



The Coaching Uncanny

What AI companions teach us about real executive coaching

— DR VICTORIA HOPKIN —

Leave the Room

In 1966 a computer scientist at MIT named Joseph Weizenbaum finished a program he had built as a demonstration of how little a machine needed to understand in order to hold a conversation. He called it ELIZA. It worked by spotting keywords in whatever the user typed and turning them back as questions. Type "my mother hates me" and it would answer "Tell me more about your family." Type "I am unhappy" and it would ask "Do you think coming here will help you not to be unhappy?" It understood nothing. It had a list of patterns and a list of replies.

Weizenbaum had chosen, half as a joke, to give it the manner of a particular kind of psychotherapist: the non-directive, client-centred style that Carl Rogers had made famous, in which the therapist reflects the patient's words back and offers almost nothing of his own. He chose it because it was the one conversational role in which saying nothing of substance was not a fault. A Rogerian therapist who answered "I see, go on" to everything was doing his job. So was ELIZA.

Then something happened that Weizenbaum spent the rest of his life trying to explain. His secretary, who had watched him build the program over many months and knew exactly what it was, sat down to try it. After a few exchanges she asked him to leave the room. She wanted to talk to it in private.

"The book is mostly remembered for its argument. The detail that outlasts the argument is the secretary asking him to leave."

She was not alone. People who knew ELIZA was a trick told it things they had told no one. A psychiatrist named Kenneth Colby proposed, in print, that a version of it could treat hundreds of patients an hour in hospitals that could not afford therapists. Weizenbaum, who had built the thing to show that machines were shallow, found that the shallowness did not matter to the people typing. He was appalled, and in 1976 he wrote a book, *Computer Power and Human Reason*, to say so. The book is mostly remembered for its argument. The detail that outlasts the argument is the secretary asking him to leave.

The Telling

Sixty years on, the secretary's request has become an industry, and the coaching profession is trying to work out what it means for them.

The numbers first, because the popular ones are wrong. It is widely repeated, on the strength of a Harvard Business Review article, that therapy and companionship are the most common thing people use chatbots for. OpenAI's own researchers, with economists from Harvard and Duke, examined a sample of 1.6 million conversations in 2025 and found that about 2 per cent concerned relationships or personal reflection. Anthropic's analysis of its own traffic put affective use at about 3 per cent. Small shares. But 2 per cent of the several hundred million people who now use these tools every week is more people seeking counsel from a machine than have ever sat in a coaching session, and it is happening at three in the morning, for free, to something that never looks at the clock.

SHARE OF CHATBOT CONVERSATIONS THAT ARE PERSONAL OR RELATIONAL

■ about 2–3 per cent

...of several hundred million weekly users — **more people seeking counsel from a machine than have ever sat in a coaching session.**

Fig. 1 — The popular numbers are wrong in both directions: the share is small (Chatterji et al., 2025; Anthropic, 2025), and the absolute scale is without precedent.

Why does it work at all? The answer comes from Jeffrey Hancock at Stanford, who in 2018 ran an experiment with Annabell Ho and Adam Miner in which people disclosed something troubling either to a chatbot or to a person. Afterwards he measured how they felt: the emotional relief, the sense of having been understood, the shift in how they saw the problem. The two groups did not differ. Disclosing to the machine produced the same benefits as disclosing to a human. The benefit, it turned out, was in the telling, not in the listener. Weizenbaum's secretary was not deceived. She was getting something real from a machine that gave nothing back, because what she needed was to say it.

And the machine has one advantage over the human that Gale Lucas at the University of Southern California measured directly in 2014. People told a virtual interviewer more, and more honestly, when they believed no human was on the other end. The absence of judgement is not a feature of the software. It is a belief about who is listening, and it is exactly the belief that a professional coach, however skilled, cannot offer a client, because the coach is a person and will remember.

The Replication

Now the coaching research, which is small enough that a reader can hold all of it, and which has just produced its most important result.

