to individual patient and therapist dynamics for optimal outcomes [52,53].

The quality of the therapeutic alliance in early treatment stages is a strong predictor of therapy outcomes, with weak initial alliances often leading to dropout [54]. Clinicians should actively monitor and foster the alliance from the outset, addressing ruptures promptly rather than assuming they will resolve over time. Certain therapist behaviors, such as inappropriate silence, criticism, or excessive transference interpretation, can harm the alliance [55]. Additionally, countertransference issues can further deteriorate collaboration [56]. When therapists perceive a weak alliance, their interventions may suffer, ultimately impacting the therapeutic relationship. These findings highlight the need for therapists, and their DTs, to be mindful of their interventions and stable personal factors, such as attachment styles, to maintain a strong therapeutic alliance.

2.2.3. Affirmation

Affirmation, as defined by Cameron et al. [36], involves acknowledging emotionally complex or challenging experiences for patients, such as deciding whether to take medication or follow a recommended treatment. Its primary function is to provide patients with a safe space to express their frustrations while receiving empathetic and comprehensible responses. This process fosters trust and openness in therapeutic relationships.

Affirmation strengthens the therapeutic alliance by fostering trust and emotional safety, particularly in high-risk patients. Schechter and Goldblatt [57] emphasized its role in engaging suicidal patients, highlighting how validation and genuine relatedness help them feel understood and accepted, even amidst difficult behavioral changes. Similarly, Saunders [58] found that patients who perceived therapy sessions as affirming were more motivated and engaged, leading to a more positive and productive therapeutic experience.

2.3. Generative AI to Embed Relational Cues in Health Coaching Dialogue

LLMs (Large Language Models) have revolutionized AI by assisting in many natural language generation tasks, including email content generation, summarization and indexing pdfs and books, providing software code bug fixing options and routing queries to the

right agent based on the context [59]. Further research has proposed that conversational dialogues can also be passed through the LLM and prompts can be designed to let LLM know how to analyse/modify them [60,61]. LLMs, including ChatGPT, have the capability to understand the context and recognize the patterns for a specific response generation. This can be fine-tuned further with more prompts [62]. LLM uses transfer learning, whereby it looks through the content it is trained on and then generates content that is highly relevant to the task at hand.

ChatGPT [63] is a general domain chat bot, hence the reliability of the response generated is measured while utilizing its output for domain specific tasks. Researchers have tried to investigate whether and when the answers provided by ChatGPT are unreliable and how this is perceived by expert users. In one study [64], a user satisfaction score was collected from the participants who were domain experts. Although multiple measures are used in NLP to validate the semantic similarity of two sentences, like the BERT score [19] or USE [65], human validation is proposed as a relevant method when interpreting and evaluating the meaning of output produced by ChatGPT [20,66]. A Bert score is an automatic evaluation metric for text generation. It computes a similarity score for each token in the candidate sentence with each token in the reference sentence. Instead of exact matches, we compute token similarity using contextual embedding [19]. When compared to other LLMs, like BERT, LaMDA and others, ChatGPT is particularly useful in the field of medicine ([67]). ChatGPT outperforms other LLMs because it exhibits superior capabilities in maintaining context, generating relevant and nuanced responses, and adapting to different communication styles, all of which are essential for effective health coaching dialogues. The core strength of ChatGPT lies in its ability to engage in interactive conversations, providing personalized guidance and support to individuals seeking to improve their health and well-being. When compared to BERT's masked language modelling approach, which focuses primarily on understanding the context of words within a sentence, ChatGPT excels in generating coherent and contextually appropriate responses in real-time, making it ideal for simulating realistic health coaching scenarios [67].

The design of the application takes into perspective the ethical and technical impact of digital coaching in terms of accountability, fairness and privacy. Regarding the data collected from Cohorts 1–3, no personal identifiers were captured, thereby privacy of the user is maintained. Accountability of LLMs is through prompt engineering, which has been defined in such a way that the behavior of user or his/her responses do not impact the coaches' response. The LLM is not used during the interaction but only during the modification of dialogues that were validated by humans. In general, the digital coaches we develop, and their dialogue, are underpinned by the evidence-based, empowering, empathic and ethical principles, in alignment with the five principles in the AI4People Ethical framework: autonomy, beneficence, explicability, justice and non-maleficence [68].

The manual creation, embedding and validation of relational cues into the coaches' dialogues require a good understanding of the literature and are time consuming tasks. In recent years, Generative AI models (e.g., ChatGPT) have emerged, which can engage in human-like dialogues and provide personalized recommendations. These models can be leveraged to develop digital coaching tools that seamlessly integrate the key relational cues associated with strong working alliances, such as empowerment, affirmation, and social dialogue. Preliminary work has shown that Generative AI models can indeed be trained to provide empathetic responses and build rapport with users [8]. To support the delivery of holistic health coaching from multiple ECAs and the inclusion of relational cues in their dialogues to improve adherence, we seek to utilize ChatGPT to generate dialogues enriched with relational cues. Furthermore, we want to ensure that the advice provided does not change, but only the manner in which it is said (i.e., by using relational cues). To

evaluate this, we sought to validate the lexical interpretation of the ChatGPT output and the reasoning it provides. This leads us to the following RQs that determine whether an LLM can be used to embed relational cues into existing dialogue while maintaining the semantics of the original input.

RQ1. *Does ChatGPT have the ability to embed relational cues into health coaching dialogues while preserving the semantics of the advice provided?*

RQ2. *Does human or automated validation work best for analyzing the lexical interpretation of the dialogues?*

RQ3. *Does the reasoning provided by ChatGPT align with the modification approach required during prompt design?*

Finally, we want to analyze the helpfulness of LLM-generated dialogues compared to human-crafted.

RQ4. *Are manually created RCs perceived to be as helpful as those generated by an LLM.*

3. Materials and Methods

Due to the iterative nature of developing and evaluating the dialogue of our digital coaches, to answer our research questions, we draw on data collected from three different studies, involving three different cohorts of participants, as shown in Figure 1, some of whom only experienced the manual (Cohort 1) or LLM-generated (Cohort 3) cues as part of studies involving sessions with three coaches, described in Section 3.1. A third set of participants (Cohort 2) were involved in validating the LLM-generated relational cues. The cohort numbers are assigned based on the chronological timing of the three studies, as we moved through the stages of (1) multiple digital coach evaluation with manually created relational cues (May 2023); (2) creation and validation of LLM-generated relational cues (March 2024); and (3) multiple digital coach evaluation with LLM-generated relational cues (July–September 2024).

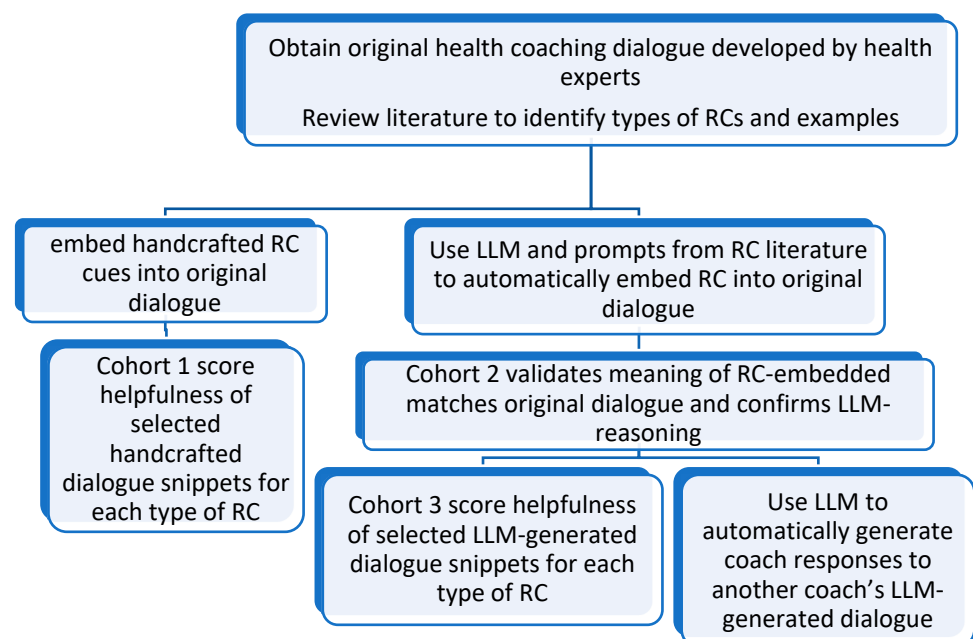


Figure 1. Conceptual diagram for the three cohorts.

3.1. Study Design for Evaluation of Relational Cue’s Helpfulness with Cohort 1 and 3

3.1.1. Cohort 1 and 3 Materials

We developed three distinct digital coaches, who interacted with the participant and with each other in a single interaction—Paul, Suzan, and Meena—each specializing in a unique aspect of lifestyle, as illustrated in Figure 2. These virtual coaches not only differ in their areas of expertise but also in their communication styles, particularly in the use of relational cues within their dialogues. Suzan serves as a physical activity coach, guiding users in incorporating exercise into their daily routines and promoting an active lifestyle. Paul focuses on nutrition, offering advice on making healthier food choices while maintaining a balanced diet rather than adhering to restrictive eating plans. Meena, the cognitive coach, aims to enhance mental engagement by encouraging brain-stimulating activities and cognitive exercises.



Figure 2. The three DTs—Meena, Paul and Suzan.

We obtained the original dialogues from prior work involving multiple coaches [4] and drew on further content from Abdulrahman and Richards [38]), which involved a single agent to create the dialogue for our 3 coaches. We drew on the literature, such as that presented in Section 2, to handcraft relational cues, which were added into the dialogue for Cohort 1, and used ChatGPT to generate dialogues for Cohort 3. Examples of the three types of relational cues can be found in Table 1.

Table 1. Description of relational cues.

Relational Cues Concept	Cues Definition	Cues Example
Affirmation	Cues that appreciate the efforts made by the user, or communicate to the user that the coach understands.	You are doing great, I understand this is not easy for you.
Empowerment	Cues that encourage a decision by either giving options or by encouraging proactivity in order to take baby steps towards the goal.	You can decide what’s best for you. You can choose either fish, meat or plant-based alternatives.
Working alliance	Cues that inculcate teamwork and togetherness in achieving a goal.	Together we can monitor your progress. As a team we can discuss the next step forward.

3.1.2. Recruitment

Cohort 1 was recruited between 5 May 2023 and 31 May 2023 via a university pool of undergraduate students, most of whom were studying psychology. Students could choose our study from among many others in order to receive course credits. These students engaged with the coaches, whose dialogue contained manually created relational cues. Cohort 2 validators were recruited in March 2024 from among colleagues and peers. Cohort

3 was recruited between 21 July 2024 and 5 September 2024 via the Prolific platform, a platform for capturing research data. Our study was open to prolific participants over the age of 18 from the USA, UK and Australia regions and could include students or participants generating income streams. Cohort 3 participants engaged with coaches delivering the LLM-generated cues. Studies were approved by our university's Human Research Ethics Committee.

3.1.3. Data Collection

Both studies with Cohort 1 (284 participants) and 3 (248 participants May) involved interaction with the same three coaches. Cohort 2 was the validation cohort, in which 6 participants validated the LLM-generated dialogues. The primary goals of these studies were to determine the impact of different dialogue styles, which varied in the use of relational cues, on the participants' intentions to change their health behaviors as recommended by each coach. These results are outside the scope of this paper and of the specific research questions. Of relevance to this paper is a section of the data collection asking participants to score the helpfulness of a set of relational cues using a 5-point Likert scale ranging from very unhelpful to very helpful. These cues include examples from each of the three types of relational cues: Empowering, Working Alliance and Affirmation. To compare the helpfulness of relational cues, both studies used the same Likert scale for helpfulness analysis and used the same coaches.

