


Academic Paper

Coaching at Scale: Investigating the Efficacy of Artificial Intelligence Coaching

Nicky Terblanche  (University of Stellenbosch Business School)

Joanna Molyn  (Oxford Brookes University)

Erik De Haan  (VU University, Amsterdam)

Victor O. Nilsson  (Hult International Business School)

Abstract

There is limited empirical efficacy evidence on the confluence of artificial intelligence (AI) and organisational and life coaching. Coaching “works” but is often unavailable or unaffordable. AI could scale coaching to reach a wider audience, however, we do not yet know how well AI coaching “works”. This replication randomised controlled trial longitudinal study tested the efficacy of a chatbot AI coach called Vici. An experimental group (n=75) used Vici for six months. Eight measurements on goal attainment, resilience, psychological wellbeing, and perceived stress were collected from the experimental and control group (n=94). Data was collected at baseline, after each of the six chatbot usage months, and three-months later. The experimental group showed a statistically significant increase in goal attainment, while all other measures yielded non-significant results. Using AI, goal and control theories we interpret these results to indicate that AI coaching is effective in a narrow application, suggesting that AI could democratise coaching in a cost-effective, scalable manner.

Keywords

artificial intelligence coaching, AI coaching, chatbot coach, goal attainment, psychological wellbeing, resilience, perceived stress, positive psychology

Article history

Accepted for publication: 14 July 2022

Published online: 01 August 2022



© the Author(s)

Published by Oxford Brookes University

Introduction

Internationally, coaching is a fast-growing multi-billion dollar industry (International Coaching Federation [ICF], 2020a) with wide-ranging, proven benefits to individuals (Athanasopoulou & Dopson, 2018; Blackman, Moscardo and Gray, 2016; Grover & Furnham, 2016). However, coaching is not widely available (Shoukry & Cox, 2018) or affordable (Terblanche, Passmore and Myburgh, 2021) in all societies. Artificial intelligence (AI), able to replace certain human expertise at a vastly reduced cost, could be a potential game changer for coaching (Acemoglu & Restrepo, 2018).

AI as a concept originated in the 1950s and has seen a number of surges and declines (Haenlein & Kaplan, 2019). However, in recent years there has been renewed interest in the application of AI in numerous contexts including helping professions such as psychology, healthcare (Kamphorst, 2017) and, more recently, coaching.

Coaching is defined as a one-on-one structured conversation between a coach and client with the aim of facilitating sustainable change for the individual and potentially other stakeholders (Bachkirova, Cox and Clutterbuck, 2014). Research in coaching has grown substantially (De Haan, 2021; De Haan & Nilsson, in press). Although it is recognized that proving coaching efficacy is difficult and expensive research, as is the case for therapy, this type of research is imperative to verify the claims of practitioners and to understand the factors, processes and mechanisms underlying the interventions (Grover & Furnham, 2016). Examples of recent, well designed efficacy studies include Jones, Woods and Zhou (2021) who found that individual characteristics influence coaching efficacy, and Fontes and Dello Russo (2021) who found that coaching was associated with increases in psychological capital, job attitudes and aspects of job performance. In addition, there are a number of meta-studies that provide compelling evidence that coaching has positive outcomes for the individuals and their organisations (Athanasopoulou & Dopson, 2018; Blackman et al., 2016; De Haan, 2021; Grover & Furnham, 2016; Jones, Woods and Guillaume, 2016; Theeboom, Beersma and van Vianen, 2014).

There appears to be broad agreement on coaching efficacy. Evidence of AI efficacy in the coaching domain however is sparse. There are studies on virtual and e-coaching (see for example Geissler, Hasenbein, Kanatouri and Wegener, 2014), but these studies involve human coaches and cannot be considered AI coaching (Terblanche, 2020). In the coaching domain AI research seems to be limited to conceptual papers (Graßmann & Schermuly, 2020; Terblanche, 2020), technology adoption studies (Terblanche & Cilliers, 2020), or the use of AI entities that employ pure psychological approaches (Ellis-Brush, 2021). Studies where human coaches are replaced entirely by AI are more common in the related fields of psychology and healthcare and will be discussed later in this paper (see Gaffney, Mansell & Tai 2019; Lattie, Adkins, Winquist, Stiles-Shields, Wafford and Graham, 2019).

