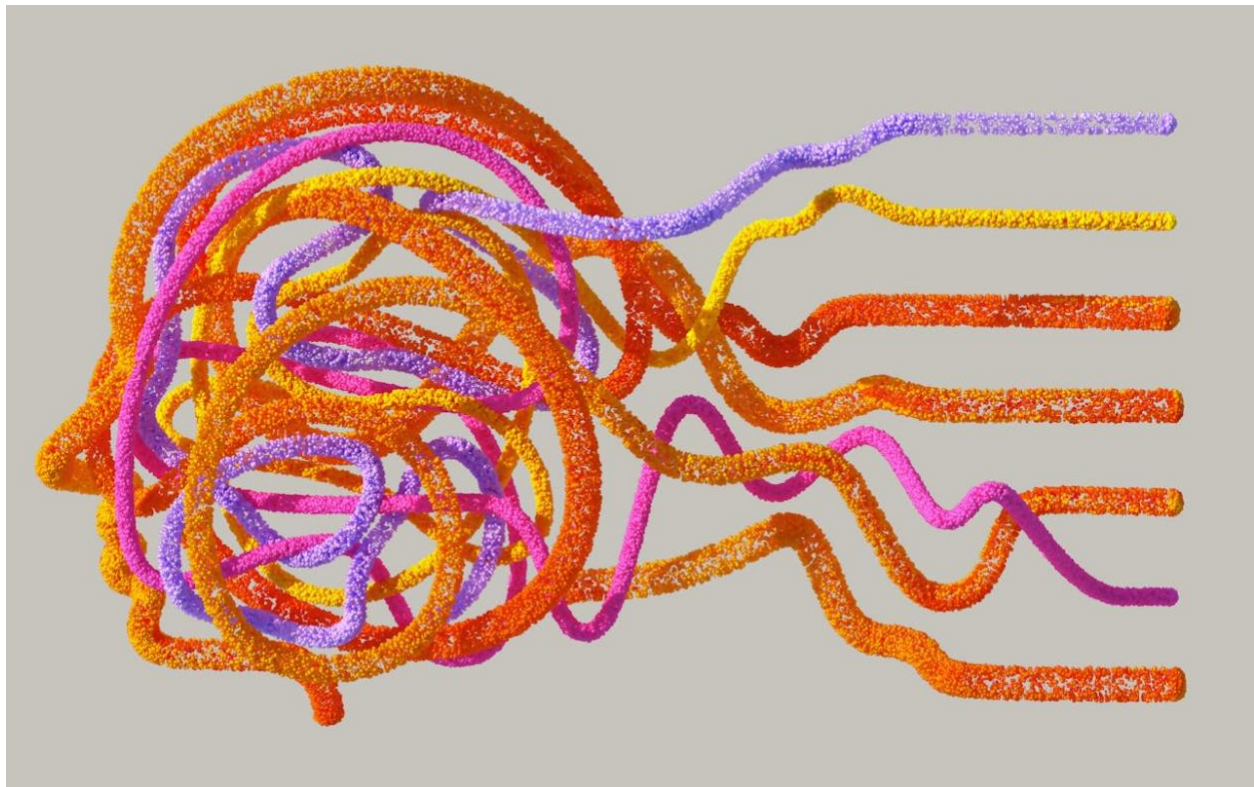


A systematic literature review of artificial intelligence (AI) in coaching: Insights for future research and product development



Jonathan Passmore, Bergsveinn Olafsson & David Tee

Citation: Passmore, J., Olafsson, B. Tee, D. & (2025). A systematic literature review of artificial intelligence (AI) in coaching: Insights for future research and product development, *Journal of Work – Applied Management*. Doi:10.1108/JWAM-11-2024-0164

Abstract

Purpose

Artificial intelligence (AI) has the potential to dramatically change the human approaches to work, and specifically to learning and development. While AI coaching can reduce costs and increase accessibility, it also presents both opportunities and threats to human coaches. The objective of this study was to conduct a systematic literature review of peer-reviewed research on the use of AI in coaching.

Design

A SLR method was used to search eight databases for articles produced up to March 2024. Data extraction was conducted, with Quality Assessment undertaken independently, in parallel, using two researchers and a third arbiter. The ROBINS-I tool was used to assess risk of bias in the included studies. A narrative synthesis of a total of 16 quantitative, qualitative or mixed-methods studies covering $n = 2312$.

Findings

The SLR identified four key themes: Research design and AI integration, AI usefulness in coaching, impact of AI coaching, and ethical considerations. The findings suggest that AI coaches can be effective, accepted, useful and match human coaches in competence for specific tasks.

Practical Implications

AI coaching is a growing area of practice and research. This paper brings together the literature and identifies future research priorities and potential next steps in AI Coach development.

Originality

The paper uses clinical research SLR methods applying these robust processes to the field of organisational research, to set a new standard through the use of a pre-determined research protocol, quality assessment and ROB, well providing a comprehensive literature review of AI coaching.

Keywords Systematic Literature Review, AI coaching, AI ethics, coaching outcomes, technology coaching, coaching avatars, AI coach effectiveness

Introduction

Artificial intelligence (AI), defined as machines that can mimic human capabilities such as problem-solving, pattern recognition, and decision-making (Russel & Norvig, 2020), has been forecast to put 300 million jobs at risk over the next decade (Briggs & Kodnani, 2023). Simultaneously, AI is projected to create many new jobs and is currently dramatically affecting various sectors, from healthcare, finance, education, manufacturing, retail, to the creative industries (Anantrasirichai & Bull, 2022; Dwivedi et al, 2023; Lee & Yoon, 202).

One application of AI is large language models (LLMs), which can understand, generate, and interact with human language based on vast amounts of data (Brynjolfsson & McAfee, 2014). The creation of the Transformer architecture (Vaswani et al., 2017) accelerated the development of LLMs such as Google's BERT, OpenAI's GPT-3 and GPT-4, Meta's LLaMA, and Microsoft's Co-pilot (Giattino et al., 2023). Notably, completion LLMs like GPT-3 and BERT differ from chat-based LLMs like GPT-4 and LLaMA in their training and development. This paper focuses on the latter, emphasizing conversational LLMs that have evolved significantly with the help of reinforcement learning for aligning models, giving rise to more interactive AI systems.

LLMs can be highly effective in tasks such as text summarization, coding, language translation and generating creative content (Liu et al., 2023). In the context of coaching, LLMs can provide real-time feedback, and assist with administrative tasks, thereby potentially enhancing the overall coaching experience while allowing coaches to focus more on complex, human-centric aspects of their work (Passmore & Tee, 2023).

LLMs are undoubtedly going to affect professional services, including the coaching industry. In this article coaching is defined as: a voluntary intervention involving a series of future-focused, structured, purposeful conversations characterised by open questions, listening, summaries, reflections and affirmations, intended to facilitate the client in generating and acting upon strategies which result in developing greater self-awareness, enhancing personal responsibility and achieving meaningful progress towards a desired change. Coaching is used across various contexts, including sports, education, healthcare and business (Athanasopoulou & Dopson 2018; Smith & Smoll, 2017; van Nieuwerburgh, 2016; Wolever et al, 2013). The application of coaching in work settings has accelerated; in 2022, 89% of organisational buyers of coaching estimated an increased investment in coaching in the upcoming year (McKenzie et al., 2022). Within the empirical literature, meta-analyses indicate coaching having a medium to large effect on various desirable outcomes, such as performance, goal attainment, subjective well-being, self-efficacy, hope, resilience and optimism (Cannon-Bowes et al, 2023; De Haan & Nilsson, 2023; Jones et al., 2016).