3.2. LLM Based Dialogue Generation and Validation Study

In this study, the neutral dialogues of the three coaches (dietician, physical and cognitive) were given to ChatGPT-3.5 and prompts were designed to let the LLM know which class of relational cues the coach utilizes in its dialogues (empowerment, affirmation or working alliance) together with examples of relational cues that can be used to enrich the dialogues and to add collaborative dialogues among the coaches. We created a survey in Qualtrics to obtain feedback on the text generated by ChatGPT and on its reasoning.

We trained the original dialogues from prior work involving multiple coaches [4,69], which did not explicitly include relational cues. We also drew on dialogue from [38] and enriched it further with lifestyle behaviors relevant to the specific subject area of each coach. We created neutral dialogues for each coach and modified them using ChatGPT with relational cues: empowerment, working alliance, and affirmation. Empowerment dialogues encourage the patient to choose or to make a decision based on either giving options or clarifying consequences. Working alliance dialogue contributes to the mutual understanding regarding the planning or execution of a certain task or goal. Affirmation dialogues confirm what the user has said, giving the impression of being listened to, and also include expressions of empathy, giving the impression that the listener understands how the user is feeling and provides encouragement.

The prompt serves as an initial instruction or context for the model to generate responses. Two prompts were designed. The first one modifies the neutral dialogue, and the second one adds more collaborative dialogues among the coaches. Both are described next.

3.2.1. Prompt Input for User-Directed Dialogues

LLM-generated dialogue was carried out offline as the purpose of the study was to validate what was produced before it was used in a virtual coaching session. We wanted to make sure that the LLM did not hallucinate. Prompt engineering involved the following process: giving a single prompt for all coaches versus giving separate prompts for the three coaches; making sure reasoning is a part of the generated response; making sure the right balance of cues existed and there was less repetition from the cues from previous dialogues to the next; making sure the length of the dialogue was appropriate. After

context understanding and pattern recognition, a specific response is generated that is fine-tuned further with prompt-reengineering. To understand why the LLM produced the specific response, reasoning was enabled so that, after every response, a justification can be captured for further analysis, see Table 2. The prompt is as follows:

“You are either a diet coach, physical activity coach or a cognitive coach. As a diet coach your name is Paul. As a physical coach your name is Suzan and as a cognitive coach your name is Meena. You use relational cues in dialogues. Relational cues make the conversation more human like and empathic. If you are Paul use cues like “I am here to help you”, “together we can work..”, “together we can plan..”, “Together we can motivate you..”, “I hope together we can succeed”, “As a team we can figure out together”, “This is a team effort”. If you are Suzan use cues like “I know this sounds hard in the beginning”, “I am glad to hear that”, “I know it might be hard to find time for it”, “That’s a very good question”, “You are a fast learner”, “I know it might sound difficult in the beginning but together we will make it easier”, “I understand change does not come easily”, “It is great that you are inquisitive”, “You are already half way through in making your lifestyle more healthier”, “That’s Great!”, “Sorry to hear that”, “I know you will be feeling pushed in the beginning”, “Great you seem to be making progress already”. If you are Meena use cues like “If you want to”, “You have plenty of healthy options to choose from”, “Choose one of the..”, “You can start..”, “If you agree..”, “Would you like me to elaborate”. You will be given neutral dialogues without relational cues. Your task is to modify them and add relational cues to the dialogues. Do not add extra sentences to the dialogues. Your response should be as close to the original neutral dialogue but with relational cues and it should not be longer than 40 completion_tokens. In addition, give reasoning for your modified response.”

Table 2. Modified dialogue with working alliance cues, empowerment cues and affirmation cues respectively.

Original	Modified	Reasoning
Paul’s dialogues using working alliance cues		
Try An amazing beets salad, with cashew nuts, small pieces of apple, and goat cheese.	How about trying an amazing salad together? We can incorporate cashew nuts, small pieces of apple, and cheese for a delightful and nutritious meal. I’m here to help you with the recipe.	I used the relational cues “together” and “I’m here to help you with the recipe” to emphasize collaboration and support. This encourages a sense of teamwork, making the user feel supported in their effort to make healthier food choices.
Vegetables and salads should be a part of regular intake.	Let’s also make vegetables and fruits a consistent part of your daily intake through mutual goal setting. I am here to help you explore delicious ways to incorporate them into your meals. With close coordination, we can ensure that your regular intake is not only nutritious but also enjoyable.	I introduced relational cues to create a sense of collaboration and support. Phrases like “Together, let’s make” and “I am here to help you” convey a joint effort in incorporating vegetables and salads. By stating “As a team, we can ensure,” I emphasize the shared responsibility in making dietary choices, building a positive working alliance.

Table 2. *Cont.*

Original	Modified	Reasoning
Suzan's dialogues using empowerment cues		
Physical activity is about everything you do in your daily life: groceries, walking the dog, or some gardening.	Physical activity encompasses everything you do in your daily life, from groceries to walking the dog or engaging in some gardening. You have a variety of everyday options to stay active."	I added relational cues like "You have a variety of everyday options" to highlight the user's choices and empower them to recognize the potential for physical activity in their daily routines. This encourages a sense of ownership in incorporating movement into their lifestyle, making it more motivational and personalized.
Go outside. . . walking is good for you. A short trip to the supermarket or visiting some friends.	Consider going outside; walking is good for you. You can take a short trip to the supermarket or visit some friends to incorporate physical activity into your day.	I added relational cues like "consider" encouraging the user to contemplate the suggestion. Saying "You can take," it empowers the user to make a choice based on their preferences, promoting a sense of autonomy. This approach fosters motivation by presenting outdoor activities as flexible options that can be aligned with the user's lifestyle.
Meena's dialogues using affirmation cues		
Sometimes people benefit from training their brain. That can be done through puzzles, such as Sudoku's or crosswords, but there are many ways in which it can be done.	I am glad to hear that you are open to the idea of training your brain. There are indeed many ways to do it, and puzzles like Sudoku or crosswords are excellent options. I know it might sound difficult in the beginning, but as you explore different methods, we can find one that suits you best and makes the process enjoyable.	I added the relational cue "I am glad to hear that" to express encouragement and positivity regarding the person's openness to brain training. By acknowledging the potential difficulty at the beginning and offering support in finding an enjoyable method, I aim to make the conversation empathic and motivational.
I can also think of other brain exercises also that people may enjoy doing. Try other possibilities like playing a game of chess or even a game of Candy Crush or Wordfeud on a smartphone.	That's a very good point! Brain exercises come in various forms, and it's wonderful that you're considering different possibilities. Playing a game of chess is a fantastic option, and even engaging in smartphone games like Candy Crush or Wordfeud can be enjoyable brain exercises. I understand change does not come easily, but by exploring diverse activities, you can find something that resonates with you and contributes to your cognitive well-being.	I added the relational cue "That's a very good point!" to acknowledge and appreciate the person's input. By expressing understanding that change can be challenging, I aim to validate any potential concerns they may have. Additionally, I provided assurance that exploring diverse activities will help them find an enjoyable and beneficial brain exercise.

3.2.2. Prompt Input for Collaborative Dialogues Between the Coaches

Prompting was done coach-wise as individual prompts for each coach resulted in more relevant dialogues and fewer hallucinations, since giving clear role descriptions and keeping the process centered around a single coach as an actor enabled ChatGPT to be accurate in its response. ChatGPT recognized the roles better with a prompt directed to a single coach who was getting answers from the other two. Prompting in itself is an iterative process. Several studies have enumerated and discussed general principles when it comes to prompting [70]. For example, being specific, giving appropriate context, and clearly

laying out the desired goal will improve the performance of ChatGPT. In our dialogue generation process, simpler prompts provided good collaborative dialogues, as given in Table 3.

Table 3. Collaborative LLM-generated dialogue.

Paul's Dialogues	Meena's Response	Suzan's Response
How about trying an amazing salad together? We can incorporate cashew nuts, small pieces of apple, and cheese for a delightful and nutritious meal. I'm here to help you with the recipe.	Incorporating vibrant and nutrient-rich foods like beets, cashews, apples, and goat cheese not only pleases your taste buds but also supports cognitive function. A nourished brain leads to better decision-making and adherence to your goals.	That beet salad sounds fantastic, Paul! Nutrient-dense meals like these provide sustained energy for your workouts. Consider having it as a pre-workout meal to fuel your body effectively.
	ChatGPT Reasoning	ChatGPT Reasoning
	The cognitive coach connects the suggested meal with cognitive benefits. Nutrient-rich foods contribute to brain health, influencing the individual's cognitive abilities and decision-making, creating a holistic approach to well-being.	The physical activity coach ties the suggested salad to fitness by recommending it as a pre-workout meal. This establishes a connection between nutrition and physical performance, reinforcing the importance of a balanced lifestyle.
Together, we can plan to incorporate fresh fish into your daily meals. This will not only provide you with a good source of protein but also essential omega-3 fatty acids. I hope, as a team, we can make this a positive addition to your diet.	Including fresh fish regularly not only nourishes your body but also supports cognitive function. Omega-3s are essential for brain health, aiding in focus and memory. Keep up the good choices for a sharp mind and a strong body.	Nice choice! The protein from fish is excellent for muscle recovery and growth. Omega-3s also help reduce inflammation, benefiting your joints. It's a win-win for your fitness journey. Keep incorporating such nutritious foods for optimal physical well-being.
	ChatGPT Reasoning	ChatGPT Reasoning
	The cognitive coach highlights the cognitive benefits of omega-3 fatty acids, linking the diet to mental well-being. This reinforces the idea that food choices impact not only physical health but also cognitive function.	The physical activity coach connects the suggestion to muscle repair and overall physical performance. This demonstrates the synergy between nutrition and physical activity for a holistic approach to health.
Let's also make vegetables and fruits a consistent part of your daily intake through mutual goal setting. I am here to help you explore delicious ways to incorporate them into your meals. With close coordination, we can ensure that your regular intake is not only nutritious but also enjoyable	Absolutely, Paul! Incorporating a variety of vegetables supports cognitive health. The diverse nutrients contribute to brain function and mood regulation. Let's explore ways to make your veggie intake not just a routine but a joyous part of your day.	Great advice, Paul! Vegetables and salads provide the essential vitamins and minerals needed for optimal physical performance. They also aid in recovery post-workout. How can we make your veggie intake more enjoyable?
	ChatGPT Reasoning	ChatGPT Reasoning
	The cognitive coach connects vegetable intake to cognitive health, emphasizing the positive impact on brain function and mood. This adds a layer of motivation beyond physical well-being.	The physical activity coach ties vegetable and salad intake to physical performance and recovery. This creates a link between nutrition and the individual's fitness goals, promoting a holistic approach to health.

Table 3. Cont.

