University of Reading

Can Chatbots like GPT-4 replace human coaches: Issues and dilemmas for the coaching profession, coaching clients and for organisations

Article

Accepted Version

Passmore, J. ORCID: https://orcid.org/0000-0003-0832-7510 and Tee, D. (2023) Can Chatbots like GPT-4 replace human coaches: Issues and dilemmas for the coaching profession, coaching clients and for organisations. The Coaching Psychologist, 19 (1). pp. 47-54. ISSN 2396-9628 doi: 10.53841/bpstcp.2023.19.1.47 Available at https://centaur.reading.ac.uk/112848/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

Identification Number/DOI: 10.53841/bpstcp.2023.19.1.47
<https://doi.org/10.53841/bpstcp.2023.19.1.47>

Publisher: The British Psychological Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in

the End User Agreement.

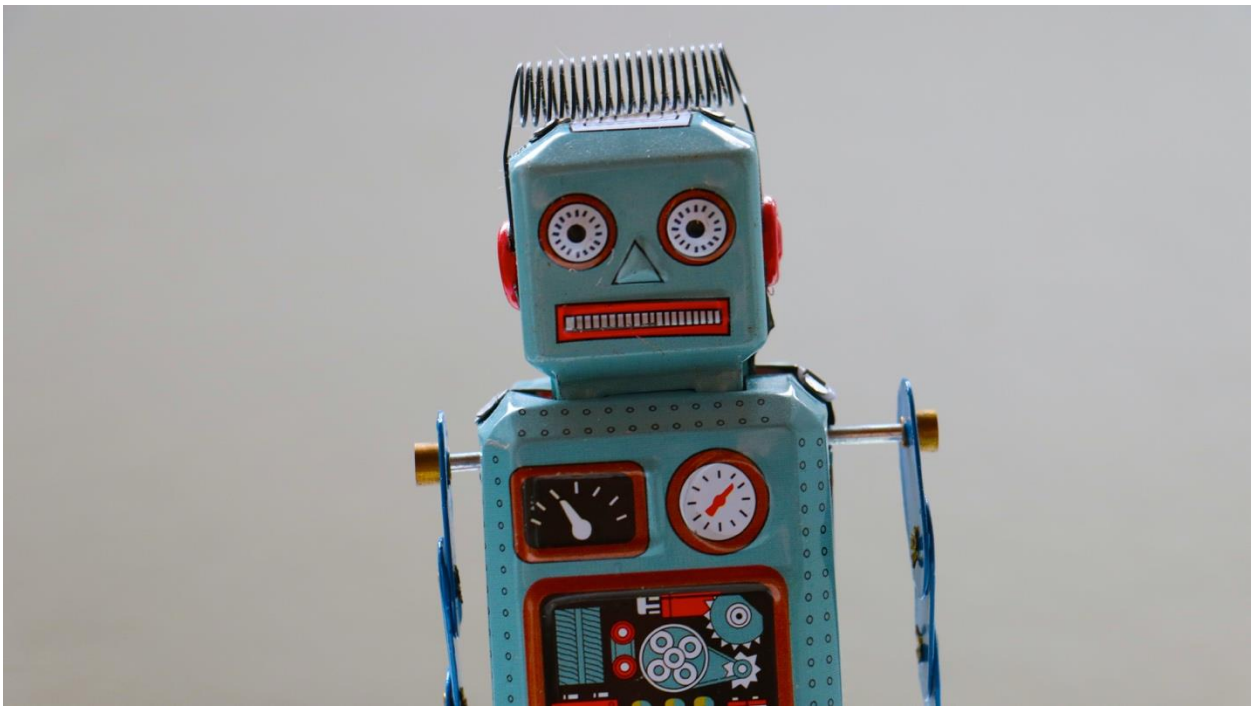
www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Can Chatbots like GPT-4 replace human coaches: Issues and dilemmas for the coaching profession, coaching clients and for organisations



Jonathan Passmore & David Tee

Citation: Passmore, J. & Tee, D. (2023) Can Chatbots like GPT-4 replace human coaches: Issues and dilemmas for the coaching profession, coaching clients and for organisations, *The Coaching Psychologist*. *19(1), 47-54.* DOI: 10.53841/bpstcp.2023.19.1.47

Key words: *Coaching chatbot; AI coaching definition; GPT-4; artificial intelligence; AI coaching; human coaches.*

Note: This text is made available by the authors as part of their commitment to knowledge sharing. This version contains the original pre-publication version of the paper, as a result page numbers and layout may vary from the published version.