The uptake of LLMs in coaching will have major implications for coach training schools, coaching clients, and coaches. From the coaching client side, LLM applications in coaching have the potential to make coaching available and affordable for more people (Terblanche, 2022b). Traditional coaching requires two human beings to meet in person or through digital means, which requires energy, time and cost for coaching clients. By leveraging LLMs in coaching through AI text or voice chatbots, which enables users to engage in conversational exchanges (Chung & Park, 2019), more people can gain inexpensive 24/7 access to AI coaching, defined as a “Synchronous coaching experience, where the machine replaces the role of the human coach, facilitating their human client in goal setting, issue exploration, personal reflection and developing insights and actions” (Passmore & Tee, 2024).

From the coaches’ perspective, scholars have previously suggested that an AI coach is capable of replacing a beginner coach, who relies heavily on existing coaching models and techniques, while being unable to replace an experienced coach who has moved beyond a simplistic model, such as GROW or other equivalent frameworks (Graßmann & Schermuly, 2021; Terblanche et al, 2022b). This is unsurprising given AI bots were script-based until November 2022. However, the arrival of Chat GPT has reignited progress and early generative AI bots still show problematic behaviour (Passmore & Tee, 2024). However, AI has the potential to both radically change people development at work and creates an existential threat to human coaches in the medium term. In the short term AI can also be seen to complement human coaches by automating repetitive tasks such as writing notes from a session, writing up client’s agreed actions or intersession activities, as well as undertaking coach-client matching, analysing coach performance during a session, advising coaching clients on which intersessional activities may best support their development, monitoring progress and providing nudges to clients (Passmore & Tee, 2023). These complimentary applications have the potential to make coaches, and coaching providers more effective. Given these potential uses, AI is also likely to impact coaching training schools, as upcoming coaches need to better understand how these tools work, and how they can design or leverage them to enhance their own practice (Passmore & Woodward, 2023).

Existing studies on AI in coaching are nascent but are expected to develop rapidly within the next decade. This systematic literature review aims to provide a platform for future research, focusing on a review of peer review published papers, as well as inform the development of coachbots which are starting to emerge from multiple coaching providers.

In formal terms the primary objective of this study is to conduct a systematic literature review of peer-reviewed research regarding the use of artificial intelligence (AI) in coaching. The secondary objective of the study is to determine whether there is evidence for the effectiveness of AI as a coaching tool to generate desired and intended outcomes. Consequently, the research questions that guided this systematic review were:

1. How has AI been integrated into coaching in the empirical literature?
2. What methods and research designs have been used to study AI in coaching?
3. What does the empirical evidence suggest on the usefulness and impact of AI in coaching?
4. What ethical considerations do studies emphasize?
5. How can future studies and product developers advance the understanding and integration of AI in coaching?

By systematically reviewing the empirical literature on AI in coaching guided by these questions, we inform researchers and practitioners of the evidence that supports or does not support the effectiveness of certain AI applications when integrated into coaching. Consequently, the implications of the systematic review are insight to evidence of the effectiveness of AI in coaching, which can direct product developers and future studies within this emerging field with evidence-based applications of AI in coaching.

Method

This systematic review sought to be conducted in accordance with the 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses' (PRISMA) guidelines (Moher, Liberati, Tetzlaff & Altman, 2009). This included a systematic review (SR) protocol being created, submitted for inclusion in online protocol databases and circulated amongst the researchers to guide the data gathering and analysis stages.

Eligibility criteria

The type of studies for consideration in the SR were based on the following criteria: (1) Produced in the medium of English, (2) Published in peer-reviewed publications, (3) The intervention was labelled as coaching, (4) The coach was not required to have technical knowledge of client's role or desired outcome, (5) The term 'AI' or 'Artificial intelligence' was included in the title or abstract text, (6) The research produced primary data, or was a systematic review/meta-analysis (7) The study was published between 1st January 1990 (an arbitrary date assumed to precede all AI coaching research) and 1st March 2024.

Various techniques and approaches are used within coaching, typically rooted in psychological approaches. This SR did not restrict itself to any specific approach or domain. Therefore, the coaching could have been conducted by self-coaching, another human or by technology, and could have been a one-with-one or one-with-many intervention.

Search strategy and information source

The search term coach* + artificial intelligence was used to identify peer-reviewed qualitative, quantitative and mixed methods studies in the following databases: Business Source Complete, Emerald Journals, ProQuest, PubMed, PsycArticles, PsycINFO, Science Direct and Web of Science. A forward and backward search of the

identified studies was subsequently conducted. In addition, the researchers checked specialist coaching publication websites for recently accepted but unpublished articles. Finally, coaching researchers with published studies featuring coaching using AI were contacted to ensure that any 'in press' studies could be included in this SR.

Selection process

Once duplications from the searches across different databases were removed, the search strategy produced an initial set of 60 peer-reviewed articles. The titles and abstracts for each of these articles were independently screened against the eligibility criteria by two reviewers using reference manager software. The full texts were obtained for the 42 remaining relevant articles, which were again independently screened by two reviewers. At both screening stages, any disagreement regarding eligibility was resolved through escalation to the third researcher. As detailed in Figure 1, following the screening of the full texts, a backward/forward search then identified three further articles.

Quality assessment and Risk of bias

The 'Quality Assessment Tool for Studies with Diverse Designs' (QATSDD) (Sirriyeh, et al., 2012) was used to assess the quality of the included studies. In both instances, two researchers conducted independent parallel assessments of the included studies. Any disagreements were resolved by discussion or escalated to the third researcher, which also ensured appropriate judgement had been reached by the other researchers.

The 'Risk of Bias in Non-randomised Studies - of Interventions' (ROBINS-I) tool (Sterne et al., 2016) was selected for assessing any risks of bias, as it is specifically designed to include non-randomised studies. This was undertaken by one researcher, with a 20% sample undertaken by a second researcher to ensure consistency.

Results

The full-text review resulted in 16 articles being identified that met the eligibility criteria. There were seven quantitative, three qualitative and five mixed-methods studies, published between 2019-2024.

Following the PRISMA process (Page et al., 2021) a total of 27 articles were excluded. Six articles were excluded because they were not primary research or synthesis studies (predominantly being 'thought piece' articles) and 21 were excluded because the 'coaching' intervention was more akin to one-to-one instruction or training, requiring the 'coach' to have specific technical knowledge. The forward/backward search identified 433 articles that were excluded, mainly because the intervention in these referenced studies did not involve AI or meet our definition of coaching. This high rejection rate points to the nascent nature of this field of research, with researchers as yet unable to draw on a large body of extant research and, instead, needing to 'cast their net' over tangential fields of inquiry. A summary of each included study is provided in Table 1.