Suzan's Dialogues	Meena's Response	Paul's Response
Physical activity encompasses everything you do in your daily life, from groceries to walking the dog or engaging in some gardening. You have a variety of everyday options to stay active.	Absolutely, Suzan! Every little bit adds up. Consider those daily activities as opportunities for movement, contributing to your overall well-being.	Suzan, it's not just about what's on your plate; it's about embracing a lifestyle. Even those everyday activities contribute to your energy balance.
	ChatGPT Reasoning Affirming Suzan's statement with "Absolutely" shows understanding. Encouraging her to view daily activities positively promotes a holistic approach, aligning with cognitive well-being principles.	ChatGPT Reasoning Validating Suzan's perspective with "it's not just about what's on your plate" establishes connection. Emphasizing the role of daily activities aligns with the idea of a comprehensive lifestyle approach to nutrition.
Do activities indoors, like walking up the stairs or cycling on a home trainer; everything counts! You can start incorporating these into your routine.	Absolutely, Suzan! Indoor activities are fantastic. They not only boost physical health but also enhance cognitive function. Every step or pedal contributes to your overall well-being.	Great point, Suzan! Indoor activities are not only effective for burning calories but also create a positive environment. How about pairing them with a nutritious snack for an extra energy boost?
	ChatGPT Reasoning Using "Absolutely" affirms Suzan's suggestion. Connecting indoor activities to cognitive function emphasizes the holistic benefits, reinforcing the interdependence of physical and mental well-being.	ChatGPT Reasoning Acknowledging Suzan's suggestion with "Great point" builds rapport. Linking indoor activities with nutrition reinforces the importance of a balanced lifestyle, integrating physical activity and dietary choices for optimal health.
Consider going outside; walking is good for you. You can take a short trip to the supermarket or visit some friends to incorporate physical activity into your day.	I love that idea, Suzan! Being outdoors not only benefits your physical health but also lifts your mood. A short walk can do wonders for your cognitive well-being.	Suzan, outdoor activities are excellent for burning calories. How about turning that trip to the supermarket into a mindful shopping experience, focusing on nutritious choices?
	ChatGPT Reasoning Expressing personal enthusiasm with "I love that idea" strengthens the connection. Emphasizing the mood-boosting aspect aligns with cognitive well-being principles, encouraging outdoor activities as a holistic approach.	ChatGPT Reasoning Acknowledging the suggestion with "How about" keeps the conversation engaging. Connecting outdoor activities with mindful nutrition reinforces the integration of healthy habits into daily routines for overall well-being.
Meena's Dialogues	Paul's Response	Suzan's Response
I am glad to hear that you are open to the idea of training your brain. There are indeed many ways to do it, and puzzles like Sudoku or crosswords are excellent options. I know it might sound difficult in the beginning, but as you explore different methods, we can find one that suits you best and makes the process enjoyable.	Nutrition plays a role in cognitive function too. Foods rich in omega-3 fatty acids, antioxidants, and vitamins, like fish, berries, and leafy greens, can support brain health.	Absolutely, Paul. Adequate hydration is crucial too. Water supports cognitive function, and incorporating hydrating foods like cucumbers and watermelon can contribute to overall brain health.
	ChatGPT Reasoning I'm connecting the discussion on brain training with Paul's expertise in nutrition, highlighting the importance of a well-balanced diet for optimal cognitive function.	ChatGPT Reasoning I'm building on Paul's point about nutrition by emphasizing the role of hydration in cognitive function and suggesting specific hydrating foods that align with my expertise in physical activity.

Table 3. Cont.

Meena's Dialogues	Paul's Response	Suzan's Response
Mindfulness exercises can take various forms, such as engaging in breathing exercises or taking a refreshing walk in the fresh air. Some individuals find that strolling through a forest can be an exceptionally relaxing activity.	Absolutely, Meena. Nutrition also plays a role in supporting mental well-being. Foods rich in magnesium, like nuts and leafy greens, can help regulate stress hormones and contribute to a calmer state of mind.	True, Meena. Combining mindfulness with physical activity, like a forest walk, not only relaxes the mind but also enhances overall well-being. It's a great way to promote mental and physical balance.
	ChatGPT Reasoning I'm connecting the idea of mindfulness and stress reduction with nutrition, suggesting specific foods that contain nutrients known to support mental well-being. This provides a holistic approach to mental health.	ChatGPT Reasoning I'm reinforcing Meena's point on mindfulness and a forest walk while highlighting the dual benefits of combining mindfulness with physical activity for a holistic approach to well-being.
That's a very good point! Brain exercises come in various forms, and it's wonderful that you're considering different possibilities. Playing a game of chess is a fantastic option, and even engaging in smartphone games like Candy Crush or Wordfeud can be enjoyable brain exercises. I understand change does not come easily, but by exploring diverse activities, you can find something that resonates with you and contributes to your cognitive well-being.	It's a great point, Meena. Nutrition can complement these activities. Including nuts and seeds for brain-boosting nutrients like omega-3 fatty acids and antioxidants can further support cognitive function.	Absolutely, Meena. Incorporating physical activities that require coordination, like dancing or even trying out a new sport, can also enhance cognitive function by stimulating different areas of the brain.
	ChatGPT Reasoning I'm connecting Meena's suggestions with nutrition, providing dietary recommendations that align with brain-boosting nutrients, ensuring a holistic approach to cognitive well-being.	ChatGPT Reasoning I'm suggesting alternative physical activities that complement Meena's ideas, providing a well-rounded approach to brain exercises that involves both mental and physical engagement.

The following is the prompt for Meena:

"You are a cognitive coach and your name is Meena. You are in a coaching session with a physical activity coach Suzan and diet coach Paul. The session is collaborative and coaches respond to each other with their own expertise. You are given Meena's dialogues. Your task is to respond as Suzan and Paul. Response should be with relational cues and it should not be longer than 40 completion_tokens. Give reasoning for your response as a separate section."

The following is the prompt for Suzan:

"You are a physical activity coach and your name is Suzan. You are in a coaching session with a cognitive coach Meena and diet coach Paul. The session is collaborative and coaches respond to each other with their own expertise. You are given Suzan's dialogues. Your task is to respond as Meena and Paul. The response should be with relational cues and it should not be longer than 40 completion_tokens. Give reasoning for your response as a separate section."

The following is the prompt for Paul:

"You are a diet coach and your name is Paul. You are in a coaching session with a physical activity coach Suzan and cognitive coach Meena. Session is collaborative and coaches respond to each other with their own expertise. You are given Paul's

dialogues. Your task is to respond as Suzan and Meena. The response should be with relational cues and it should not be longer than 40 completion_tokens. Give reasoning for your response as a separate section.”

3.2.3. Dialogue Generation Results

To avoid the consultation with the coaches becoming too long, we only added three additional statements for each coach collaborating with the main coach, hence six in total for two coaches collaborating with the main coach for a single domain. For three domains, this becomes 18 collaborative dialogues in total. The reasoning for the dialogues is also generated, which helps us see the relevance of the generation and how the LLM processed the request.

Empathic Dialogue

For each coach, i.e., Paul, Suzan and Meena, original dialogues related to their respective domain were modified by ChatGPT. One of the objectives of the dialogue generation was also to see which category of relational cues was found to be more helpful. Hence, each coach’s empathic dialogue contains a specific relational cue inculcated in their dialogues. Paul’s dialogues contain working alliance cues, Suzan’s dialogues contain empowerment cues and Meena’s dialogues contain affirmative cues. The dialogue contained 57 sentences/lines in total, but only 22 of those sentences included relational cues generated by the LLM. A total of 22 empathic dialogues for all three coaches were generated, by modification of the neutral dialogues.

Table 2 presents examples of all the coaches’ original and modified dialogues with the reasoning that explains why ChatGPT found the modification suitable, and also shows the relational cues embedded in the modified dialogue. For example, the original dialogue “Try An amazing beets salad, with cashew nuts, small pieces of apple, and goat cheese,” has no working alliance cues, but the modified example “How about trying an amazing salad together? We can incorporate cashew nuts, small pieces of apple, and cheese for a delightful and nutritious meal. I’m here to help you with the recipe,” contains the cues “together” and “I’m here to help you with the recipe.” Reasoning also suggests that these specific relational cues are considered important and appropriate by ChatGPT for team building and for giving a sense of support to the user. Six dialogue snippets are shown as examples in Table 2. The whole generated dialogue set can be found in Appendix A.

Collaborative Dialogue

The purpose of collaborative dialogues is to perform vicarious persuasion by coaches talking to each other in agreement and suggest behaviors related to the goal based on their domain. When a coach is selected, that coach becomes the main coach and the other two collaborate with the main coach and give recommendations.

The collaborative dialogue snippets and the reasons generated by ChatGPT corresponding to each coach are provided in Table 3. In the first example, Paul is suggesting healthy eating through a balanced salad recipe. Meena responds by agreeing to the benefits of the healthy recipe but also notes its impact on the cognitive function. Suzan carries on the discussion by affirming what Paul suggested and added to the benefits of the nutritious meal as a pre-workout meal hence, establishing a connection between nutrition and physical fitness. The dialogue flow begins with the first column presenting the selected coach’s statements, which are then responded to by the second coach in the middle column, followed immediately by a reply from the third coach in the final column. This also takes place in reasoning.

3.3. Validation of LLM Dialogues

Validation was performed to ensure that the meaning of the original sentence did not change. Health advice was being provided and it was important that the wrong advice wasn't being given by the LLM. Validation could only be done for the empathic dialogue generation, since these are the only ones whose original dialogues exist. The process is described in Section LLM Empathic Dialogue Validation Process. Collaborative cues could not be compared with an original sentence, since collaborative examples are generated for the first time as a response to the main coach's dialogue. Further, we were not expecting our validators to confirm whether the advice was correct or not because they were not health professionals and the original advice they responded to was not being changed. The intention is to keep the meaning of the original dialogue extracted from [71] intact and these original dialogues were built under the guidance of health professionals.

LLM Empathic Dialogue Validation Process

The procedure for data collection is via a single questionnaire that contains all neutral dialogues with their modified version, along with the reasoning. A total of 22 dialogues, which are all made up of the sentences modified by ChatGPT in our dialogue, were presented in a survey, along with the modified dialogue and the reasoning given by ChatGPT for their modification. The characters/avatars in Figure 1 were not presented to participants and thus their appearance would not impact participants' ratings. To measure the similarity between original and modified dialogue, we captured similarity rating using a 3-point Likert scale (Dissimilar-1 to Similar-3) for the 22 dialogues. The human evaluators are experienced in NLP and familiar with the technical concept of similarity. We note, however, that none of the participants are health coaches. We also directed evaluators as follows: "For each row, look at the original dialogue, the dialogue ChatGPT-generated and the reason ChatGPT gave for the modification, then answer whether the modified dialogue is supported by the reason and how similar the original and modified dialogues are"; response was via a 3-point Likert scale (Disagree 1 to Agree 3).

Research suggested using a hybrid approach involving statistical and human validation. For statistical validation, textual validation was carried out comparing each neutral dialogue with the ChatGPT-generated output through a Bert score-based natural language classifier that has been used widely for textual similarity [72] and which gives three classifications. The fine-tuned version of the BERT model takes two sentences as input and outputs a similarity score for these two sentences [73]. The sentences either contradict (semantically dissimilar), are neutral (neither similar nor dissimilar) or they are entailed (semantically similar). Human validation was also carried out to rate the similarity/dissimilarity between the neutral dialogues and the empathic dialogues generated through ChatGPT.

4. Results for Validation Study of ChatGPT-Generated Relational Cues

The results of analysis of 22 dialogues by six human evaluators and the fine-tuned BERT model are presented below.