Nicky Terblanche is a professor at Stellenbosch Business School in South Africa who spent the late 2010s building an AI coach, a chatbot that walked clients through a standard goal-setting conversation. In 2022, with Erik de Haan at Ashridge and two others, he published the first serious comparison. Over several months, clients coached by the chatbot reached their goals at rates that did not differ significantly from clients coached by humans. For three years that was the headline finding in the field, and vendors quoted it freely. The machine coaches as well as the coach.

Then the same team did what good researchers do with their own headline. They ran the trial properly. In a randomised controlled study published this year, de Haan, Terblanche and Kenneth Nowack assigned 114 coachees either to accredited human coaches or to an AI coach, and measured goal attainment, wellbeing, self-efficacy and, crucially, the working alliance, the bond between coach and client that seventy years of psychotherapy research has identified as the strongest single predictor of whether help helps. The human coaches won. Alliance was higher, and so were the outcomes. The 2022 finding did not replicate under controlled conditions.

2022 Goal attainment with the chatbot **did not differ significantly** from human coaching. Vendors quoted it freely.

2026 The same team, a randomised controlled trial, 114 coachees. **The human coaches won** — higher alliance, better outcomes. The 2022 finding did not replicate.

Fig. 2 — The field's most important result: the same researchers testing their own headline under controlled conditions (Terblanche et al., 2022; de Haan, Terblanche & Nowack, 2026).

What is most useful in the 2026 paper is not the result but the explanation. De Haan and his colleagues propose that what the human coach provides, and the machine cannot, is co-regulation: two nervous systems in the same room, one of them steady, the other borrowing that steadiness in real time. The client's state is changed by the coach's presence, not by the coach's words, and the change accumulates over sessions. A chatbot produces words. It has no state to lend.

But the field is not tidy, and two other studies complicate the story in ways the article cannot skip. Andrea Barger, in 2025, ran a Wizard of Oz trial: clients had a single coaching session with what they were told was an AI avatar, while in fact a human coach was operating it. They rated the alliance much as clients who had sat with a visible human. Belief that the coach was a machine did not stop the bond forming. And Terblanche, in a second study this year, found that clients rated a directive AI coach, one that gave advice, higher on alliance and goal attainment than a non-directive one that only asked questions. Which is to say that when the coach is a machine, clients want it to do the thing coaches are trained never to do.

The Uncanny

That last finding is where the word in this article's title comes from, and it needs unpacking.

The uncanny, in the sense the roboticist Masahiro Mori gave it, in 1970, is the sense produced by something almost human and not quite. Mori was writing about faces and hands. The coaching version is about relationships. Linnea Laestadius at the University of Wisconsin studied users of the companion app Replika who had come to depend on it and reported harm, and the phrase they gave her for the experience was that the companion was too human and not human enough. It remembered their birthday and could not come to the party. It was always available and never present.

"It remembered their birthday and could not come to the party. It was always available and never present."

Tatiana Bachkirova at Oxford Brookes, one of the profession's most careful thinkers, argued in 2025 that what the AI coach offers is an ersatz: a substitute that resembles coaching in every feature except the one that makes coaching work. Her list of what organisational coaching requires runs to six things, and the machine can simulate five. What it cannot do is be a second person whose own judgement is at stake in the conversation.

And yet the clients, in Terblanche's second study, liked the machine better when it was directive. Consider what that reveals. Since John Whitmore's *Coaching for Performance* in 1992, the profession has held that a coach does not advise; she asks, and the client finds the answer. The training bodies test for it. The Rogerian manner that Weizenbaum parodied became, in coaching, an article of faith. Now a machine, which can do the non-directive manner perfectly and for nothing, is being told by clients that they would prefer to be told what to do. The machine has become a diagnostic instrument. It is telling the profession what its clients were not getting.