Figure 1. *PRISMA flow diagram of the screening process* (from Page et al., 2020)

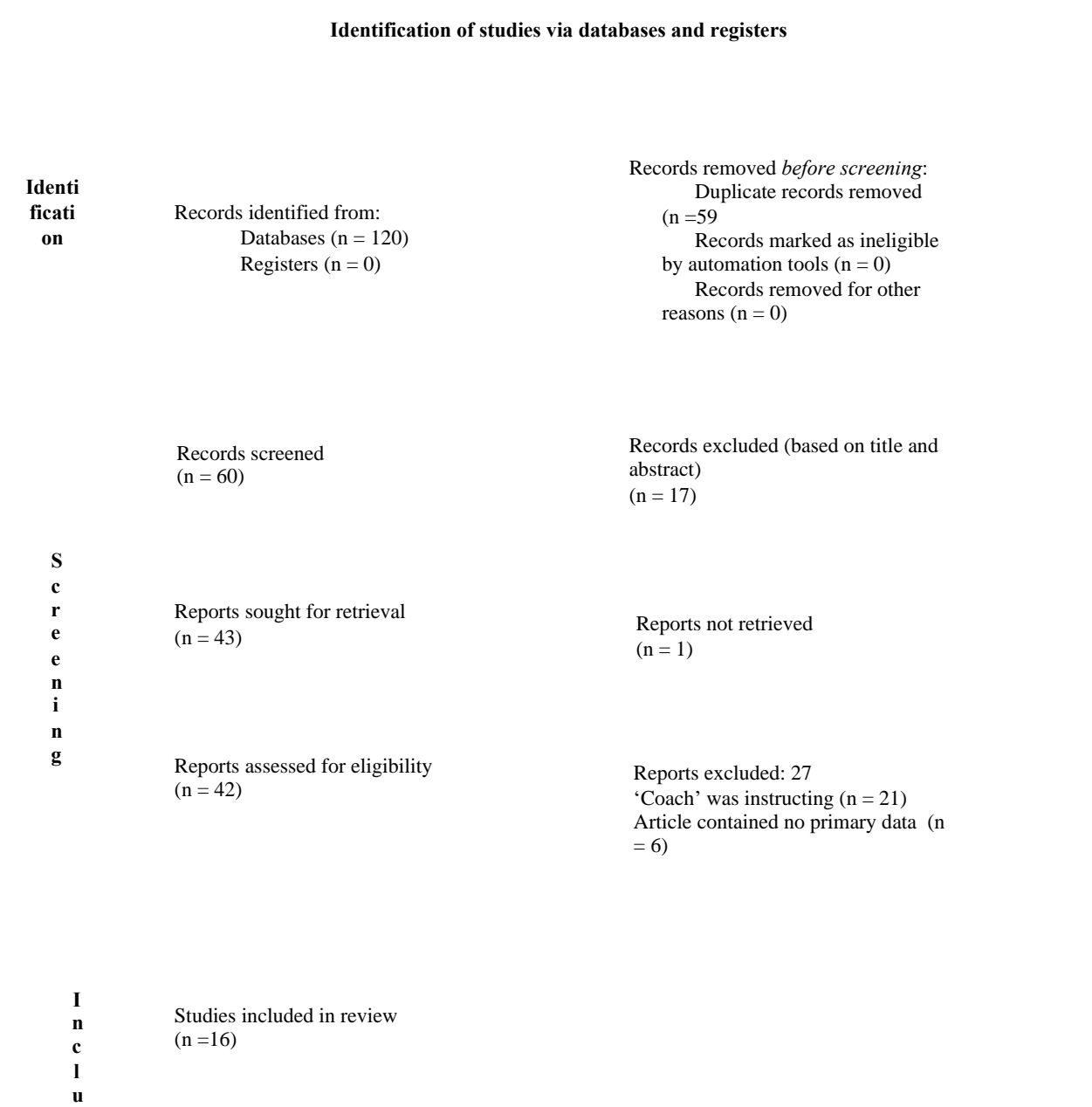


Table 1: Summary of papers

Authors (doi)	Paper type (Qual, RCT, Quasi- Experimental design)	Sample	n	AI coaching delivery	AI coaching training	Results / Findings / Insight
<i>Ellish-Brush (2021)</i> 10.24384/er2p-4857	<i>Mixed-methods: Quasi- experimental design and qualitative</i>	<i>Banking sector, No age reported</i>	48	<i>Text-based</i>	<i>Cognitive behavioural therapy</i>	<i>Improved self- resilience, No significant Working alliance</i>
<i>Figuerola et al (2021)</i> 10.3389/fdgth.2021.747153	<i>Qualitative</i>	<i>Low-income women, aged 27-41, majority Hispanic/ Latine</i>	18	<i>Text-based</i>	<i>Behavioural activation, motivational interviewing, acceptance and commitment, and solution- focused therapy</i>	<i>Positive perception of the chatbot, showed interest in using chatbots for health improvements , concerns about data privacy</i>
<i>Hassoon et al (2021)</i> 10.1038/s41746-021-00539-9	<i>Randomised control Trial (RCT)</i>	<i>Overweight or obese cancer survivors, mean age 62.1 years, 90% female, various cancer types</i>	42	<i>Text-based, Voice- based</i>	<i>Physical activity interventions</i>	<i>Improved step count by voice bot compared to text bot and control, no significant difference between text bot and control</i>
<i>Kannampallil et al (2022),</i> 10.2196/38092	<i>Mixed methods (Observational): Quantitative and qualitative</i>	<i>Patients with mild to moderate depression and/or anxiety, mean age 43.9 years, 77% female, 73% racial or ethnic minorities</i>	26	<i>Voice- based</i>	<i>Problem- Solving treatment</i>	<i>High pragmatic usability and favourable user experience, higher temporal workload during a problem- solving session</i>
<i>Mai et al (2021)</i>	<i>Mixed-methods: Quasi-</i>	<i>Students, aged 21-39,</i>	12	<i>Text-based</i>	<i>Exam anxiety,</i>	<i>Disclosure to the chatbot</i>

10.1007/978-3-030-90328-2_29	experimental design and qualitative	majority male (75%)			solution-focused coaching	and rapport, more self-disclosure and rapport found in the chatbot informational disclosure versus self-disclosure
Mai et al (2022) 10.1038/s41746-021-00539-9	Exploratory quantitative study	University students	21	Conversational AI (writing), Rule-based (clicking)	Exam anxiety, solution-focused coaching	Moderate to high working alliance, higher value observed for bonding in the conversational AI, coaching through chatbots well accepted
Movsumova et al (2020) https://doi.org/10.15219/em86.1485	Mixed-methods (qualitative and quantitative)	Varied demographics, including men and women of different ages, occupations, and positions	33	AI-based tool (Mentorbot) through Telegram, assisting human coaches in dialogue and session analysis	No specific training details for coaches on Mentorbot mentioned.	Positive dynamics clarity, willingness to act and stress reduction. Mentorbot was effective for novel and confidential requests, while human coaches were stronger in reducing stress and perceived overall usefulness
Passmore et al. (2021) 10.53841/bp_stcp.2021.17.2.41	Quantitative: Survey	Coaches from 79 countries, average age 54, 66% female	1200	N/A	N/A	Mixed views of the role of AI in coaching, equally divided seen as providing benefits and disbenefits