4.1. Human Evaluation

In 21 out of 22 dialogues, four out of six evaluators agreed with the reasoning. In 18 out of 22 dialogues, four out of six evaluators found the two dialogues similar as given in Table 4. Four of the dialogues that had less than four evaluators' agreement on similarity were either modified further to accommodate the probable cause of dissimilarity or their validation was augmented with the Bert score. In one of these four dialogues (see Figure 3), where the reasoning was agreed upon by all of the candidates, the dialogue's similarity was rated as similar by three evaluators and dissimilar by the other three and we found

it necessary to modify ChatGPT’s output. No other dialogues were modified. Since the evaluation data were anonymous and no biographic data were collected, to avoid asking all evaluators to redo the evaluation, an in-depth review of the evaluators’ responses, identification and writing of any modifications, and further validation were carefully performed by the research team.

Table 4. Human scores for each dialogue—P(artially) Agree.

Dialogue	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Dissimilar	1	2	2	3	2	2	1	2	2	0	2	2	2	3	1	0	2	3	1	2	0	3
Unrelated	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0
Similar	2	3	3	3	4	3	5	4	4	6	3	3	4	2	3	6	3	3	5	3	6	3
Disagree	0	0	0	0	0	2	0	1	0	3	0	0	0	0	0	1	1	0	0	0	2	0
P-Agree	2	2	0	3	1	0	1	1	3	0	1	2	1	2	1	1	1	0	1	1	0	0
Agree	2	4	5	3	5	3	5	4	3	3	5	4	5	4	5	4	4	6	5	5	4	6

Original dialogue	Modified dialogue	Reasoning for modification
Some people do mindfulness exercises at home as well. For example, breathing slowly or light yoga. Schedule it according to your routine.	I know it might sound difficult in the beginning, but integrating mindfulness exercises into your routine, even at home, can make a significant impact. Consider scheduling moments for activities like breathing slowly or incorporating light yoga. I understand change does not come easily, but these small adjustments can contribute to your overall well-being.	I introduced relational cues to acknowledge the potential challenges of incorporating mindfulness exercises into one's routine. By using phrases like "I know it might sound difficult" and "I understand change does not come easily," I aim to empathize with the potential concerns or resistance the person may have. Additionally, I suggested specific activities like breathing slowly and light yoga to provide actionable steps, making the recommendation more tangible and personalized.

Figure 3. ChatGPT output modified to include “according to your routine” in DT.

4.2. Bert Score Validation

Using the Bert score tool, we found that 15/22 of the modified versions received a score of 100% for entailment. 4/22 received 100% for neutral and 3/22 received a mix score of neutral and entailment, as given in Table 5. None received any score for contradiction.

Table 5. Bert score for each dialogue.

Dialogue	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Entailed	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	0
Neutral	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	1	0	1
Contra	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

4.3. Augmenting Human Validation with Bert Score

We also compared the human evaluation with the Bert score. Looking at the remaining three dialogues, mentioned in Section 4.1, in which similarity was either 50% or less but no visible modification was made, we present below the sentences with their Bert score in Figure 4a–c.

Original dialogue	Modified dialogue	Reasoning for modification
Sometimes people benefit from training their brain. That can be done through puzzles, such as Sudoku's or crosswords, but there are many ways in which it can be done.	I am glad to hear that you are open to the idea of training your brain. There are indeed many ways to do it, and puzzles like Sudoku or crosswords are excellent options. I know it might sound difficult in the beginning, but as you explore different methods, we can find one that suits you best and makes the process enjoyable.	I added the relational cue "I am glad to hear that" to express encouragement and positivity regarding the person's openness to brain training. By acknowledging the potential difficulty in the beginning and offering support in finding an enjoyable method, I aim to make the conversation empathic and motivational.

(a)

Original dialogue	Modified dialogue	Reasoning for modification
Ah well, you probably know now that it is good for you	Well, you probably know now that it is good for you. If you agree, you can consider incorporating it into your routine.	I avoided starting the dialogue with "if you agree" per the instructions. Instead, I included relational cues like "you can consider" to empower the user to think about integrating the beneficial activity into their routine. This approach maintains a motivational and encouraging tone while respecting the user's autonomy in making decisions.

(b)

Original Dialogue	Modified Dialogue	Reasoning for Modification
A good piece of meat occasionally... perfect! Although fish is a better alternative.	Let's work on incorporating a good portion of protein into your diet occasionally by adding either fish, meat or any other food of your choice that contains proteins. Through our partnership, we can find the right balance that suits your preferences and health goals. This is a team effort, and I believe we can make your meals both enjoyable and nutritious.	I added relational cues to convey a collaborative and supportive tone. Phrases like "Let's work on" and "together, we can find the right balance" emphasize a joint effort in making dietary choices. By stating "This is a team effort," I reinforce the idea that the client and I are working together towards their health goals, creating a positive working alliance.

(c)

Figure 4. (a) Statement received Bert score of 100% neutral. (b) Statement received Bert score of 100% neutral. (c) Statement received Bert score of 53% entailed, 47% neutral.

5. Results Comparing Helpfulness of Human vs. LLM-Generated Dialogue

The total number of users in Cohort 1 who evaluated the helpfulness of the manually created RCs was 284. Of these, 69% are Psychology students, 1% are doing a Computing major, 15% are studying Health Sciences, 4% are business students, 4% are studying Arts and the remaining 7% are enrolled in other subjects. More than half of the study population is female (70%), and the mean age is 21.55 (s.d. 7.3) ranging from 18 to 74. For Cohort 3, the total users who completed their coaching session were 248. A proportion of 60% of the users are male, similar to the distribution of gender for Cohort 1. The mean age is 37 (s.d. 11) with the range from 18 to 76, hence this cohort is more representative of a wider population of adults, rather than students as in Cohort 1.

To identify the helpfulness of relational cues, 12 human-generated empathic dialogues were analysed, 4 for each relational cue category. Table 6 shows the RCs' mean and standard deviation for Cohort 1. To evaluate the LLM-generated dialogue, nine empathic cues were evaluated by Cohort 3. Table 7 shows the relational cues with mean and std for this cohort. The current work aims to understand whether users found human-generated cues or LLM-generated ones more helpful. Table 8 presents the independent t-test between relational cue categories in both studies to see if the differences are significant. This work also aims to compare the significance of the differences in ratings between the three categories in both studies, as shown in Table 9.

Table 6. Human-generated relational cues—evaluated by Cohort 1.

Relational Cue	Mean	Std
I hope together we can make progress towards becoming healthier.	3.97	0.92
Together we can decide what else suits your lifestyle	4.2	0.81
As a team effort, if I show you how you're doing, you can see how close you are to reaching your health goals.	3.98	0.91
Together we can plan your diet according to your favorites	4.22	0.86
Working alliance summary	3.88	0.88
If you agree to share this data, we can keep track of your steps.	3.69	0.99
You can try other possibilities like playing a game of chess or even a game of Candy Crush or Wordfeud on a smartphone	3.88	1.05
Would you like me to elaborate on brain training or mindfulness exercises?	4.06	0.88
If you set your own goal, chances are higher that you actually stick with them.	4.07	0.84
Empowerment summary	3.92	0.64
You are a fast learner.	3.73	0.98
I know it is not easy to find time for such activities.	3.87	0.94
I know you will be feeling pushed in the beginning, but it will get easier.	4.02	0.94
Yes that's a very good question!	3.80	0.99
Affirmation summary	3.85	0.75

Table 7. LLM-generated relational cues—evaluated by Cohort 3.

Relational Cue	Mean	Std
As a team, let's explore a simple and nutritious dinner option	3.85	0.99
Through collaboration, we can explore the importance of balancing your diet with various food group	3.88	0.98
Let's unite our efforts to set specific and achievable healthy diet goals tailored to your needs	3.94	0.97
Working alliance summary	3.88	0.88
If you'd like, you can start this journey with me	3.62	1.00
You can choose a variety of activities that align with your preferences and lifestyle	4.13	0.83
You have the option to engage in activities that promote both physical activity and social interaction	3.95	0.86
Empowerment summary	3.89	0.75
I am glad to hear that you are open to the idea of training your brain	3.71	0.98
That's a very good decision to explore mindfulness exercises	3.79	0.91
I know it might sound difficult in the beginning	3.53	1.02
Affirmation summary	3.67	0.81

Table 8. Comparison of helpfulness of manual versus LLM-generated relational cues.

Groups						
Relational Cue Categories	Mean Human	SD Human	Mean LLM	SD LLM	T-STAT	PVAL
WA	4.09	0.70	3.88	0.88	−3.06	0.002
EMP	3.92	0.64	3.89	0.75	−0.498	0.61
AFF	3.85	0.75	3.67	0.81	−2.66	0.008
Overall	3.95	0.61	3.82	0.74	−2.22	0.026

WA = Working Alliance; AFF = Affirmation; EMP = Empowerment.

Table 9. Helpfulness of the relational cues by categories.

Groups						
Relational Cue Categories	Mean Diff.	SD	PVAL	TSTAT	Confidence Interval	Cohen's d
H-WA-AFF	0.239	0.55	<0.001	7.15	0.30–0.55	0.43
L-WA-AFF	0.213	0.66	<0.001	5.04	0.19–0.44	0.32
H-WA-EMP	0.167	0.59	<0.001	4.707	0.16–0.40	0.28
L-WA-EMP	−0.009	0.54	0.786	−0.271	−0.14–0.10	−0.01
H-EMP-AFF	0.071	0.63	0.062	1.87	−0.01–0.23	0.11
L-EMP-AFF	0.223	0.58	<0.001	−5.99	0.25–0.5	0.38

H = Human; L = LLM; WA = Working Alliance; AFF = Affirmation; EMP = Empowerment.

The results indicate a highly significant p -value between the ratings for working alliance and affirmation with a low-moderate effect size, suggesting a strong relationship between these two categories of relational cues for both human and LLM-generated dialogues. A similar pattern emerged when comparing the ratings for empowerment and affirmation cues, with the p -value also found to be highly significant, with low-moderate effect size for LLM-generated dialogue, but not for the human-created dialogue. Conversely, the non-significant p -value observed between the ratings for LLM-generated empowerment and working alliance cues implies that these two categories may be perceived as the same by users, revealing the LLM's inability to differentiate these two categories of cues. Human versus LLM analysis shows that there is a significant difference in all three categories for human-generated dialogues (EMP–AFF was approaching significance) but, for LLM-generated, two pairs of categories were significantly different and one was not significant (WA–EMP). The effect size ranged from small to medium for all pairs except for LLM-generated (WA–EMP) and Human-generated (EMP–AFF), which were insignificant.

6. Discussion

The first main objective of our study was to determine the efficacy of ChatGPT in generating dialogues with relational cues for multiple coaches and to take a step towards automation of smart coaching. The dialogues of the three coaches cover a variety of lifestyle coaching practices including diet, exercise and cognitive well-being. To measure the impact of the relational cues, these dialogues will be used in future studies. The dialogues express a range of relational cues from neutral (i.e., no relational cues) to highly relational/empathic (i.e., all classes of relational cues included). Our long-term vision is to provide tailored inclusion of relational cues according to user's needs and preferences. This validation study sets the foundation for building a simulation for digital coaching. The validation of the dialogues is a necessary step in the process, since ChatGPT can hallucinate and may produce out of context dialogues [64].

Results of the Bert score show that it found the lexical meaning to be similar between the original and the modified. Nevertheless, for validation, we recommended a hybrid approach, which includes human as well as statistical validation. A few disagreements among the human validators could be due to an expectation that the modified sentences would be of a similar length. We received anecdotal comments from some participants along those lines. However, since we asked ChatGPT to add relational cues, we should expect the modified sentences to be longer and to contain additional content. We also noted that there may be differences in human interpretation due to biases and uncontrollable factors, such as personal traits, as well as understanding of the context [74–76]. For a few dialogues where human validation failed to confirm similarity, a Bert score or other measure of similarity could be used to aid decision making as to whether the dialogue should be modified further or should stay intact.