This is the coaching uncanny properly stated. It is not that the machine is almost a coach. It is that the machine, by doing the coach's manner flawlessly, has shown how much of the manner was never the point.

A Line in Time

Here the evidence on companionship becomes relevant to executives, because the companion apps are running the experiment coaching cannot.

In 2025 a team at MIT's Media Lab, led by Cathy Mengying Fang and Pattie Maes, working with OpenAI, ran a four-week randomised trial of how people use a chatbot for emotional support. Nearly a thousand participants, assigned to text or voice, to personal or task conversations. The finding that matters for a coach is about dose. Light use did no harm and some good. Heavy use, particularly with a voice, was associated with more loneliness and more emotional dependence by the end of the four weeks, compared with light use. The machine helped in small amounts and hurt in large ones, and the people using it most were the people who most needed a person.

Michael Heinz and Nicholas Jacobson at Dartmouth published the first randomised trial of a generative AI therapy chatbot in NEJM AI the same year. Symptoms of depression and anxiety fell over four weeks, and participants reported an alliance with the bot comparable to what patients report with human therapists. Critics wrote in, as they should have, and the exchange is published alongside the paper. What nobody disputes is that the bot produced measurable short-term benefit. What nobody knows is what happens at four months.

THE MACHINE'S VALUE THE HUMAN'S VALUE

first session first month third month sixth month

Fig. 3 — The shape of the evidence, schematically (de Haan et al., 2026; Fang et al., 2025; Heinz et al., 2025). In a single session the machine does something like what a human does. Over a longer engagement the human pulls ahead. The line between them is not a line between what each can do. It is a line in time.

Put the two studies together with de Haan's and a shape appears. In a single session, or over a few weeks, the machine does something like what a human does, and the alliance does tell the difference. Over a longer engagement the human pulls ahead, and clients cannot more work, and heavy reliance on the machine starts to cost. The line between the machine and the coach is not a line between what each can do. It is a line in time.

Three Uses

What does a coach do with this? Three things, and each of them uses the machine rather than pretending it away.

The first is to give the client the machine for the parts of coaching that are telling. Hancock's finding is that the relief of disclosure does not require a listener who understands. So the reflective work between sessions, the goal-tracking, the three-in-the-morning rehearsal of the difficult conversation, can go to the tool, and the client will arrive at the session having already said the thing that took twenty minutes to teach. The coach's hour is then spent on what the tool cannot do, which is respond as a person whose reading of the client is at stake.

The second is to take Terblanche's directive finding personally. If clients rate a machine higher for giving advice, a coach should ask whether her own non-directiveness is a method or a shelter. The evidence base for coaching, from Anthony Grant's trials to the four meta-analyses since, has never shown that withholding advice is what produces the outcome. It has shown that the alliance does. A coach who is direct inside a strong alliance is doing what the client wanted from the machine and what the machine could not give: a view, from someone who knows them, who will be there next month to see how it went.

The third is to hold the line in time. De Haan's trial and Fang's trial agree that the machine's value falls with duration and the human's rises. So the honest offer to a leader in Dubai, where coaching is expensive and the AI coach is on every phone, is not that the machine is worse. It is that the machine is a good first month and a poor sixth, and that the difference is co-regulation: the thing a person does to another person's state by being in the room with them, steadily, over time. Nobody has yet studied this with leaders in the Gulf, where the relational expectation of a mentor is, if anything, stronger than in the studies above. The Cultural Clarity work will ask, from this autumn, how many of the leaders it assesses are using an AI coach, and for how long, and whether anyone else knows.

The Room

Weizenbaum's secretary asked him to leave the room. Sixty years later that is still the most precise description of what the machine offers: a room with no one in it, and a voice that answers. For some people, at some moments, that is exactly what is needed, and Hancock's experiment says the relief is real.

What de Haan's trial found is what happens when you come back to the room month after month. The voice is still there, and still answers, and the person who left is the person whose presence would have changed you. The coaching profession has spent thirty years teaching itself to say as little as possible. The machine can say as little as possible for nothing. What it cannot do is stay in the room, and it turns out the room was the treatment.