Abstract

The advent of artificial intelligence (AI) and machine learning (ML) has led to the speculation that chatbots could revolutionise the coaching industry in the coming decade, replacing humans as the main provider of coaching conversations. The development of GPT4 has led to these bots becoming increasingly sophisticated and effective at providing support and guidance in various fields. Coaching providers have been quick to operationalise these generative language tools to create new products like AIMY, evoach and Vici. This paper examines the potential of AI chatbots and their integration into coaching tools. It will review the advantages and current limitations of AI coaching chatbots and offer a preliminary definition for the field, seeking to differentiate chatbots from human coaching. The paper also explores the role of coaching psychology, professional bodies and governments in the development and evolution of AI systems and coaching chatbots, and suggests the urgent need for action to protect clients and organisations from unregulated and unethical practices.

Introduction

Coaching psychology has emerged as a central discipline that seeks to understand the processes and mechanisms underlying coaching practice and effectiveness (Grief et al., 2022). As technology advances, AI-driven systems such as chatbots are increasingly being used in various fields including coaching, raising questions about their potential to replace human coaches (Khandelwal & Upadhyay, 2021). Given the pace of change, the lack of under- standing of AI and the complexity of algorithms, it is hard for managers, organisations, professional bodies and national governments to put in place the regulation and controls needed to manage these emerging technological applications.

GPT-4 launched in March 2023 and, like its predecessors developed by OpenAI, it uses natural language processing (NLP). Early reviews have demonstrated NLP's remarkable power in generating coherent and contextually relevant text, as well as occurrences of falsification of data and practices not aligned to coaching. This article explores the potential of chatbots using natural language processing to replace human coaches and examines the advantages and limitations of AI Chatbot-delivered coaching. It goes on to discusses the role of coaching and psychological science in the development and evaluation of these machines and the actions needed by professional bodies and governments to regulate the field to protect clients and organisations.

What is Artificial Intelligence?

The term Artificial intelligence (AI) may have its origins in classical works such as the Greek myth of Talos, a giant bronze automaton. The term in its present form emerges from the work of Alan Turin and that of Warren McCulloch and Walter Pitts, 1940s US cyberneticians interested in developing Turin's theories into artificial operating 'brains'.

Over the subsequent eighty years, machines have increased in their ability to perform cognitive functions, including learning, interacting, problem-solving and displaying creativity (Rai et al., 2019). Brynjolfsson and McAfee (2014) have suggest that AI are reaching a stage of development where their ability to improve performance no longer relies on humans to explain how to accomplish the specific task.

Writers have now delineated two types of AI: narrow AI and general AI, which differ in the number of tasks that they can perform (Schlegel & Uenal, 2021). Narrow AI refers to systems that perform one or a few specific tasks. Narrow examples include a patient cancer detection system or a system capable of producing written text in various forms. These narrow applications of AI have witnessed rapid progress in the last decade, culminating with GPT-4. In contrast, general AI, which is intended to perform most activities that humans can do, has made almost no progress.

What is AI coaching?

AI coaching has come to refer to the application of artificial intelligence methodologies to the coaching environment and we would argue is different from digital coaching, which may be defined as, *'Synchronous coaching with a human coach via digital technology',* AI (Chatbot delivered) coaching is *'Synchronous or asynchronous coaching using AI or a computer as a coach instead of a human coach'* (Diller & Passmore, In Press).

Definition of AI Coaching

Synchronous or asynchronous coaching using AI or a computer as a coach instead of a human coach'

In reality while the focus of thinking has tended to centre on the role of the AI chatbot as a medium for coaching delivery, the coaching environment is diverse and there are many processes where AI can add value (Table 1).