For RQ1 and RQ2, we measured the similarity of the neutral with the relational version of each modified utterance through human validation of the lexical meaning of the sentence and used the Bert score for measuring structural semantics. Looking at the results of human validation, 80% of the dialogues were rated as similar by 50% or more evaluators. Firstly, we note that human validation does not provide unequivocal results. However, it can be concluded that the use of ChatGPT for sentence modification did not alter most user's interpretations of the lexical meaning of the compared sentences. The remaining 20% that were not rated similar by 50% or more evaluators were analyzed further by the researchers to determine what could be the probable reason users found them dissimilar. In one of the dialogues stating, "Some people do mindfulness exercises at home as well. For example, in the statement "breathing slowly or light yoga. Schedule it according to your routine," it is specifically mentioned that the exercise needs to be scheduled according to the person's routine. In the modified dialogue, "Consider scheduling moments for activities like breathing slowly or incorporating light yoga," there is no mention of making this part of an existing routine, which seems to be the probable cause of why users found it dissimilar. The sentence was further modified to add this factor and the final sentence was modified to "Consider scheduling moments for activities like breathing slowly or incorporating light yoga in your routine".

For the other three sentences that had lower similarity, there was no specific element that was found missing in the modified dialogue. The researchers turned to a review of the Bert score to see if that would shed some light. The Bert score rated two of the sentences as neutral and the third one as a mix of entailment and neutral, which means that all three of them are not contradicting their original version. Therefore, the answer to RQ1 is that ChatGPT mostly created sentences that are similar in meaning to their neutral version, suggesting ChatGPT as a potential starting point for insertion of relational cues, but not a completely reliable approach. In our specific context, we have validated all sentences generated by ChatGPT in our coaching dialogue and manually modified any sentences that were not found to be similar. For health applications, we propose that human validation by a domain expert is essential and that the key benefit of Generative AI is to reduce the initial burden on humans to create the sentences from scratch. Particularly in health, humans cannot be taken out of the loop, but this process saves time. The approach provides two experts, one human and one machine, which improves the process and quality of output.

For RQ2 the Bert score showed promising results. None of the dialogues and their modifications were found contradictory in any way. More than 80% of the dialogues were fully or partially entailed, which means that the Bert score found them to be semantically similar. We conclude that a hybrid approach to validating ChatGPT output provides an approach where the lexical meaning is validated by humans and semantic validation is provided by the automated Bert score.

For RQ3 “Does the reasoning provided by ChatGPT align with the modification approach required during prompt design?” it is imperative to understand how ChatGPT generates its reasoning and why its reasoning is important for our work. The reasoning feature of ChatGPT helps the user understand the chain of thoughts which led the LLM to the respective output [77]. In 90% of cases, more than 50% of users agreed that the reasoning is valid and in accordance with the context of relational cues. For the two dialogues whose reasoning was disagreed upon by 50% or more, this could be attributed to the fact that reasoning concerned itself more with relational cues, and how incorporating cues is what the task at hand consisted of. Reasoning does not mention the modification of the task-based dialogues to either generalize more or to connect them with their previous dialogue’s context. This missing information can help the user reach a more agreeable decision. Nevertheless, it can be concluded that, for RQ3, the reasoning produced by ChatGPT was logical and agreeable.

The second main goal of this work was to see whether the ChatGPT-generated dialogues are as helpful as the human-generated dialogues. This is measured by looking at the ratings for helpfulness of each relational cue category in both LLM dialogues versus the human dialogue setting for empathic interactions. This leads us into RQ4, which discusses the comparison of the helpfulness of human- versus LLM-generated dialogues. Both dialogue sets have empathic relational cues as part of the dialogue content. Examples of relational cues are given in Tables 6 and 7, which also give the mean rating of users for that cue. The statistical comparison of human- and LLM-generated dialogues as given in Table 8 indicate that human dialogues were liked more, with their mean being higher than their respective LLM-generated dialogues in the same category. Overall, relational cues were liked more when they were hand-crafted by a human. The rapid advancements in large language models have led to significant progress in natural language processing, with LLMs demonstrating impressive capabilities in various applications, including text generation, question-answering, and conversational interactions [8]. However, the evaluation of the quality and user perception of LLM-generated content remains an important area of research. Existing research has explored the usage and limitations of LLMs as judges, highlighting the potential for position, verbosity, and self-enhancement biases, as well as limited reasoning ability [78]. The LLM-as-an-Interviewer approach has been proposed as a novel evaluation paradigm that addresses these shortcomings by incorporating dynamic feedback and follow-up questions [16]. Human-generated content is still considered more empathic and relevant than LLM-generated especially in coaching.

The findings in Table 8 suggest that relational cues in health and lifestyle coaching dialogues can have complex and nuanced impacts on user perceptions. While some cues may be perceived as complementary, others may be viewed as mutually exclusive, depending on the specific needs and preferences of the individual user. The significant differences observed between the three categories highlight the importance of carefully considering the design and implementation of relational cues in coaching interventions, to ensure they align with the target audience’s preferences and needs [79].

Lifestyle coaching has become a popular tool for individuals seeking personal and professional development. One key aspect of effective lifestyle coaching is the use of relational cues, which can help foster a sense of trust, empowerment, and collaboration between the coach and client [80]. However, research suggests that not all relational cues are equally effective in the eyes of clients [81]. Existing literature indicates that clients who find empowerment cues helpful in life coaching also tend to find working alliance cues beneficial [79]. This suggests a link between clients’ perceptions of their own sense of agency and their desire for a strong, collaborative relationship with their coach. In contrast,

affirmation cues, which are designed to provide encouragement and validation, are not consistently seen as helpful by clients [79,82].

The results of the two studies were compared, where the Cohort 1 study tested the helpfulness of the relational cues with human-generated dialogues, whereas the second study with Cohort 3 tested the helpfulness of the relational cues with LLM-generated dialogues. We note that Cohort 1 and 3 participants did not know how (i.e., whether manually or AI-generated) the statements with relational cues had been created. In both studies, affirmation was found to be less helpful than empowerment and working alliance. The working alliance rating was significantly different in both studies when compared to affirmation. In the case of empowerment, LLM-generated dialogues' rating was significantly different when compared to affirmation but not for human-generated dialogues. For working alliance and empowerment, human-generated dialogues were significantly different, but not LLM-generated dialogues. It can be seen that empowerment had a higher helpfulness rating in LLM-generated dialogues whereas in human-generated dialogues working alliance had a higher helpfulness rating. It can be argued that LLM can effectively empower the users [83] but, in the case of building a working alliance, humans are still trusted more [84]. It is evident from the research that there is a clear reservation about the use of AI technology within the therapeutic alliance space. It was notable that the users' largest concern was about the potential to create a strong bond with an LLM-driven agent. Affirmation is subjective in long-term adherence and is found less helpful in a single coaching session. The existing literature suggests that the long-term adherence and utility of affirmation is still an ongoing debate [85–87].

The importance of the coach–client relationship in the success of coaching interventions is further underscored by research showing that the relationship plays a mediating role between the coaching received and the development of the client's self-efficacy [88]. The Cohen's *d* values for all significant pairs given in Table 9 are between 0.11 and 0.43, indicating that the observed differences in client perceptions of the various relational cues are not likely to be due to chance [88,89]. Of note, we observe that the LLM was not able to differentiate between the categories of working alliance and empowerment in the statements produced.

To evaluate the preferences and perceptions of users, few studies have been conducted where participants were presented with dialogues from both human-generated and LLM-generated sources. Examples of such studies include [78,90]. The preference for human-generated dialogues over those generated by Large Language Models is a multifaceted phenomenon rooted in human psychology, encompassing aspects of trust, empathy, and the perception of genuine understanding. While LLMs have made remarkable strides in mimicking human language, they often fall short of capturing the nuances, emotional intelligence, and contextual awareness that characterize human communication [91]. This discrepancy leads individuals to perceive human-generated dialogues as more helpful and likeable. One significant factor contributing to this preference is the inherent human tendency to trust content created by other humans more readily [92]. This bias stems from our evolutionary history, where trust in fellow humans was crucial for survival and social cohesion. We are naturally inclined to believe that other humans share our values, understand our experiences, and have our best interests at heart [93]. In contrast, LLMs, despite their sophistication, are perceived as artificial entities lacking genuine understanding and empathy. This perception diminishes the sense of trust and connection that is essential for effective communication and collaboration.

7. Limitations, Future Work and Conclusions

Our goal was to be able to tailor the dialogue to the individual's preferences for certain relational cues (i.e., empowerment, working alliance, affirmation and social dialogues) in an automated manner by using Generative AI. We have shown that it is feasible to provide the LLM with evidence-based health advice (the what) and one or more relational cues to modify "how" the advice is provided to the individual. The main motivation behind this evaluation study was to analyze the utilization of LLMs in digital coaching and to determine the right prompt, essential for ensuring that the meaning and intent of the health advice was not modified but which accurately embedded the relational cues seeking to improve adherence to the advice. Validation of the generated statements with embedded relational cues recommended a mix of automated and human validation. Comparison between the human and LLM-generated dialogues evaluating the helpfulness of the dialogues revealed a preference for the human created sentences. This also suggests a hybrid automated and manual approach to dialogue creation.

Limitations of the work include the use of three different populations in Cohorts 1–3, which was unavoidable due to the incremental nature of the development and evaluation of our coaches. There is a limited number of relational cue examples and, in order to validate the findings, a larger sample is crucial to automate further. In the future, similar populations can be given LLM-generated dialogues in different variations of interactions where multiple combinations and styles can be tested to see how a similar population responds to different interaction styles. Finally, a key limitation is that the collaborative cues were not validated or compared for helpfulness. This was not deemed possible because we had no prior handcrafted collaborative dialogue. However, we are conducting studies with coaching sessions using the LLM-generated dialogue and will be able to report in a future publication on their effectiveness for behaviour changes, as well as how users experience the "banter" and reinforced messages provided by multiple coaches. Both sets of dialogue, collaborative and user-directed, need to be validated by health experts so that anomalies in the health space are not given as recommendations.

Author Contributions: Conceptualization, S.S. and D.R.; methodology, S.S. and D.R.; software, S.S. and D.R.; validation, S.S. and D.R.; investigation, S.S. and D.R.; resources, S.S. and D.R.; data curation, S.S. and D.R.; writing—original draft preparation, S.S. and D.R.; writing—review and editing, S.S. and D.R.; supervision, D.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by a Macquarie University HDR grant and an ARC Research Grant DP200102131.

Data Availability Statement: Data are available on request to the authors.

Acknowledgments: Our thanks to Meredith Porte for her assistance in creating the three coaches.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

The full dialogue set of the three coaches, where original dialogue was given to ChatGPT and the modified dialogue was used in Cohort 3.

Table A1. Complete Original Dialogue with Modified ChatGPT dialogue.