Passmore & Tee (2023a) 10.1108/JWgAM-06-2023-0057	Cross-sectional, mixed-method study	Experts in coaching, academic program directors, experience in reviewing and marking coaching assignments	14	Text-based	Various prompts to evaluate GPT-4's ability to define coaching, compare ethical codes, summarise meta-analyses, and conduct coaching	GPT-4 is capable of generating plausible content but often contains inaccuracies and falsified information, Concerns over CPT-4 ethical judgement were highlighted
Stephens et al. (2019) 10.1093/tbm/ibz043	Feasibility study	Youth enrolled in a weight management program, mean age 15.2 years, 57% female, 43% Hispanic	23	Text-based	Behavioural coaching	AI coach was feasible and helpful; high engagement (4,123 messages), 96% found it useful, 81% reported positive progress toward goals
Terblanche & Cilliers (2020), 10.22316/poc/05.1.06	Exploratory study: Survey	Online participants, no age or demographics reported	226	Text-based	Goal-attainment theory, GROW model	Performance expectancy, social influence, and attitude significantly influence behavioural intent to use AI coach
Terblanche et al. (2022a), 10.24384/5c-gf-ab69	Longitudinal RCT	Undergraduate students, diverse demographics, average age 22 years	168	Text-based	Goal-attainment theory, GROW model	Improved goal attainment, no significant changes in psychological well-being, resilience, or perceived stress
Terblanche et al.	Longitudinal RCT	Business school	478	Text-based	Goal-attainment	Improved goal attainment

(2022b), 10.1371/journal.pone.0270255		students, diverse demographics			theory, GROW model	compared to the control group, same effect on goal attainment as human coaches
Terblanche et al. (2023) 10.1080/17521882.2022.2094278	Qualitative study	Final year undergraduate students, aged 20-22, diverse cultures, low socioeconomic background	31	Text-based	Goal-setting	Positive attitude and performance expectations promoted engagement; AI coach perceived as accessible, easy to use, and intelligent; minimal perceived risk; social influence and information about the AI coach influenced adoption
Terblanche et al. (2024) 10.1080/17521882.2024.2304792	Qualitative study	Coaches and clients from a financial services organisation	16 (9 coaches, 7 clients)	Text-based	GROW model	Coaches were concerned about potential negative interference with the coach-client bond, while clients found the chatbot useful for goal tracking, accountability, and convenience. Clients felt psychologically safe with the chatbot and appreciated its non-judgmental nature

<i>Tropeg et al (2019 10.2196/12805)</i>	<i>Scoping review</i>	<i>Mostly within medical care</i>	<i>Not applicable (Review of 49 studies)</i>	<i>Text, voice and avatars, ECA</i>	<i>Various methods, focusing on health improvement</i>	<i>Effectively support physical activity and weight management. ECA have significant potential in promoting healthy behaviours</i>
--	-----------------------	-----------------------------------	--	-------------------------------------	--	--

Risk of bias and Quality results

Eleven of the 16 studies were rated as having a 'low' overall risk of bias, with two studies awarded a 'high' categorisation and two rated as 'unclear', due to key information not being reported in the source studies. Using the 16 items in the QATSDD tool, an average quality rating of 77% was determined, with an inter-rater reliability (IRR) of .67. The most prevalent indicator of quality was a clear rationale for the choice of data collection tool', which was detected in all 16 studies. The lowest scoring quality determinant was 'Statistical assessment of reliability and validity of measurement tool(s)', which was absent from six of the quantitative data studies.

Research design and AI integration

The summary table reveals a heterogeneity in the design characteristics across the included studies. Sample sizes ranged from four to 1200, multiple dependent variables have been studied and research settings include low-income Hispanic women, banking sector employees, and overweight/obese cancer survivors. An important realization was the usage of AI 'AI coaching' as an umbrella construct. AI technology has advanced markedly over the five-year period that the included studies were published, meaning the term 'AI' is used across a wide range of studies from those which use a 'Wizard of Oz' (with a remote researcher creating faux 'AI' responses in real time that appear on the participant's screen), to text-based models and voice-based avatar interactions. Further AI coaches have been trained on diverse approaches, ranging from cognitive behavioural therapy, goal theory, motivational interviewing and solution-focused coaching.

AI usefulness in coaching

Our systematic review suggests that various methodologies in AI coaching are useful for and accepted by people, ranging from voice, text, script- and conversational-based (Figuera et al, 2021; Mai et al, 2021, Kannampallil et al, 2022; Stephens et al., 2019). AI coaches can provide feedback, help with goal-tracking, serve as accountability partners, and are seen by participants in the studies published to date as accessible, convenient and psychologically safe (Terblanche et al, 2023; 2024). Furthermore, AI coaches can create a non-judgmental environment that encourages self-disclosure (Ellish-Brush et al, 2021; Mai et al, 2021; Terblanche et al., 2024).

In addition, an AI-based bot, assisting human coaches with questions to ask during a coaching session and creating a report to reflect after a coaching session, has been shown to broaden coaches' perspectives (Mosvumov et al, 2020). This hybrid approach to coaching was more effective for novel and confidential requests compared to human coaching without AI bot assistance. However, while both showed a positive influence on clarity, willingness to act and stress reduction, human coaches were in the studies included in this review were stronger in reducing stress and perceived overall usefulness

The Impact of AI Coaching

The four studies that utilized either randomized control trials or quasi-experimental research methodologies suggest that AI coaching chatbots designed with a narrow sense of purpose aimed at improving specific outcomes are effective in increasing beneficial outcomes. Terblanche et al (2022a; 2022b) showed that Vici, who was solely trained in goal attainment, was able to match the level of impact of human coaches on goal attainment while being ineffective in increasing resilience and psychological well-being. Elish-Brush (2021), showed that an AI coaching chatbot trained in psychological techniques and resources via cognitive behavioural therapy increased self-resilience. Hassoon et al (2021) showed that an AI coach trained in delivering physical interventions was able to increase the step count of cancer survivors compared to a control group.

In terms of working alliance, Mai et al (2022) found that both script-based (clicking) and conversational (writing) AI coaches demonstrated a medium to a high working alliance, supported by other studies in the review (see Kannampallil et al, 2022; Terblanche et al. 2024). However, Elish-Brush (2021) found that an AI coach was able to increase self-resilience without demonstrating a working alliance. Rather than focusing on working alliance, Terblanche & Cilliers (2020) argued that AI coaches should rather have technology acceptance. Their findings suggest that performance expectations, positive attitude, and social influence predict behavioural intentions of using an AI coach. These findings were supported in a later study, where positive attitude and performance expectations promoted engagement while social influence affected adoption (Terblanche et al., 2023).

Ethical considerations

Surprisingly, few of the papers included in this the review discuss ethical concerns or implications arising from participants. However, a few of them discussed and reported ethical principles and risks including data privacy, security, lack of harm and bias.

Figuera et al (2021) reported that most participants had privacy concerns relating to sharing their location and had limited technological understanding, Elish-Brush (2021) noted that biases in AI can result in greater social prejudice and that no ethical framework exists to guide the applications of digital agents. Passmore & Tee (2023) in their review of GPT-4 expressed concerns over its ability to make ethical judgments and to identify risk during conversations. While GPT-4 communicated empathy, it did not provide referrals to crisis hotlines, emergency services, or relevant professional

services. Terblanche (2023) reported that the perceived risk of using an AI coach was low, as it was seen by students as being unbiased, impartial, and private to use. Terblanche (2020) noted 'best practices' as suggestions for designing an AI coach, including ethical considerations which include trust, transparency, reliability, benevolence, and integrity.

Discussion

We identified 16 papers from the review that met our inclusion criteria for the period from the first of January 2000 to the first of March 2024. As can be noted from the summary table (Table 1) the first papers emerged in 2019, with the volume of papers building during 2021 and 2022. We expect that process to continue given the developments of AI and, in part, this was a motivator for us to provide a platform for future research.