Table 1: Possible AI applications in coaching.

I	Assessment of client readiness for coaching
II	Coach - client matching
III	Recommendation of inter-sessional coaching assignments (sometimes called homework)
IV	Supporting goal attainment between sessions through client nudges
V	AI coach supervisor (asynchronous coach supervision based on session transcript)

VI	AI coach advice (synchronous advice to the coach during the session from the AI bot)
VII	AI chatbot delivered coaching (machine replacement of the human coach)

Coaches may have different responses to each of these scenarios (applications), with a different level of perceived threat to their livelihoods from each of them.

Very little research has been undertaken on the majority of these applications. To date, only one paper has compared the impact of AI delivered coaching with human coaching (Terblanche et al., 2022). However, it would be fairer to describe this as 'proof of concept' research, as the coaching delivery for the two participants groups (AI delivered and human delivered) was undertaken at different times, clients were drawn from different samples (leaders and students) and participants (clients) were not randomly assigned. However, the study does suggest that AI Chatbot-delivered coaching has potential to deliver positive impact.

Further, AI Chatbot-delivered coaching has the potential to quickly scale, learning from a large dataset of human-to-human coaching processes, and some coaching providers have already been recording coaching conversations and transcribing the content with a view towards using this data for future machine learning, while others have created commercial AI Chatbots which offer coaching (CoachHub, 2023).

With a sufficiently large dataset, say of a 100,000 or 1,000,000 coaching conversations, the bot could have access to more data than the most highly experienced coach with 40 years of coaching experience. But operationalising this data also requires a system capable of turning data into a conversational response which can be sustained for a prolonged period. The recording and use of such data also raises ethical concerns for most coaches, alongside clients and organisational coaching sponsors.

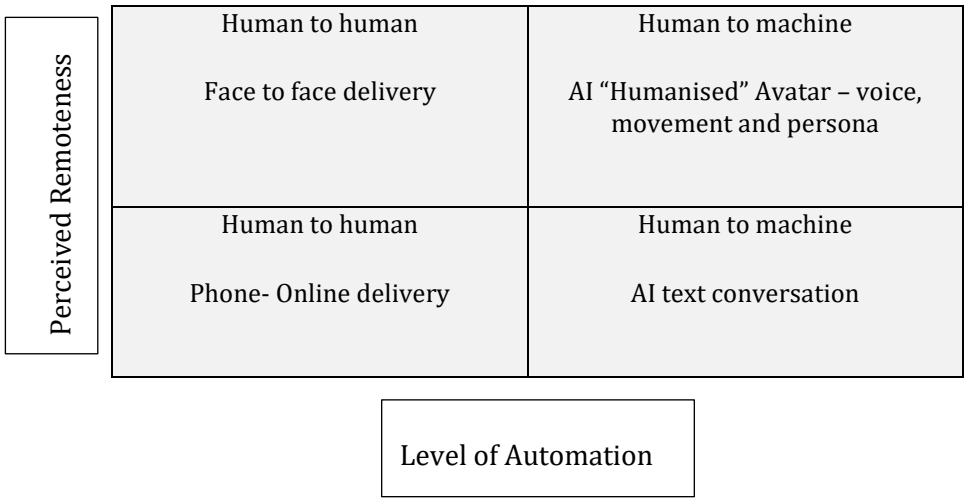
The shifts in the coaching industry from a human-to-human face-to-face model, has seen what we describe as a pluralisation of coaching, with new forms of delivery emerging. This was initially through phone coaching (audio) and what has been a gradual and continual growth of online delivery (audio and video) using Zoom-like technology and the popularisation of digital coaching platforms. These technologies have enabled the coach to be physically remote while delivering a synchronous coaching session.

The emergence of AI technologies now offers the opportunity for further pluralisation. Options include text-based tools, which involve the client interacting via their keyboard with the AI application, to humanised coach avatars with movement, voice and a persona.

The first step in this pluralisation process was the advent of geographically removed coaching, initially using mobile (or cell) phones (audio), moving to small scale online

(audio and video) and a position in 2023 where the majority of coaching conversations are now taking place online.