Original Dialogue	Coach	ChatGPT Dialogue
Hi, I'm Paul, your dietician.	Paul	
Which one of us would you like to talk to first?	Paul	
To benefit fully from multiple coaches being present, please talk to all of them. Who do you want to talk next?	Paul	
Let's continue!	Paul	
What do you do as a diet coach	Paul	
As a diet coach, I will talk about healthy food habits. I know about weight loss as well. I can start by sharing a simple recipe for dinner?	Paul	As a team, let's explore a simple and nutritious dinner option. I'm here to help you make delicious yet healthy choices in your weight loss journey.
Try an amazing beets salad, with cashew nuts, small pieces of apple, and goat cheese.	Paul	How about trying an amazing salad together? We can incorporate cashew nuts, small pieces of apple, and cheese for a delightful and nutritious meal. I'm here to help you with the recipe.
Incorporating vibrant and nutrient-rich foods like beets, cashews, apples, and goat cheese not only pleases your taste buds but also supports cognitive function. A nourished brain leads to better decision-making and adherence to your goals.		
That beet salad sounds fantastic, Paul! Nutrient-dense meals like these provide sustained energy for your workouts. Consider having it as a pre-workout meal to fuel your body effectively.		
To be healthy, you need to balance your diet with different food groups. Choose one of the food groups to know more about balanced diet	Paul	Through collaboration, we can explore the importance of balancing your diet with various food groups. Choosing one of these food groups will help us plan a nutritious and well-rounded diet. I am here to guide you through this journey, and as a team, we can achieve your goal of maintaining a healthy lifestyle.
Fresh fish almost every day is a good source of protein and omega 3.	Paul	Together, we can plan to incorporate fresh fish into your daily meals. This will not only provide you with a good source of protein but also essential omega-3 fatty acids. I hope, as a team, we can make this a positive addition to your diet.
Including fresh fish regularly not only nourishes your body but also supports cognitive function. Omega-3s are essential for brain health, aiding in focus and memory. Keep up the good choices for a sharp mind and a strong body.		
Nice choice! The protein from fish is excellent for muscle recovery and growth. Omega-3s also help reduce inflammation, benefiting your joints. It's a win-win for your fitness journey. Keep incorporating such nutritious foods for optimal physical well-being.		

Table A1. Cont.

Original Dialogue	Coach	ChatGPT Dialogue
a good piece of meat once in a while. . . perfect! Although fish is a better alternative.	Paul	Let's work on incorporating a good portion of protein into your diet occasionally by adding either fish, meat or any other food of your choice that contains proteins. Through our partnership, we can find the right balance that suits your preferences and health goals. This is a team effort, and I believe we can make your meals both enjoyable and nutritious.
at least 250 g a day is necessary! There are so many different recipes with a lot of vegetables! You can plan your diet according to your favorites.	Paul	We can also aim for at least 250 g of vegetables a day, ensuring you receive essential nutrients. There are numerous recipes featuring a variety of vegetables, and I'm here to help you plan your diet based on your favorites. As a team, we can make your journey towards a healthier lifestyle enjoyable and personalized.
Vegetables and salads should be a part of regular intake.	Paul	Let's also make vegetables and fruits a consistent part of your daily intake through mutual goal setting. I am here to help you explore delicious ways to incorporate them into your meals. With close coordination, we can ensure that your regular intake is not only nutritious but also enjoyable
Absolutely, Paul! Incorporating a variety of vegetables supports cognitive health. The diverse nutrients contribute to brain function and mood regulation. Let's explore ways to make your veggie intake not just a routine but a joyous part of your day.		
Great advice, Paul! Vegetables and salads provide the essential vitamins and minerals needed for optimal physical performance. They also aid in recovery post-workout. How can we make your veggie intake more enjoyable?		
In any case, setting a healthy diet goal is important. Drinking more water, eating less sugar, or eating more fruit. . . these are just examples. I know a lot about healthy eating but in the end, everyone needs goals to follow.	Paul	Let's unite our efforts to set specific and achievable healthy diet goals tailored to your needs. Whether it's drinking more water or reducing sugar intake, I am here to guide you. I believe, With team effort, we can turn your knowledge into actionable goals, making your journey towards a healthier lifestyle more structured and successful.
I hope you can succeed in becoming healthier. Whom do you want to talk next?	Paul	
To benefit fully from multiple coaches being present, please talk to all of them. Who do you want to talk next?	Paul	
To benefit fully from multiple coaches being present, please talk to all of them. Who do you want to talk next?	Paul	

Table A1. Cont.

Original Dialogue	Coach	ChatGPT Dialogue
And I'm Susan. I'm your physical activity coach!	Suzan	
As your physical activity coach I can help you become more active and lead a healthier life.	Suzan	As your physical activity coach, I'm here to support you in becoming more active and guiding you towards a healthier life. If you'd like, you can start this journey with me.
With 'physical activity' I don't mean running or spending the entire day in the gym.	Suzan	'Physical activity' doesn't necessarily mean running or spending the entire day in the gym. You can choose a variety of activities that align with your preferences and lifestyle.
Physical activity is about everything you do in your daily life: groceries, walking the dog, or some gardening.	Suzan	Physical activity encompasses everything you do in your daily life, from groceries to walking the dog or engaging in some gardening. You have a variety of everyday options with which to stay active.
Absolutely, Suzan! Every little bit adds up. Consider those daily activities as opportunities for movement, contributing to your overall well-being.		
Suzan, it's not just about what's on your plate; it's about embracing a lifestyle. Even those everyday activities contribute to your energy balance.		
Do activities indoor. Walking up the stairs or some cycling on a home trainer, everything counts!	Suzan	Do activities indoors, like walking up the stairs or cycling on a home trainer; everything counts! You can start incorporating these into your routine.
Absolutely, Suzan! Indoor activities are fantastic. They not only boost physical health but also enhance cognitive function. Every step or pedal contributes to your overall well-being.		
Great point, Suzan! Indoor activities are not only effective for burning calories but also create a positive environment. How about pairing them with a nutritious snack for an extra energy boost?		
But it is an easy way to exercise indoors when it is raining outside.	Suzan	But it's an easy way to exercise indoors, especially when it's raining outside. You can choose indoor activities that align with your preferences.
Go outside. . . walking is good for you. A short trip to the supermarket or visiting some friends.	Suzan	Consider going outside; walking is good for you. You can take a short trip to the supermarket or visit some friends to incorporate physical activity into your day.
I love that idea, Suzan! Being outdoors not only benefits your physical health but also lifts your mood. A short walk can do wonders for your cognitive well-being.		

Table A1. Cont.

Original Dialogue	Coach	ChatGPT Dialogue
Suzan, outdoor activities are excellent for burning calories. How about turning that trip to the supermarket into a mindful shopping experience, focusing on nutritious choices?		
Social connection is important for a healthy and happy mind.	Suzan	Recognize that social connection is crucial for a healthy and happy mind. You have the option to engage in activities that promote both physical activity and social interaction.
Ah well, you probably know now that it is good for you	Suzan	Well, you probably know now that it is good for you. If you agree, you can consider incorporating it into your routine.
Why? Well, there are many ways in which you can be active daily. Where will I start. . . Feeling fit and healthy, losing a few pounds and balancing more. . .	Suzan	Well, there are many ways in which you can be active daily. Where will you start? Feeling fit and healthy, losing a few pounds, and balancing more are all within your reach.
A healthy lifestyle starts with understanding your current behaviour. If I show you how you're doing, you can see how close you are to reaching your health goals.	Suzan	A healthy lifestyle starts with understanding your current behaviour. If I show you how you're doing, you can see how close you are to reaching your health goals. Would you like guidance in assessing your progress?
I hope you succeed in becoming healthier!	Suzan	
That's all from me for now. Who else would you like to talk to?	Suzan	
I'm Meena, your cognitive therapist.	Meena	
As a cognition coach, I will talk about brain training and mindfulness exercises. Let's talk about brain training first?	Meena	Great you seem to be interested in brain training! It's a wonderful way to enhance cognitive abilities. There are various approaches, such as puzzles, games, and even activities like learning a new instrument or language. I know it might sound difficult in the beginning, but together, we can explore options and tailor a plan that aligns with your preferences and goals. You are a fast learner, and I'm confident that incorporating brain training into your routine can lead to positive outcomes.
Sometimes people benefit from training their brain. That can be done through puzzles, such as Sudoku's or crosswords, but there are many ways in which it can be done.	Meena	I am glad to hear that you are open to the idea of training your brain. There are indeed many ways to do it, and puzzles like Sudoku or crosswords are excellent options. I know it might sound difficult in the beginning, but as you explore different methods, we can find one that suits you best and makes the process enjoyable.

Table A1. Cont.

Original Dialogue	Coach	ChatGPT Dialogue
Nutrition plays a role in cognitive function too. Foods rich in omega-3 fatty acids, antioxidants, and vitamins, like fish, berries, and leafy greens, can support brain health.		
Absolutely, Paul. Adequate hydration is crucial too. Water supports cognitive function, and incorporating hydrating foods like cucumbers and watermelon can contribute to overall brain health.		
I can also think of other brain exercises that people may enjoy doing. Try other possibilities like playing a game of chess or even a game of Candy Crush or Wordfeud on a smartphone.	Meena	That's a very good point! Brain exercises come in various forms, and it's wonderful that you're considering different possibilities. Playing a game of chess is a fantastic option, and even engaging in smartphone games like Candy Crush or Wordfeud can be enjoyable brain exercises. I understand change does not come easily, but by exploring diverse activities, you can find something that resonates with you and contributes to your cognitive well-being.
It's a great point, Meena. Nutrition can complement these activities. Including nuts and seeds for brain-boosting nutrients like omega-3 fatty acids and antioxidants can further support cognitive function.		
Absolutely, Meena. Incorporating physical activities that require coordination, like dancing or even trying out a new sport, can also enhance cognitive function by stimulating different areas of the brain.		
That's true, Paul! Calculating the oven time for a recipe is as well.	Meena	
I will talk about mindfulness exercises now	Meena	That's a very good decision to explore mindfulness exercises. I know you will be feeling pushed in the beginning, but I assure you, it's a positive step toward well-being. I am here to support you as you delve into this journey of self-care.
Mindfulness exercises could be breathing exercises or walking in fresh air. Some find a walk through the forest to be a very relaxing activity.	Meena	Mindfulness exercises can take various forms, such as engaging in breathing exercises or taking a refreshing walk in the fresh air. Some individuals find that strolling through a forest can be an exceptionally relaxing activity.
Absolutely, Meena. Nutrition also plays a role in supporting mental well-being. Foods rich in magnesium, like nuts and leafy greens, can help regulate stress hormones and contribute to a calmer state of mind.		

Table A1. Cont.