Our systematic review suggests that people are willing to use a range of methodologies in AI coaching and benefit from doing so (Mai et al., 2021; Stephens et al., 2019; Terblanche et al., 2021a;2021b). Combining these insights with the prediction that AI coaching methodologies and features will improve and become more efficient, AI has a huge potential in learning and development and specifically within coaching.

One such potential future use case is the development of embodied conversational agents (ECA) using generative AI with multimodal features (i.e., visual, audio, text, which have become the most popular interfaces in healthcare and psychology (Tropea et al., 2019). Other features of ECAs include gender, aesthetics, voice, personality and behavioural patterns (Tropea et al., 2019). Thus, ECAs might even represent celebrities, political figures or intellectuals in the future, allowing users to receive AI coaching from Marshall Goldsmith, Albert Einstein or Donald Trump.

Independently from AI coaching methodologies and features, the papers in this systematic review indicate that AI coaches that are designed with a narrow sense of purpose aimed at improving specific outcomes are effective in enhancing desirable outcomes such as goal attainment, self-resilience and physical activity (Ellish-Brush 2021, Hassoon et al, 2021; Terblanche et al, 2022a; 2022b). This is echoed by evaluations of coachbots against the ICF coach competencies which have shown AI coaches are capability of meeting ICF credentialing standards at ACC, and show evidence of PCC behaviours (Passmore, 2024).

These findings might suggest that AI coaching are a genuine threat to novice coaches who rely heavily on simple models and approaches in their coaching practice (Graßmann & Schermuly, 2021; Terblanche et al., 2022b). However, caution is needed about the reliability and validity of the findings of the above studies. There are three reasons for this. Firstly, samples have generally been based on student samples. Secondly, they have more typically been based on simple, often health related goals. Finally, there are questions which can be raised about individual studies which are often cited as key evidence. One example, Terblanche et al.'s (2021a;2022b) studies were conducted among undergraduate students and the AI data was gathered during the

Covid-19 lockdown, which may have influenced participants reactions given they were unable to leave home. More research is need to validate these results. Furthermore, the rigorousness of the Elish-Brush (2021) study was questionable, as it featured a within-effect design, did not report the validity and reliability of the questionnaire on self-resilience and had a limited sample size. Similarly, although using randomized control trial design, Hassoon et al. (2021) had a limited sample size of cancer patients.

That said, these insights have some important implications for human coaches and coaching training schools. First, coaches should equip their students with knowledge about AI and support them in gaining the skills to design and deploy AI coachbots. This may involve collaboration with programmers to leverage Retrieval Augmented Generation (RAG) architecture to enhance the quality of the coaching bot performance. Although just two studies in the review covered client benefits associated with AI-assisted methodologies in complementing coaching praxis (Terblanche et al., 2024), coaches should be open to exploring how AI systems can enhance the quality and efficiency of their coaching through reflective practice, client goal-tracking and accountability, automation of repetitive tasks (coaching session reports), inter-sessional activities, assessments and client discovery before the human coaching session (Passmore & Tee, 2023).

Second, although AI's potential is enormous, currently coaches have unique human features that AI does not have. For instance, while AI can demonstrate empathy through data-driven responses, arguably better than human professionals (Ayers et al, 2023), a machine does not feel anything. Empathy is an emotional response that implies you feel what the other person is feeling (Cuff et al, 2016). Thus, 'AI empathy' might be better viewed as cognitive empathy (understanding) compared to affective empathy (emotional response), and thus may create a different effect in the human participant (client).

Other unique human coaching features include adaptability, cultural sensitivity, ethical judgement and privacy. Human coaches can adapt to the needs and change of direction in the coaching conversation while relying on various approaches, tools and techniques learned from coaching, which will be harder for an AI coach. Current AI coaches also lack the 'meta-intervention' of using different coaching approaches and techniques based on the culture in which the coaching occurs or reflecting the cultural background of the client. Furthermore, clients sometimes bring challenging ethical dilemmas to explore in the coaching conversation which coaches might be better to facilitate compared to an AI chatbot. In addition, humans can draw on intuition, which AI will struggle to replicate. Lastly, while AI bots may be lower priced than human coaches, many people will prefer a human coach that shares their identity and lived experience, and to whom they feel accountable.

In terms of the client-AI coach relationship, the review suggest that an AI coach can demonstrate a working alliance, but it might not be as important to outcomes in AI coaching as it is in human coaching (Mai et al; 2021; Terblanche & Cilliers 2020). Although we are unable to make a strong conclusion based on the scarce studies in the review, technology acceptance might be a contender to working alliances when it

comes to the effectiveness of AI coaching. Expecting that an AI coach will perform effectively and having a positive attitude towards AI coaches and other people in one's social environment using it does seem to be particularly important when it comes to behaviour intentions, engagement and adoption (Terblanche & Cilliers, 2020; Terblanche et al., 2023). Future studies could help discern the role of working alliance and technology acceptance in moderating the relationship between AI coaching and desirable client outcomes.

Related to human-AI relationships, ethical considerations are needed to prevent undesirable outcomes and unintended consequences. Although not emphasised in the majority of studies, our review suggests there are several ethical concerns when it comes to AI coaching, including data privacy and security, lack of harm and biases (Ellish-Bruch 2021; Figuera et al, 2021; Passmore & Tee, 2023a). Given these concerns from researchers and participants alike and the multiple ethical questions AI raises, we believe there is a role for professional bodies in managing and regulating the development and applications in AI coaching. Both the ICF and EMCC have published ethical standards relating to AI and digital coaching (see EMCC, 2024; ICF, 2024). These standards, developed with AI developers and coaching platform providers, offer advice and guidance but have avoided setting compliance standards such as those set for human coaches through competency frameworks or ethical codes of conduct. A possible next step is for such frameworks to become quality assurance standards, using frameworks such as ISO, providing information and comfort to organisational buyers and to individual users on data privacy, data storage, bias and ethical practice. Further change may also result from legislation such as The EU's (2024) AI Act and changes to GDPR regulations. Such standards could bring AI bots into line with human coaches, reducing the risk of harm from AI coaches while promoting a focus on human flourishing and holding AI developers accountable for the unintended consequences of their AI coaches (Passmore, et al 2024)

Future Research

Given the expected rise in research on AI coaching, there are many fruitful insights that future research could explore. Due to the lack of strong empirical research in the field, the main focus of future research should be to test the effectiveness of AI coaching by using robust research design including randomised control trials (RCT) and quasi-experimental methods with different participant groups and sufficient participant numbers to ensure enough power. In addition, comparing the effects of human coaches with different levels of expertise (beginner versus experienced coach or managers) to AI coaching on outcomes including goal attainment, self-insight, solution-focused thinking, subjective well-being and psychological well-being will be invaluable. Such studies will reveal greater clarity on in what circumstances and with what clients can AI coaching match or outperform human coaching.

Next, future research should compare different modalities of the delivery of AI coaching, including voice, text and ECAs. Only one study in the review compared voice and text-driven AI bots (Hassoon et al., 2021). While the voice bot was shown to be more

effective in increasing physical activity compared to the text bot, the mechanism of interactions between the user and the AI coaches was different. Therefore, future research should test text versus voice bots with different user groups, while ensuring that the interaction between the bot and the user is identical (e.g. bidirectional or initiated mainly by the AI coach or the participant).

In terms of ECAs, future research will need to discern which features of ECAs are most productive for coaching and which ones lead to unintended consequences, such as dependency, harm and the 'uncanny valley' phenomenon, which refers to negative impressions of machines with overly human-like qualities and negatively correlates with the trustworthiness of ECAs (Passmore et al., 2024; Sasaki et al., 2017; Tropea et al., 2019). This would include comparing the effectiveness of animated-, human-, machine-like (robot) and animal ECAs to important coaching outcomes.