Figure 1: Pluralisation of Coaching: Remoteness and Automation.



More recently there has been a movement towards automation. This process started with text-based chat bots such as Vici (Terblanche et al., 2022) and evoach (Mai & Rutschmann, 2023), and has progressed to the incorporation of voice recognition software to create conversational bots such as AIMY (CoachHub, 2023). Future steps may include avatar bots, which offer voice and movement and potentially persona. It is only a short step to imagine a provider offering a coaching conversation with a fictional character or a client's hero, such as Nelson Mandela, Anita Roddick, Ada Lovelace or Mahatma Gandhi, and the app, using software such as Deepswap and Fakeyou and drawing on historic movie and voice recordings, creating the sensation of the bot sounding and appearing as if the client is engaging with the character selected.

These movements to remoteness and to automation have been viewed by coaches as less desirable, with coaches seeing the online experience as attractive for its convenience, but ultimately less enjoyable, and feeling threatened by the emergence of coaching chatbots.

However, little research has explored client perspectives, and at what price points automation starts to appear a more attractive option over a human coach. Is this at 50 per cent of the cost of a human, 10 per cent or even zero cost? Funding models in other automated spheres, such as knowledge search

(Google), client-created content (YouTube) and music (Spotify) offer zero cost options to consumers, instead using advertising and personal data to fund their services. Coaching could potentially have a similar future, with free at the point of use coaching, and the service funded based on targeted advertising from the information collected during the session or in pre-session assessments.

Graßmann and Schermuly (2021) argue that the entire coaching process is open to including artificial intelligence capabilities, and the main barrier to the complete replacement of face-to-face and online delivery modes of coaching is the development of technologies. However, to paraphrase Jeff Goldblum's character from 1993s 'Jurassic Park', just because we can now create AI coaches, it doesn't excuse us not stopping to think if we should. These are questions worthy of discussion. Specifically, what are the advantages of AI chatbot-delivered coaching, what are its risks and how can the ethical issues be managed in a sector with no regulatory constraints?

Advantages of AI Chatbot-delivered coaching

One of the primary advantages of AI Chatbot-delivered coaching is its scalability, as it can be deployed to large numbers of users simultaneously, unlike human coaches where, even when provided through a digital coaching platform, capacity may be limited to a specific number of clients. This scalability provides the opportunity for coaching to be available for those unable to access coaching due to their remote geographical location.

AI Chatbot-delivered coaching can be more cost-effective than human coaches, as the coaches do not require ongoing renumeration and can operate 24 hours a day without breaks, refreshment, supervision, sick leave or holidays. Scalability and minimal costs each make coaching services more accessible to individuals who do not have the economic power to access human coaches.

Thirdly, tools like GPT-4 can provide consistent and objective feedback unencumbered by emotional states, fatigue, prejudices and biases or other factors that can impact on human coach performance. Their programming can also mean consistent adherence to coaching maxims or principles: AI chat- bots programmed to 'ask – don't tell', for example, will dependably observe this principle in a way that human coaches may not.

A fourth factor is that the AI Chatbot-delivered coaching can leverage vast amounts of data to provide personalised evidence-based insights, making them 1000 or 10,000 times more experienced than the most practised coaching psychologist. This data driven approach can thus enhance the effectiveness of coaching by identifying patterns and trends that may not be apparent to human coaches who may both forget or have a limited data sample from which to draw.

Limitations of AI based coaching systems

One of the most significant limitations of AI coachbot-delivered coaching is its current inability to demonstrate empathy and emotional intelligence or operate with a sense of humour at the level of most humans. Skills such as emotional under- standing, the building of rapport and an ability to adapt to the client's emotional state – considered crucial skills for human coaches – are presently beyond the capabilities of AI Chatbot-delivered coaching. However, over time – and maybe much sooner than we can anticipate at the time of writing – AI chatbots will become more sophisticated, possibly to the point where the client cannot tell the difference between the responses of the bot and that of an empathetic human.