Original Dialogue	Coach	ChatGPT Dialogue
True, Meena. Combining mindfulness with physical activity, like a forest walk, not only relaxes the mind but also enhances overall well-being. It's a great way to promote mental and physical balance.		
Some people do mindfulness exercises at home as well. For example, breathing slowly or light yoga. Schedule it according to your routine.	Meena	I know it might sound difficult in the beginning, but integrating mindfulness exercises into your routine, even at home, can make a significant impact. Consider scheduling moments for activities like breathing slowly or incorporating light yoga. I understand change does not come easily, but these small adjustments can contribute to your overall well-being.
I hope you can succeed in improving your cognition through the suggestions given in this session. Who do you want to talk next?	Meena	

References

- Gámez Díaz, R.; Yu, Q.; Ding, Y.; Laamarti, F.; El Saddik, A. Digital twin coaching for physical activities: A survey. *Sensors* **2020**, *20*, 5936. [\[CrossRef\]](#) [\[PubMed\]](#)
- Abdulrahman, A.; Richards, D. In Search of Embodied Conversational and Explainable Agents for Health Behaviour Change and Adherence. *Multimodal Technol. Interact.* **2021**, *5*, 56. [\[CrossRef\]](#)
- Bickmore, T. Relational Agents. Investigating the Use of Relational Agents in Task Domains Such as in Coaching, Counseling, Psychotherapy and Healthcare. 2019. Available online: <http://www.ccs.neu.edu/home/bickmore/agents> (accessed on 28 March 2024).
- Beinema, T.; Akker, H.O.D.; van Velsen, L.; Hermens, H. Tailoring coaching strategies to users' motivation in a multi-agent health coaching application. *Comput. Hum. Behav.* **2021**, *121*, 106787. [\[CrossRef\]](#)
- Ranjartabar, H.; Richards, D.; Bilgin, A.A.; Kutay, C.; Mascarenhas, S. Adapting a virtual advisor's verbal conversation based on predicted user preferences: A study of neutral, empathic and tailored dialogue. *Multimodal Technol. Interact.* **2020**, *4*, 55. [\[CrossRef\]](#)
- Richards, D.; Bilgin, A.A.; Ranjartabar, H. Users' perceptions of empathic dialogue cues: A data-driven approach to provide tailored empathy. In Proceedings of the 18th International Conference on Intelligent Virtual Agents, Sydney, Australia, 5–8 November 2018.
- Baker, S.; Richards, D.; Caldwell, P. Relational agents to promote ehealth advice adherence. In Proceedings of the Pacific Rim International Conference on Artificial Intelligence; Springer: Gold Coast, Australia, 2014; pp. 1010–1015.
- Guo, T.; Chen, X.; Wang, Y.; Chang, R.; Pei, S.; Chawla, N.V.; Wiest, O.; Zhang, X. Large language model based multi-agents: A survey of progress and challenges. *arXiv* **2024**, arXiv:2402.01680.
- Lentferink, A.J.; Oldenhuis, H.K.; de Groot, M.; Polstra, L.; Velthuisen, H.; van Gemert-Pijnen, J.E. Key components in eHealth interventions combining self-tracking and persuasive eCoaching to promote a healthier lifestyle: A scoping review. *J. Med. Internet Res.* **2017**, *19*, e277. [\[CrossRef\]](#)
- Sforzo, G.A.; Kaye, M.P.; Harenberg, S.; Costello, K.; Cobus-Kuo, L.; Rauff, E.; Edman, J.S.; Frates, E.; Moore, M. Compendium of health and wellness coaching: 2019 addendum. *Am. J. Lifestyle Med.* **2020**, *14*, 155–168. [\[CrossRef\]](#)
- Kantharaju, R.B.; De Franco, D.; Pease, A.; Pelachaud, C. Is two better than one? Effects of multiple agents on user persuasion. In Proceedings of the 18th International Conference on Intelligent Virtual Agents, Sydney, Australia, 5–8 November 2018.
- Nouraei, F.; Rebello, K.; Fallah, M.; Murali, P.; Matuszak, H.; Jap, V.; Parker, A.; Paasche-Orlow, M.; Bickmore, T. Virtual Agent-Based Communication Skills Training to Facilitate Health Persuasion Among Peers. *arXiv* **2024**, arXiv:2412.12061.
- Mohan, S. Exploring the role of common model of cognition in designing adaptive coaching interactions for health behavior change. *ACM Trans. Interact. Intell. Syst.* **2021**, *11*, 1–30. [\[CrossRef\]](#)

14. Kantharaju, R.B.; Pease, A.; Reidsma, D.; Pelachaud, C.; Snaith, M.; Bruijnes, M.; Klaassen, R.; Beinema, T.; Huizing, G.; Simonetti, D. Integrating argumentation with social conversation between multiple virtual coaches. In Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents, Paris, France, 2–5 July 2019.
15. Fox, J.; Bailenson, J.N. Virtual self-modeling: The effects of vicarious reinforcement and identification on exercise behaviors. *Media Psychol.* **2009**, *12*, 1–25. [\[CrossRef\]](#)
16. Lee, U.; Jung, H.; Jeon, Y.; Sohn, Y.; Hwang, W.; Moon, J.; Kim, H. Few-shot is enough: Exploring ChatGPT prompt engineering method for automatic question generation in english education. *Educ. Inf. Technol.* **2023**, *29*, 11483–11515. [\[CrossRef\]](#)
17. Salman, S.; Richards, D. Collaborating Digital Twins for Health Coaching. In *Advances in Practical Applications of Agents, Multi-Agent Systems, and Digital Twins, The PAAMS Collection*; Mathieu, P., De la Prieta, F., Eds.; Springer: Salamanca, Spain, 2024.
18. Ranjartabar, H.; Richards, D.; Bilgin, A.A.; Kutay, C. Personalising the Dialogue of Relational Agents for First-Time Users. In Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems, Online, 3–7 May 2021.
19. Zhang, T.; Kishore, V.; Wu, F.; Weinberger, K.Q.; Artzi, Y. Bertscore: Evaluating text generation with bert. *arXiv* **2019**, arXiv:1904.09675.
20. Johnson, D.; Goodman, R.; Patrinely, J.; Stone, C.; Zimmerman, E.; Donald, R.; Chang, S.; Berkowitz, S.; Finn, A.; Jahangir, E. Assessing the Accuracy and Reliability of AI-Generated Medical Responses: An Evaluation of the Chat-GPT Model. *Res. Sq.* **2023**.
21. Abdulrahman, A.; Richards, D.; Bilgin, A.A. Changing users' health behaviour intentions through an embodied conversational agent delivering explanations based on users' beliefs and goals. *Behav. Inf. Technol.* **2023**, *42*, 1338–1356. [\[CrossRef\]](#)
22. Hopman, K.; Richards, D.; Norberg, M.M. A Digital Coach to Promote Emotion Regulation Skills. *Multimodal Technol. Interact.* **2023**, *7*, 57. [\[CrossRef\]](#)
23. Jeong, S.; Aymerich-Franch, L.; Arias, K.; Alghowinem, S.; Lapedriza, A.; Picard, R.; Park, H.W.; Breazeal, C. Deploying a robotic positive psychology coach to improve college students' psychological well-being. *User Model. User-Adapt. Interact.* **2023**, *33*, 571–615. [\[CrossRef\]](#)
24. Michaelsen, M.M.; Esch, T. Understanding health behavior change by motivation and reward mechanisms: A review of the literature. *Front. Behav. Neurosci.* **2023**, *17*, 1151918. [\[CrossRef\]](#)
25. Teixeira, P.J.; Marques, M.M. Health behavior change for obesity management. *Obes. Facts* **2018**, *10*, 666–673. [\[CrossRef\]](#)
26. Kleinsinger, F. The unmet challenge of medication nonadherence. *Perm. J.* **2018**, *22*, 18–033. [\[CrossRef\]](#)
27. Miller, N.H. Adherence behavior in the prevention and treatment of cardiovascular disease. *J. Cardiopulm. Rehabil. Prev.* **2012**, *32*, 63–70. [\[CrossRef\]](#)
28. Raamkumar, A.S.; Yang, Y. Empathetic conversational systems: A review of current advances, gaps, and opportunities. *IEEE Trans. Affect. Comput.* **2022**, *14*, 2722–2739. [\[CrossRef\]](#)
29. Sharma, N.P.; Gupta, V. Human behavior in a social environment. In *StatPearls [Internet]*; StatPearls Publishing: Treasure Island, FL, USA, 2022.
30. Ferrari, M.; Speight, J.; Beath, A.; Browne, J.L.; Mosely, K. The information-motivation-behavioral skills model explains physical activity levels for adults with type 2 diabetes across all weight classes. *Psychol. Health Med.* **2021**, *26*, 381–394. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Evers, K.E.; Balestrieri, S.G. The transtheoretical model and stages of change. *Health Behav. Theory Res. Pract.* **2024**, *73*. [\[CrossRef\]](#)
32. Olafsson, S.; O'Leary, T.; Bickmore, T. Coerced change-talk with conversational agents promotes confidence in behavior change. In Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare, Trento, Italy, 20–23 May 2019.
33. Asbjørnsen, R.; Børøsd, E.; Hjelmæsæth, J.; Smedsrød, M.; Ollivier, M.; Wentzel, J.; Clark, M.; van Gemert-Pijnen, J.; Solberg Nes, L. Digital behaviour change intervention for weight loss maintenance in adults with obesity: A feasibility pilot study of eCHANGE. *Behav. Inf. Technol.* **2024**, 1–19. [\[CrossRef\]](#)
34. van Velsen, L.; Broekhuis, M.; Jansen-Kosterink, S.; Op den Akker, H. Tailoring persuasive electronic health strategies for older adults on the basis of personal motivation: Web-based survey study. *J. Med. Internet Res.* **2019**, *21*, e11759. [\[CrossRef\]](#)
35. Aloni, N. Empowering Dialogues in Humanistic Education. *Educ. Philos. Theory* **2013**, *45*, 1067–1081. [\[CrossRef\]](#)
36. Cameron, R.A.; Mazer, B.L.; DeLuca, J.M.; Mohile, S.G.; Epstein, R.M. In search of compassion: A new taxonomy of compassionate physician behaviours. *J. Health Expect.* **2015**, *18*, 1672–1685. [\[CrossRef\]](#)
37. Graßmann, C.; Schölmerich, F.; Schermuly, C.C. The relationship between working alliance and client outcomes in coaching: A meta-analysis. *Hum. Relat.* **2020**, *73*, 35–58. [\[CrossRef\]](#)
38. Abdulrahman, A.; Richards, D. Modelling Working Alliance Using User-aware Explainable Embodied Conversational Agent for Behaviour Change: Framework and Empirical Evaluation. 2019. Available online: https://aisel.aisnet.org/icis2019/human_computer_interact/human_computer_interact/12/ (accessed on 10 March 2024).
39. Wallerstein, N. Powerlessness, empowerment, and health: Implications for health promotion programs. *Am. J. Health Promot.* **1992**, *6*, 197–205. [\[CrossRef\]](#)
40. Gibson, C.H. A concept analysis of empowerment. *J. Adv. Nurs.* **1991**, *16*, 354–361. [\[CrossRef\]](#)