We hypothesises that the closer to the human experience, and the closer to known, or respected individuals the AI bot appears, the better the outcomes that the AI coaching tool will be able to achieve as users project human qualities onto the experience and feel more accountable for their commitments.

Given the lack of ethical considerations in the coaching studies in this review, qualitative and quantitative studies with prospective users and AI designers on ethical principles in AI coaching will be important. From the user's perspective, this means assessing what ethical features of the AI coach increase trust, engagement, adoption and behavioural intentions. This could be done with interviews alongside carefully designed quantitative assessments on ethical principles users deem valuable. On the AI design front, conducting a similar quantitative study, interviews, or a document review analysis on the ethics of existing AI coaching systems according to ethical principles could provide important information on current ethical status and applications (Passmore et al., 2024).

Future product development

Given the rapid advancements and immense potential of AI in coaching, the scope for future product development is vast. While numerous features could enhance the coaching experience, several key developments stand out for their potential to benefit prospective users.

Multimodal interaction: Having various multimodal interaction features in AI coaching that users can pick and choose from may offer enhanced benefits. These multimodal features include text, voice, and video, and a range of types of ECAs. Users should be able to pick and choose what interaction method best meets their need. While most people are expected to talk to and look at the AI coach, others may prefer to interact through writing. In addition, adding various features of ECAs may add value, such as different animated figures, animals or different human images, which offer both known personas and also diversity across race, gender and nationality. These ECAs should also have different behavioural characteristics, personalities, genders, colours and voices associated with them. In each case the designers should make clear what the user is interacting with. Equally interaction styles can be varied. For instance, while

some people may prefer a high level of challenge, others may prefer a more empathetic and compassionate interaction. This gives the user autonomy to pick and choose the features that are most compatible with them, and the issue they wish to explore. In addition, having multimodal features is more likely to meet the needs of neurodiverse individuals and provides greater opportunities to offer different experiences to meet diverse cultural and national backgrounds.

Elements of effective human coaching: AI coaches should incorporate the essential elements of coach-client relationships such as empathy, non-judgemental attitude, trust, transparency, benevolence and integrity (Terblanche & Cilliers, 2020). By offering cartoons or animal representations AI coaches could prevent the impression that bots are human (Sasaki et al., 2017; Terblanche, 2020). For those that use a human image, user information should make clear the user is engaging with a bot not a human. Further consistent with a coaching style, AI coaches should ask one question at a time and should be coded to avoid giving advice.

Specificity: Until we have general artificial intelligence that can perform any intellectual task that human beings can do, AI coaches should be designed with a narrow sense of purpose aimed at achieving specific outcomes. Just like human coaches who normally specialise in people development, as opposed to also offering accounting services and legal advice. These AI coaches should focus on evidenced-based approaches, such as solution-focused coaching, cognitive behavioural coaching or positive psychology coaching. By relying on theoretical models supported by research, AI coaches can provide greater confidence to users that positive outcomes can be achieved (Olafsson et al., 2024).

Advanced Personalisation: In addition to the features of the ECAs discussed above, features that increase personalisation are encouraged. These features should leverage the demographics, cultural background and context of the user. This could be incorporated with simple questions that users can fill out before engaging with the AI coach, and then the AI coach has more data to adjust their interaction based on the client's background and context. Furthermore, having the ability to attach documents, such as resumes, journal entries, pictures and videos while engaging with the AI coach can enhance personalisation. In addition, the integration of wearable technology which gives data on things such as sleep, heart rate variability and exercise habits, is valuable since it gives AI data-driven information on the physiology and lifestyle habits of the user, allowing it to personalize the questions and feedback it gives with greater effectiveness. The balance needed here is, that while increased personalisation is beneficial for the user, the data should not be used to exploit nor manipulate the users with marketing, create dependency or lead to discriminatory outcomes.

Ethical principles: Covering all AI Coach ethical principles is too vast a topic to articulate here (see Passmore, & Tee, 2023; Passmore et al, 2024). That said, transparency with informed consent is encouraged, which explicitly states simply the non-humanness of the AI coach, acknowledges imperfection and biases, and has crystal clear information on data privacy, security and use. This allows prospective users to be aware and

mindful of the implications and risks associated with interacting with the AI coach. Furthermore, AI coaches should only work within their domain of training and expertise, meaning that if the user brings up topics that are not an AI coach's role to facilitate exploring, such as suicidal ideation, or irrelevant out-of-scope question, the AI should make appropriate referrals in the first case and direct the conversation back to relevance in the second case.

In addition, future AI coaching tools must address potential biases, which can arise from the data used to train these systems, leading to unfair or inappropriate outputs. Implementing robust bias detection and mitigation strategies is essential to ensure equitable and inclusive interactions across a global user population (Ferrer et al., 2021; Leong & Sung., 2024). Addressing biases not only enhances the fairness of AI tools but also fosters trust and reliability among users. By proactively managing these issues, developers can create AI systems that align with ethical principles while supporting effective coaching outcomes for varied global audiences.

Although various fruitful features for product development could benefit coaches, they remain out of the scope of the findings of this systematic review.

Limitations to the study

While the study set high standards in developing a research objectives, a research protocol, and using dual evaluations throughout the process, along with an Extraction Template, Quality Assessment process and ROB, all unusual in organisational research but common in clinical studies, this study has several limitations that warrant consideration. First, the researchers established a protocol of the systematic literature review, it was not possible to publish the protocol on an Open Access site, as no such site exists for this type of research. Failing to publish the protocol in advance reduces transparency and increases the risk of design deviation.

Second, there was a change in personnel over the course of the study, as two researchers active at the planning stage of the research dropped out and were replaced by one substitute. Although this change occurred relatively early in the research process, it could have introduced some inconsistencies from protocol to the data collection stage.

Third, while the average quality rating of the studies in the systematic review was 77%, the inter-rater reliability was 0.67, indicating some inconsistencies in how the researchers assessed the quality characteristics of the source studies, which could influence the reliability of the review findings. However, to mitigate this lack of reliability, a third researcher was involved in the arbitration process, in which he independently assessed the quality of the studies and resolved any discrepancies observed by the initial researchers. Including the third researcher increased the likelihood that the final quality assessment was more reliable and thereby enhancing the credibility of the systematic review.

Conclusions

In conclusion, by 2035, the coaching industry, like many others, will look radically differently as a result of AI. Our review indicates that various versions of AI coaches can be useful, accepted and effective, and increasingly will match human levels of competence. Researchers should focus on continuing to assess how AI coaching can contribute to human development and wellbeing, through robust research methods. Product developers should emphasize designing effective and ethical AI coaching tools which serve the needs of their users.