Secondly, although AI Chatbot delivered coaching has made significant strides, with tools such as Vici, evoach and AIMY, these apps still struggle with adaptability, creativity and playfulness, particularly when compared to human coaches. Human coaches can

draw from a wealth of personal experience, intuition and creativity to develop innovative solutions for clients, while AI chatbots are constrained by the data and algorithms on which they have been built.

Thirdly, the use of AI delivered services can lead to a potentially diminished human connection. As this spreads through service functions, from automated goods ordering, to in store service and into HRD, the needs for connection as a social species may make human coaching even more important.

Finally, there are the ethical concerns. These include data privacy, data security, potential biases in AI algorithms, as well as AI Chatbot's possible focus on goal attainment over wider desirable goals, such as harm to others, or harm to self. The evidence from other sectors in terms of data privacy suggests this should be a legitimate concern, both for individuals as well as organisations. Few organisations would wish their executives or employees to be discussing confidential organisational issues with a chatbot coach in the knowledge that data is being recorded and categorised. At an individual level, Marcus (2023) details one already known instance where GPT-3 urged a user to end their life. Further advances in AIs might make negative consequences more likely rather than less likely, depending on the goals set for it to achieve, creating unintended consequences.

Ethical considerations

All new technologies bring with them ethical considerations. We have set out some of these in Table 2. The list is not exhaustive but is an illustration of the types of issues which AI Chatbot providers, as well as professional bodies and governments, need to consider as these technologies move from laboratory to test environments and from test environments to commercial offers.

We suggest professional bodies have an important role to play in setting standards and seeking to secure engagement with commercial providers. This may involve assessing AI bots for compliance and certifying the coaching at a specific level, such as using the ICF framework to assess if the coaching meets ACC, PCC or MCC standards, and observing compliance with ethical codes of conduct. Professional bodies could also require confirmation about confidentiality levels and other standard coach practices. This could involve professional bodies offering certification of products, thus providing greater reassurance to buyers and users.

Table 2: AI Chatbot delivered Coaching: Ethical dilemmas.

Theme	Ethical dilemma
Accountability	Who is responsible for the outcomes of the coaching session if the AI app's input is harmful or its 'advice' results in negative consequences?
Bias	How is unintended bias monitored and managed?
Cross cultural	How do AI chatbots respond sensitively to gender, racial, faith and identity issues?
Data Privacy	How is data collected and stored?
Digital divide	How can greater equity be achieved when AI services may be limited by economic power and/or access to technology such as a smart phone
Empathy	How does AI respond in emotionally charged situations or when clients experience distress as a result of the conversation?
Faking	Is it ethical for the AI tool to use the voice, image or persona of a living or dead person without their express consent?
Informed Consent	How are clients informed they are interacting with a machine and have the choice whether to do so? How is consent obtained to hold client data? How and when can it be withdrawn? What happens to data when clients withdraw their consent?
Job displacement	How should AI be applied if this will lead to a reduction in opportunities for coaches?
Over reliance	How can AI be limited or restricted to prevent overuse or over reliance?
Regulation	How should the coaching industry seek to regulate standards by providers?
Security	How is access to data restricted?
Trends	How do AI chatbots stay up to date with changing (daily) trends and patterns? From which sources do they draw this data?

The role of coaching psychologists and coaches

Coaching psychologists arguably have a critical role in the development and evaluation of AI Chatbot coaching delivery. Firstly, coaching psychologists should be actively involved in the development of these systems, bringing the science of coaching psychology, psychological theory, coaching models and ethical practice to the design.

Secondly, coaching psychology should be involved in the ongoing evaluation of AI chatbot-delivered coaching and wider AI coaching applications. This may include

conducting randomised controlled trials, assessing client satisfaction and examining the impact of these modes of coaching on various outcomes from performance-related goal attainment to enhanced wellbeing or self-efficacy.

Thirdly, the wider evidence from AI developers is that such systems have in the past been heavily biased, reflecting the gender, ethnicities, age and nationalities of those involved: too frequently, white, under 35-year old, American men. However, in a world of global delivery with diverse communities, any tool needs to reflects the diverse needs and perspectives of its users. This means design teams also need to address concerns regarding intersectionality and have members drawn from various socio-economic groupings, races, genders and other important categories, and to test their tools across multiple territories and users.