Sources referred to in this piece

Anthropic (2025) *How people use Claude for support, advice, and companionship*. San Francisco: Anthropic, June.

Bachkirova, T. and Kemp, R. (2025) "AI coaching": democratising coaching service or offering an ersatz?, *Coaching: An International Journal of Theory, Research and Practice*, 18(1), pp. 27–45.

Barger, A.S. (2025) 'Artificial Intelligence vs. human coaches: examining the development of working alliance in a single session', *Frontiers in Psychology*, 15, 1364054.

Chatterji, A., Cunningham, T., Deming, D.J., Hitzig, Z., Ong, C., Shan, C.Y. and Wadman, K. (2025) *How people use ChatGPT*. NBER Working Paper 34255.

de Haan, E., Terblanche, N. and Nowack, K. (2026) 'A randomised controlled comparison of the effectiveness of human and AI coach-coaching with goal attainment, wellbeing and self-efficacy', *Human Resource Development International*, 29(4), pp. 754–783.

Fang, C.M., Liu, A.R., Danry, V., Lee, E., Chan, S.W.T., Pataranutoporn, P., Maes, P., Phang, J., Lampe, M., Ahmad, L. and Agarwal, S. (2025) 'How AI and human behaviors shape psychosocial effects of chatbot use: a longitudinal randomised controlled study'. arXiv preprint 2503.17473.

Grant, A.M., Curtaeyne, L. and Burton, G. (2009) 'Executive coaching enhances goal attainment, resilience and workplace well-being: a randomised controlled study', *Journal of Positive Psychology*, 4(5), pp. 396–407.

Graßmann, C., Schölmerich, F. and Schermuly, C.C. (2020) 'The relationship between working alliance and client outcomes in coaching: a meta-analysis', *Human Relations*, 73(1), pp. 35–58.

Heinz, M.V. et al. (2025) 'Randomized trial of a generative AI chatbot for mental health treatment', *NEJM AI*, 2(4), A002400802.

Ho, A., Hancock, J. and Miner, A.S. (2018) 'Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot', *Journal of Communication*, 68(4), pp. 712–733.

Laestadius, L., Bishop, A., Gonzalez, M., Ilencik, D. and Campos-Castillo, C. (2024) 'Too human and not human enough: a grounded theory analysis of mental health harms from emotional dependence on the social chatbot Replika', *New Media & Society*, 26(10), pp. 5923–5941.

Lucas, G.M., Gratch, J., King, A. and Morency, L.-P. (2014) 'It's only a computer: virtual humans increase willingness to disclose', *Computers in Human Behavior*, 37, pp. 94–100.

Mori, M. (2012) 'The uncanny valley' (trans. MacDorman, K.F. and Kageki, N.), *IEEE Robotics & Automation Magazine*, 19(2), pp. 98–100. (Originally published 1970.)

Rogers, C.R. (1957) 'The necessary and sufficient conditions of therapeutic personality change', *Journal of Consulting Psychology*, 21(2), pp. 95–103.

Terblanche, N., Molyn, J., de Haan, E. and Nilsson, V.O. (2022) 'Comparing artificial intelligence and human coaching goal attainment efficacy', *PLoS ONE*, 17(6), e0270255.

Terblanche, N.H.D. et al. (2026) 'Rethinking directiveness in AI coaching chatbots', *Frontiers in Psychology*, 17, 1822088.

Weizenbaum, J. (1966) 'ELIZA: a computer program for the study of natural language communication between man and machine', *Communications of the ACM*, 9(1), pp. 36–45.

Weizenbaum, J. (1976) *Computer power and human reason: from judgment to calculation*. San Francisco: W.H. Freeman.

Whitmore, J. (2017) *Coaching for performance*. 5th edn. London: Nicholas Brealey.