References

- Athanasopoulou, A., & Dopson, S. (2018). A systematic review of executive coaching outcomes: Is it the journey or the destination that matters the most? *The Leadership Quarterly*, 29(1), 70–88. <https://doi.org/10.1016/j.leaqua.2017.11.004>
- Anantrasirichai, N., & Bull, D. (2022). Artificial intelligence in the creative industries: a review. *Artificial Intelligence Review*, 55(1), 589–656. <https://doi.org/10.1007/s10462-021-10039-7>
- Ayers, J. W., Poliak, A., Dredze, M., Leas, E. C., Zhu, Z., Kelley, J. B., Faix, D. J., Goodman, A. M., Longhurst, C. A., Hogarth, M., & Smith, D. M. (2023). Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. *JAMA Internal Medicine*, 183(6), 589. <https://doi.org/10.1001/jamainternmed.2023.1838>
- Briggs, J., & Kodnani, D. (2023). *Generative AI could raise global GDP by 7 percent*. Goldman Sachs. Retrieved on 10th January 2025 from: <https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html>
- Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company.
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117(3), 497–529. <https://doi.org/10.1037/0033-2909.117.3.497>
- Cannon-Bowers, J. A., Bowers, C. A., Carlson, C. E., Doherty, S. L., Evans, J., & Hall, J. (2023). Workplace coaching: a meta-analysis and recommendations for advancing the science of coaching. *Frontiers in Psychology*, 14, 1204166. <https://doi.org/10.3389/fpsyg.2023.1204166>
- Chung, K., & Park, R. C. (2019). Chatbot-based healthcare service with a knowledge base for cloud computing. *Cluster Computing*, 22(S1), 1925–1937. <https://doi.org/10.1007/s10586-018-2334-5>

Clutterbuck, D., & Megginson, D. (n.d.). *Further techniques for coaching and mentoring*. Butterworth-Heinemann.

De Haan, E., & Nilsson, V. O. (2023). What Can We Know about the Effectiveness of Coaching? A Meta-Analysis Based Only on Randomized Controlled Trials. *Academy of Management Learning & Education*, 22(4), 641–661.
<https://doi.org/10.5465/amle.2022.0107>

Dell'Acqua, F. et al (2024). Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality, Harvard Working Paper 24-013 Retrieved on 6th June 2024 from https://www.hbs.edu/ris/Publication%20Files/24-013_d9b45b68-9e74-42d6-a1c6-c72fb70c7282.pdf

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., ... Wright, R. (2023). Opinion Paper: “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>

*Ellis-Brush, K. (2021). *Augmenting Coaching Practice through digital methods*.
<https://doi.org/10.24384/ER2P-4857>

EMCC (2024) AI & Digital. Retrieved on 20th June 2024 from <https://www.emccglobal.org/leadership-development/ai-digital/>

EU (2024). *Artificial Intelligence AI Act*. Retrieved on 24 July 2024 from: <https://www.consilium.europa.eu/en/press/press-releases/2024/05/21/artificial-intelligence-ai-act-council-gives-final-green-light-to-the-first-worldwide-rules-on-ai/>

Ferrer, X., Van Nuenen, T., Such, J. M., Coté, M., & Criado, N. (2021). Bias and discrimination in AI: a cross-disciplinary perspective. *IEEE Technology and Society Magazine*, 40(2), 72-80.

*Figueroa, C. A., Luo, T. C., Jacobo, A., Munoz, A., Manuel, M., Chan, D., Canny, J., & Aguilera, A. (2021). Conversational Physical Activity Coaches for Spanish and English Speaking Women: A User Design Study. *Frontiers in Digital Health*, 3, 747153.
<https://doi.org/10.3389/fdgth.2021.747153>

Giattino, C., Mathieu, E., Samborska, V., & Roser, M. (2023). Data Page: Parameters in notable artificial intelligence systems. In C. Giattino, E. Mathieu, V. Samborska, & M. Roser (Eds.), *Artificial Intelligence*. Data adapted from Epoch. Retrieved from <https://ourworldindata.org/grapher/artificial-intelligence-parameter-count>

Graßmann, C., & Schermuly, C. C. (2021). Coaching with Artificial Intelligence: Concepts and Capabilities. *Human Resource Development Review*, 20(1), 106–126. <https://doi.org/10.1177/1534484320982891>

Graßmann, C., Schölmerich, F., & Schermuly, C. C. (2020). The relationship between working alliance and client outcomes in coaching: A meta-analysis. *Human Relations*, 73(1), 35–58. <https://doi.org/10.1177/0018726718819725>

*Hassoon, A., Baig, Y., Naiman, D. Q., Celentano, D. D., Lansey, D., Stearns, V., Coresh, J., Schrack, J., Martin, S. S., Yeh, H.-C., Zeilberger, H., & Appel, L. J. (2021). Randomized trial of two artificial intelligence coaching interventions to increase physical activity in cancer survivors. *Npj Digital Medicine*, 4(1), 168. <https://doi.org/10.1038/s41746-021-00539-9>

ICF (2024). Technology and Artificial Intelligence in coaching. Retrieved on 19th June 2024 from <https://coachingfederation.org/about/coalition-technology-in-coaching>

Jones, R. J., Woods, S. A., & Guillaume, Y. R. F. (2016). The effectiveness of workplace coaching: A meta-analysis of learning and performance outcomes from coaching. *Journal of Occupational and Organizational Psychology*, 89(2), 249–277. <https://doi.org/10.1111/joop.12119>

Khan, A. A. et al (2022). Ethics of AI: A systematic literature review of principles and challenges. *EASE '22: Proceedings of the 26th International Conference on Evaluation and Assessment in Software Engineering*, June 2022, 383–392 <https://doi.org/10.1145/3530019.3531329>

*Kannampallil, T., Ronneberg, C. R., Wittels, N. E., Kumar, V., Lv, N., Smyth, J. M., Gerber, B. S., Kringle, E. A., Johnson, J. A., Yu, P., Steinman, L. E., Ajilore, O. A., & Ma, J. (2022). Design and Formative Evaluation of a Virtual Voice-Based Coach for Problem-solving Treatment: Observational Study. *JMIR Formative Research*, 6(8), e38092. <https://doi.org/10.2196/38092>

Lee, D., & Yoon, S. N. (2021). Application of Artificial Intelligence-Based Technologies in the Healthcare Industry: Opportunities and Challenges. *International Journal of Environmental Research and Public Health*, 18(1), 271. <https://doi.org/10.3390/ijerph18010271>

Leong, K., & Sung, A. (2024). Gender stereotypes in artificial intelligence within the accounting profession using large language models. *Humanities and Social Sciences Communications*, 11(1), 1-11.

Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., He, H., Li, A., He, M., Liu, Z., Wu, Z., Zhao, L., Zhu, D., Li, X., Qiang, N., Shen, D., Liu, T., & Ge, B. (2023). Summary of

ChatGPT-Related research and perspective towards the future of large language models. *Meta-Radiology*, 1(2), 100017. <https://doi.org/10.1016/j.metrad.2023.100017>

*Mai, V., Neef, C., & Richert, A. (2022). "Clicking vs. Writing"—The Impact of a Chatbot's Interaction Method on the Working Alliance in AI-based Coaching. *Coaching / Theorie & Praxis*, 8(1), 15–31. <https://doi.org/10.1365/s40896-021-00063-3>

*Mai, V., Wolff, A., Richert, A., & Preusser, I. (2021). Accompanying Reflection Processes by an AI-Based StudiCoachBot: A Study on Rapport Building in Human-Machine Coaching Using Self Disclosure. In C. Stephanidis, D. Harris, W.-C. Li, D. D. Schmorow, C. M. Fidopiastis, M. Antona, Q. Gao, J. Zhou, P. Zaphiris, A. Ioannou, R. A. Sottilare, J. Schwarz, & M. Rauterberg (Eds.), *HCI International 2021 - Late Breaking Papers: Cognition, Inclusion, Learning, and Culture* (Vol. 13096, pp. 439–457). Springer International Publishing. https://doi.org/10.1007/978-3-030-90328-2_29

McKenzie, L., Pavese, L. Passmore, J. Tee, D. & Macrae, L. (2022). *The CoachHub Global Survey: Business Trends in Coaching 2023*. Berlin: CoachHub. Retrieved on 24th May 2024 from: <https://www.coachhub.com/app/uploads/2022/09/Coaching-Trends23-Report.pdf>

Movsumova, E., Alexandrov, V., Rudenko, L., Aizen, V., Sidelnikova, S., & Voytko, M. (2020). Consciousness: effect of coaching process and specifics through AI usage. *E-Mentor*, 86(4), 79–86. <https://doi.org/10.15219/em86.1485>

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*, 151(4), 264-269. <https://doi.org/10.7326/0003-4819-151-4-200908180> 00135

Nieuwerburgh, C. van (Ed.). (2016). *Coaching in professional contexts*. Sage.