Conclusions

While tools such as GPT-4 have driven coaching forward in 2023, offering even greater scalability, cost reduction possibilities and enhanced data and insights, the limitations – including the current lack of empathy, emotional intelligence and humour along with ethical concerns – mean the coaching industry needs to act. Coaching psychologists, along with coaches, can complain about the changing industry from the side lines or they can participate in the design and delivery of new applications, with a focus on inclusion and ethical practice. What is certain is that the coaching industry is moving from a craft-based delivery model to a pluralist model with multiple means by which coaching interventions can be delivered, harnessing automation and physical remoteness to offer even more accessible and affordable access to coaching.

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. doi:10.1016/j.leaqua.2017.11.004

Brynjolfsson, E. & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies.* New York: Norton & Co.

CoachHub (2023) AIMY. Retrieved 1 Aril 2023 from https://www.coachhub.com/aimy/

de Haan, E. (2019). A systematic review of qualitative studies in workplace and executive coaching: The emergence of a body of research. *Consulting Psychology Journal: Practice and Research, 71*(4), 227–248. doi:10.1037/cpb0000144

Diller, S. & Passmore, J. (in press). Conceptualizing the distinctiveness of digital coaching. *Frontiers in Psychology.*

Graßmann, C. & Schermuly, C.C. (2021). Coaching with artificial intelligence: Concepts and capa- bilities. *Human Resource Development Review, 20*(1), 106–126. doi:10.1177/1534484320982891

Greif, S., Moller, H., Scholls, W., Passmore, J & Muller, F. (2022). *International Handbook of Evidenced-base Coaching: Theory, Research and Practice.* Heidel- burgh, Germany: Springer Nature.

Grover & Furnham (2019). Coaching as a Developmental Intervention in Organisa- tions: A Systematic Review of Its Effectiveness and the Mechanisms Underlying It. *PLOS One.* doi:10.1371/journal.pone.015913

Khandelwal, K. & Upadhyay, A.K. (2021). The advent of artificial intelligence-based coaching. *Stra- tegic HR Review, 20*(4), 137–140. doi:10.1108/ SHR-03-2021-0013

Mai, V. & Rutschmann, R. (2023). Best Practices im Chatbot Coaching. Einblicke in Forschung und Entwicklung des StudiCoachBots der TH Köln und in die Coaching Chatbot Plattform evoach. *Organisationsberatung, Supervision, Coaching, 30*(1), 111–125. doi:10.1007/s11613-022-00802-2

Marcus, G. (2023). Mind Your Language. *The Wired World in 2023, 14.*

Park, S-Y.C., Kim, H. & Lee, S. (2020) Do Less Teaching, Do More Coaching: Toward Critical Thinking for Ethical Applications of Artificial Intelligence. *Journal of Learning and teaching in a digital age. 6*(2) ,97–100. https://dergipark.org. tr/en/pub/joltida/issue/63505/961435

Passmore, J. (2023, June). The impact of digital on the coaching industry. Paper presented at BPS Division of Coaching Psychology Conference Keynote, London, UK.

Rai, A., Constantinides, P. & Sarker, S. (2019). Next-generation digital platforms: Toward human-AI hybrids. *Management Information Systems Quarterly, 43*(1), iii–ix. https://misq.org/ misq/downloads/

Schlegel, D. & Uenal, Y. (2021). A Perceived Risk Perspective on Narrow Artificial Intelligence. *PACIS 21 Proceedings.* 44. https://aisel.aisnet.org/ pacis2021/44

Terblanche, N. (2020). A design framework to create artificial intelligence coaches. *International Journal of Evidence Based Coaching and Mentoring, 18*(2), 152–165. doi: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. http://dx.doi.org/10.22316/poc/05.1.06

Terblanche, N., Molyn, J., de Haan, E. & Nilsson, V. (2022). Coaching at Scale: Investigating the Efficacy of Artificial Intelligence Coaching. *Inter- national Journal of Evidence Based Coaching and Mentoring 20*(2), 20–36. doi:10.24384/5cgf-ab69