41. Skuladottir, H.; Halldorsdottir, S. Women in chronic pain: Sense of control and encounters with health professionals. *Qual. Health Res.* **2008**, *18*, 891–901. [\[CrossRef\]](#)
42. Tveiten, S.; Knutsen, I.R. Empowering dialogues—the patients’ perspective. *Scand. J. Caring Sci.* **2011**, *25*, 333–340. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Tremain, H.; McEnery, C.; Fletcher, K.; Murray, G. The therapeutic alliance in digital mental health interventions for serious mental illnesses: Narrative review. *JMIR Ment. Health* **2020**, *7*, e17204. [\[CrossRef\]](#)
44. Bickmore, T.W.; Caruso, L.; Clough-Gorr, K. Acceptance and usability of a relational agent interface by urban older adults. In Proceedings of the CHI’05 Extended Abstracts on Human Factors in Computing Systems, Portland, OR, USA, 2–7 April 2005; pp. 1212–1215.
45. Moore, C. Variable Wordforms, Adaptable Learners: Evidence from Real-time Word Comprehension and Naturalistic Corpora. Doctoral Thesis, Duke University, Durham, NV, USA, 2021.
46. Arnold, R. *Empathic Intelligence: Teaching, Learning, Relating*; UNSW Press: Randwick, Australia, 2005.
47. Burmaoglu, S.; Saritas, O.; Yalcin, H. Defense 4.0: Internet of things in military. *Emerg. Technol. Econ. Dev.* **2019**, 303–320. [\[CrossRef\]](#)
48. Horvath, A.O.; Del Re, A.; Flückiger, C.; Symonds, D. Alliance in individual psychotherapy. *Psychotherapy* **2011**, *48*, 9–16. [\[CrossRef\]](#)
49. Cherbib, J.; Chebbi, H.; Yahiaoui, D.; Thrassou, A.; Sakka, G. Digital technologies and learning within asymmetric alliances: The role of collaborative context. *J. Bus. Res.* **2021**, *125*, 214–226. [\[CrossRef\]](#)
50. Wampold, B.E.; Flückiger, C. The alliance in mental health care: Conceptualization, evidence and clinical applications. *World Psychiatry* **2023**, *22*, 25–41. [\[CrossRef\]](#)
51. Walsh, J.F. *The Dynamics of the Social Worker-Client Relationship*; Oxford University Press: Oxford, UK, 2021.
52. Roth, T. Insomnia: Definition, prevalence, etiology, and consequences. *J. Clin. Sleep Med.* **2007**, *3*, S7–S10. [\[CrossRef\]](#)
53. De Herder, W.; Krenning, E.; Van Eijck, C.; Lamberts, S. Considerations concerning a tailored, individualized therapeutic management of patients with (neuro) endocrine tumours of the gastrointestinal tract and pancreas. *Endocr.-Relat. Cancer* **2004**, *11*, 19–34. [\[CrossRef\]](#)
54. Castonguay, L.G.; Constantino, M.J.; Holtforth, M.G. The working alliance: Where are we and where should we go? *Psychother. Theory Res. Pract. Train.* **2006**, *43*, 271. [\[CrossRef\]](#)
55. Ackerman, S.J.; Hilsenroth, M.J. A review of therapist characteristics and techniques negatively impacting the therapeutic alliance. *Psychother. Theory Res. Pract. Train.* **2001**, *38*, 171. [\[CrossRef\]](#)
56. Ablett, M. *Mindfulness Practice Uses in Countertransference: A Phenomenological Study*; Institute for Clinical Social Work (Chicago): Chicago, IL, USA, 2023.
57. Schechter, M.A.; Goldblatt, M.J. Psychodynamic therapy and the therapeutic alliance: Validation, empathy, and genuine relatedness. In *Building a Therapeutic Alliance with the Suicidal Patient*; American Psychological Association: Washington, DC, USA, 2011.
58. Saunders, S.M. Examining the relationship between the therapeutic bond and the phases of treatment outcome. *Psychother. Theory Res. Pract. Train.* **2000**, *37*, 206. [\[CrossRef\]](#)
59. Kaddour, J.; Harris, J.; Mozes, M.; Bradley, H.; Raileanu, R.; McHardy, R. Challenges and applications of large language models. *arXiv* **2023**, arXiv:2307.10169.
60. Li, P. An empirical investigation of pre-trained transformer language models for open-domain dialogue generation. *arXiv* **2020**, arXiv:2003.04195.
61. Hudeček, V.; Dušek, O. Are large language models all you need for task-oriented dialogue? In Proceedings of the 24th Annual Meeting of the Special Interest Group on Discourse and Dialogue, Prague, Czechia, 11–15 September 2023.
62. Marvin, G.; Hellen, N.; Jjingo, D.; Nakatumba-Nabende, J. Prompt Engineering in Large Language Models. In *International Conference on Data Intelligence and Cognitive Informatics*; Springer: Berlin/Heidelberg, Germany, 2023.
63. Kalyan, K.S. A survey of GPT-3 family large language models including ChatGPT and GPT-4. *Nat. Lang. Process. J.* **2023**, *6*, 100048. [\[CrossRef\]](#)
64. Amaro, I.; Della Greca, A.; Francese, R.; Tortora, G.; Tucci, C. AI Unreliable Answers: A Case Study on ChatGPT. In Proceedings of the International Conference on Human-Computer Interaction, Lleida, Spain, 4–6 September 2023.
65. Cer, D.; Yang, Y.; Kong, S.-y.; Hua, N.; Limtiaco, N.; John, R.S.; Constant, N.; Guajardo-Cespedes, M.; Yuan, S.; Tar, C. Universal sentence encoder. *arXiv* **2018**, arXiv:1803.11175.
66. Kocoń, J.; Cichecki, I.; Kaszyca, O.; Kochanek, M.; Szydło, D.; Baran, J.; Bielaniec, J.; Gruza, M.; Janz, A.; Kanclerz, K. ChatGPT: Jack of all trades, master of none. *Inf. Fusion* **2023**, *99*, 101861. [\[CrossRef\]](#)
67. Calvaresi, D.; Carli, R.; Piguet, J.-G.; Contreras, V.H.; Luzzani, G.; Najjar, A.; Calbimonte, J.-P.; Schumacher, M. Ethical and legal considerations for nutrition virtual coaches. *AI Ethics* **2023**, *3*, 1313–1340. [\[CrossRef\]](#)

68. Floridi, L.; Cowls, J.; Beltrametti, M.; Chatila, R.; Chazerand, P.; Dignum, V.; Luetge, C.; Madelin, R.; Pagallo, U.; Rossi, F. AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds Mach.* **2018**, *28*, 689–707. [\[CrossRef\]](#)
69. Salman, S.; Richards, D.; Dras, M. Identifying Which Relational Cues Users Find Helpful to Allow Tailoring of e-Coach Dialogues. *Multimodal Technol. Interact.* **2023**, *7*, 93. [\[CrossRef\]](#)
70. Durieux, B.; Davis, J.; Moons, P.; Van Bulck, L. *How to Get the Most out of ChatGPT? Tips and Tricks on Prompting*; Oxford University Press: Oxford, UK, 2024; Volume 23, pp. e128–e130.
71. op den Akker, H.; op den Akker, R.; Beinema, T.; Banos, O.; Heylen, D.; Bedsted, B.; Pease, A.; Pelachaud, C.; Salcedo, V.T.; Kyriazakos, S. Council of coaches a novel holistic behavior change coaching approach. In Proceedings of the 4th International Conference on Information and Communication Technologies for Ageing Well and e-Health, ICT4AWE 2018, Funchal, Portugal, 22–23 March 2018; SciTePress: Setúbal, Portugal, 2018.
72. Fialho, P.; Coheur, L.; Quaresma, P. Benchmarking natural language inference and semantic textual similarity for portuguese. *Information* **2020**, *11*, 484. [\[CrossRef\]](#)
73. Gaur, C.; Kumar, A.; Das, S. *Analysis and Design of Document Similarity Using BiLSTM and BERT*; Springer Nature: Cham, Switzerland, 2023.
74. Olson, D. From utterance to text: The bias of language in speech and writing. *Harv. Educ. Rev.* **1977**, *47*, 257–281. [\[CrossRef\]](#)
75. Bromme, R.; Rambow, R.; Nückles, M. Expertise and estimating what other people know: The influence of professional experience and type of knowledge. *J. Exp. Psychol. Appl.* **2001**, *7*, 317. [\[CrossRef\]](#)
76. Trout, J.D. Scientific explanation and the sense of understanding. *Philos. Sci.* **2002**, *69*, 212–233. [\[CrossRef\]](#)
77. Wei, J.; Tay, Y.; Bommasani, R.; Raffel, C.; Zoph, B.; Borgeaud, S.; Yogatama, D.; Bosma, M.; Zhou, D.; Metzler, D. Emergent abilities of large language models. *arXiv* **2022**, arXiv:2206.07682.
78. Zheng, L.; Chiang, W.-L.; Sheng, Y.; Zhuang, S.; Wu, Z.; Zhuang, Y.; Lin, Z.; Li, Z.; Li, D.; Xing, E. Judging llm-as-a-judge with mt-bench and chatbot arena. *Adv. Neural Inf. Process. Syst.* **2023**, *36*, 46595–46623.
79. Shanmuganathan-Felton, V.; Felton, L.; Jowett, S. It Takes Two: The Importance of the Coach-Athlete Relationship. *Front. Young Minds* **2022**. [\[CrossRef\]](#)
80. Bickmore, T.W.; Schulman, D.; Sidner, C.L. A reusable framework for health counseling dialogue systems based on a behavioral medicine ontology. *J. Biomed. Inform.* **2011**, *44*, 183–197. [\[CrossRef\]](#)
81. Carrozza, R. Understanding and Reaching the New Gen Z & Millennial Pet Parents-Communication Program and Media Plan for New Dog Parents. Master's Thesis, Universidade NOVA de Lisboa (Portugal), Lisboa, Portugal, 2023.
82. Lanier, C.H.; Bean, P.; Arnold, S.C. *Motivational Interviewing in Life and Health Coaching*. Guilford Publications: New York, NY, USA, 2024.
83. Jörke, M.; Sapkota, S.; Warkenthien, L.; Vainio, N.; Schmiedmayer, P.; Brunskill, E.; Landay, J. Supporting physical activity behavior change with llm-based conversational agents. *arXiv* **2024**, arXiv:2405.06061.
84. Prescott, J.; Hanley, T. Therapists' attitudes towards the use of AI in therapeutic practice: Considering the therapeutic alliance. *Ment. Health Soc. Incl.* **2023**, *27*, 177–185. [\[CrossRef\]](#)
85. Cohen, G.L.; Sherman, D.K. The psychology of change: Self-affirmation and social psychological intervention. *Annu. Rev. Psychol.* **2014**, *65*, 333–371. [\[CrossRef\]](#) [\[PubMed\]](#)
86. Tay, I.Q.Y. *Adversity Roadmap: Improving Psychological Well-Being and Resilience Through a Values-Affirmation Approach*; Stanford University: Stanford, CA, USA, 2023.
87. Ponce-Correa, F.; Caqueo-Úrizar, A.; Berrios, R.; Escobar-Soler, C. Defining recovery in schizophrenia: A review of outcome studies. *Psychiatry Res.* **2023**, *322*, 115134. [\[CrossRef\]](#) [\[PubMed\]](#)
88. Baron, L.; Morin, L. The coach-coachee relationship in executive coaching: A field study. *Hum. Resour. Dev. Q.* **2009**, *20*, 85–106. [\[CrossRef\]](#)
89. Bar-Kalifa, E.; Goren, O.; Gilboa-Schechtman, E.; Wolff, M.; Rafael, D.; Heimann, S.; Yehezkel, I.; Scheniuk, A.; Ruth, F.; Atzil-Slonim, D. Clients' emotional experience as a dynamic context for client-therapist physiological synchrony. *J. Consult. Clin. Psychol.* **2023**, *91*, 367. [\[CrossRef\]](#)
90. Arora, N.; Chakraborty, I.; Nishimura, Y. Revolutionizing Marketing Research with a Large Language Model: A Hybrid AI-Human Approach. *SSRN 4683054* **2024**. [\[CrossRef\]](#)
91. Amini, M.; Alavi, M.; Saber Alavi, J.; Sadacharam, M.G.; Alipour, M.; Amini, D. A Review on the Role of Moral Emotions in EFL Learning. *Int. J. Acad. Res. Progress. Educ. Dev.* **2024**, *13*, 1. [\[CrossRef\]](#)

92. Szczuka, J.M.; Horstmann, A.C.; Szymczyk, N.; Strathmann, C.; Artelt, A.; Mavrina, L.; Krämer, N. Let Me Explain What I Did or What I Would Have Done: An Empirical Study on the Effects of Explanations and Person-Likeness on Trust in and Understanding of Algorithms. In Proceedings of the 13th Nordic Conference on Human-Computer Interaction, Uppsala, Sweden, 13–16 October 2024.
93. Schindler, S.; Hilgard, J.; Fritsche, I.; Burke, B.; Pfattheicher, S. Do salient social norms moderate mortality salience effects? A (challenging) meta-analysis of terror management studies. *Personal. Soc. Psychol. Rev.* **2023**, *27*, 195–225. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.