Olafsson, B., Martin, D., & Gonzalez-Bonorino, A. (2024). Artificial Intelligence Coaching for Wellbeing. In J. Passmore, S. Diller, S. Isaaccson, M. Brantl (eds.), *The Health and Wellbeing Coaches Handbook*. Abingdon: Routledge.

Passmore, J. (2024). *Getting better all the time: Exploring AI & Human performance, outcomes and ethics*. Association of Business Psychologists Annual Conference, London, 20th June 2024

Passmore, J., & Fillery-Travis, A. (2011). A critical review of executive coaching research: A decade of progress and what's to come. *Coaching: An International Journal of Theory, Research and Practice*, 4(2), 70–88. <https://doi.org/10.1080/17521882.2011.596484>

*Passmore, J., Liu, Q., & Tewald, S. (2021). Future trends in coaching: Results from a global coach survey, 2021. *Coaching Psychologist*, 17(2).

Passmore, J., & Tee, D. (2023a). Can Chatbots replace human coaches? Issues and dilemmas for the coaching profession, coaching clients and for organisations. *The Coaching Psychologist*, 19(1), 47–54. <https://doi.org/10.53841/bpstcp.2023.19.1.47>

*Passmore, J. & Tee, D. (2023b). The Library of Babel: Assessing the powers of Artificial Intelligence in coaching conversations and knowledge synthesis. *Journal of Work Applied Management*, 16(1), 4-18. [10.1108/JWAM-06-2023-0057](https://doi.org/10.1108/JWAM-06-2023-0057)

Passmore, J., & Woodward, W. (2023). Coaching education: Wake up to the new digital and AI coaching revolution! *International Coaching Psychology Review*, 18(1), 58–72. <https://doi.org/10.53841/bpsicpr.2023.18.1.58>

Passmore, J., Olafsson, B. Chitkara, R. & Wilson, T. (2024) Hello my name in Sky: Ethical principles in for high trust conversations with AI bots. Is it time to treat workplace coaching, training and mentoring tools like humans? London: Working Paper

Russell, S. J., & Norvig, P. (2021). *Artificial intelligence: a modern approach* (Fourth edition). Pearson.

Sasaki, K., Ihaya, K., & Yamada, Y. (2017). Avoidance of novelty contributes to the uncanny valley. *Frontiers in Psychology*, 8, 1792. <https://doi.org/10.3389/fpsyg.2017.01792>

Serriyeh, R., Lawton, R., Gardener, P., & Armitage, G. (2012). Reviewing studies with diverse designs: the development and evaluation of a new tool. *Journal of Evaluation in Clinical Practice*, 18(4), 746-752. <https://doi.org/10.1111/j.1365-2753.2011.01662.x>

Smith, R & Smoll, F. (2017). Coaching Behavior and Effectiveness in Sport and Exercise Psychology. In *Oxford Research Encyclopaedia* (Online). <https://oxfordre.com/psychology/display/10.1093/acrefore/9780190236557.001.0001/acrefore-9780190236557-e-188>

*Stephens, T. N., Joerin, A., Rauws, M., & Werk, L. N. (2019). Feasibility of pediatric obesity and prediabetes treatment support through Tess, the AI behavioral coaching chatbot. *Translational Behavioral Medicine*, 9(3), 440–447. <https://doi.org/10.1093/tbm/ibz043>

Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., ... & Higgins, J. P. (2016). ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ (Clinical Research ed)*, 355:i4919. <https://doi.org/10.1136/bmj.i4919>

Suler, J. (2004). The Online Disinhibition Effect. *CyberPsychology & Behavior*, 7(3), 321–326. <https://doi.org/10.1089/1094931041291295>

Terblanche, N. (2020). *A design framework to create Artificial Intelligence Coaches*.
<https://doi.org/10.24384/B7GS-3H05>

*Terblanche, N., & Cilliers, D. (2020). Factors that influence users' adoption of being coached by an Artificial Intelligence Coach. *Philosophy of Coaching: An International Journal*, 5(1), 61–70. <https://doi.org/10.22316/poc/05.1.06>

*Terblanche, N. H. D., Van Heerden, M., & Hunt, R. (2024). The influence of an artificial intelligence chatbot coach assistant on the human coach-client working alliance. *Coaching: An International Journal of Theory, Research and Practice*, 1–18.
<https://doi.org/10.1080/17521882.2024.2304792>

*Terblanche, N., Molyn, J., De Haan, E., & Nilsson, V. O. (2022a). *Coaching at Scale: Investigating the Efficacy of Artificial Intelligence Coaching*. *International Journal of Evidence Based Coaching and Mentoring*, 20(2), 20-36.
<https://doi.org/10.24384/5CGF-AB69>

*Terblanche, N., Molyn, J., De Haan, E., & Nilsson, V. O. (2022b). Comparing artificial intelligence and human coaching goal attainment efficacy. *PLOS ONE*, 17(6), e0270255. <https://doi.org/10.1371/journal.pone.0270255>

*Terblanche, N., Molyn, J., Williams, K., & Maritz, J. (2023). Performance matters: students' perceptions of Artificial Intelligence Coach adoption factors. *Coaching: An International Journal of Theory, Research and Practice*, 16(1), 100–114.
<https://doi.org/10.1080/17521882.2022.2094278>

Theeboom, T., Beersma, B., & Van Vianen, A. E. M. (2014). Does coaching work? A meta-analysis on the effects of coaching on individual level outcomes in an organizational context. *The Journal of Positive Psychology*, 9(1), 1–18.
<https://doi.org/10.1080/17439760.2013.837499>

*Tropea, P., Schlieter, H., Sterpi, I., Judica, E., Gand, K., Caprino, M., Gabilondo, I., Gomez-Esteban, J. C., Busnatu, S., Sinescu, C., Kyriazakos, S., Anwar, S., & Corbo, M. (2019). Rehabilitation, the Great Absentee of Virtual Coaching in Medical Care: Scoping Review. *Journal of Medical Internet Research*, 21(10), e12805.
<https://doi.org/10.2196/12805>

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need*.
<https://doi.org/10.48550/ARXIV.1706.03762>

Wolever RQ, Simmons LA, Sforzo GA, et al. A Systematic Review of the Literature on Health and Wellness Coaching: Defining a Key Behavioral Intervention in Healthcare. *Global Advances in Health and Medicine*. 2013;2(4):38-57. doi:[10.7453/gahmj.2013.042](https://doi.org/10.7453/gahmj.2013.042)

Note: References marked by * were included in the systematic literature review