Although coaching has been shown to be effective in developing people, the use of coaching especially in organisational settings is still relatively limited as observed by Shoukry and Cox (2018). Cost is partly to blame as Terblanche et al. (2021), for example, found that the average fee for executive coaching in the USA is 300 USD per hour and even in lower income countries like South Africa, the average rate is around 100 USD per hour. Employing AI coaching could significantly reduce the cost of coaching and democratise this service, allowing people, who would not normally be able to afford it, to benefit. This is especially true in the developing world. Clearly AI coaching holds potential to expand the benefits of coaching beyond the current limited reach; however, the question remains whether AI coaching “works”?

To address this gap in knowledge and to take a first step towards assessing AI coaching efficacy, we used a goal-attainment coaching chatbot called Vici to replicate a previous human-coach randomised controlled trial (RCT) longitudinal study currently undergoing peer review. The original study, situated in positive psychology, cognitive behavioural and goal attainment theories found that human coaches were able to improve goal attainment, psychological wellbeing, resilience and reduce stress. Our study aimed to investigate to what extent an AI coach could also improve these aspects in order to address the important issue of the accessibility of coaching.

The literature review that follows has three objectives. Firstly, we justify why we chose to measure goal attainment, psychological wellbeing, resilience, and stress. Secondly, we explore the theoretical foundations of coaching efficacy in order to understand why AI coaching may be effective. Thirdly we present empirical evidence from recent research on the current abilities of AI to facilitate desired coaching outcomes. This overview forms the basis of interpreting the results of the present study.

Proven coaching outcomes and theoretical justifications

The continued growth of the coaching industry precipitates the effort by researchers to study and understand the coaching phenomenon. The increase in coaching research is evident from several recent meta-studies that have classified and examined coaching outcomes. As far back as 2008, the overview by Kombarakaran, Yang, Baker and Fernandes (2008) found that coaching helps to improve people management, relationships with managers, goal setting and prioritisation, engagement and productivity, and dialogue and communication. Theeboom et al.'s (2014) meta-study found that coaching has a significantly positive effect on five individual-level categories: performance and skills; wellbeing; coping; work attitudes; and goal-directed self-regulation. In their meta-study, Blackman et al. (2016, p. 469) found a host of positive outcomes including: improved work/life balance; psychological and social competencies; self-awareness and assertiveness, increasing confidence; developing relationships, networks and interpersonal skills; adapting to change more effectively; helping to set and achieve goals; role clarity; and changing behaviors.

From these meta-studies, it is apparent that while coaching clearly has positive outcomes, there is little consensus in the literature regarding the most appropriate criteria for classifying and evaluating coaching outcomes (De Haan & Nilsson, in press; Grant, Passmore, Cavanagh and Parker, 2010; MacKie, 2007; Smither, 2011). To address this issue, Jones et al. (2016) proposed a three-component classification framework for coaching outcomes comprising of cognitive, skills-based and affective categories. They stated that cognitive outcomes are typically guided by goal-setting and that many of the coaching outcomes are affective in nature, including development of self-efficacy, confidence, reduced stress, increase satisfaction and motivation. In line with this framework, the meta-study by Athanasopoulou and Dopson (2018) found that positive outcomes for the coachee include personal development specifically related to reduced stress/anxiety, increased work/life balance (wellness), improved resilience and improved goal-setting.

The choice of measurement constructs for the present study (goal attainment, psychological wellbeing, resilience, perceived stress) is therefore validated by the coaching outcomes framework of Jones et al. (2016) and the findings of Athanasopoulou and Dopson (2018). The constructs are also supported by other studies who found coaching to increase *goal attainment* (Diller, Muehlberger, Braumandl and Jonas, 2020; Grant, Curtayne and Burton, 2009; Spence & Grant, 2007), *wellbeing* (Duijts, Kant, van den Brandt and Swaen, 2007; Govindji & Linley, 2007; Grant et al., 2009; Spence & Grant, 2007; Theeboom et al., 2014), *resilience* (Grant et al., 2009; Green, Grant and Rynsaardt, 2020), and to reduce *stress* (Grant et al., 2009; Junker, Pömmmer and Traut-Mattausch, 2020). The measurement constructs for the present study were also selected in order to compare the results directly with a similar study involving human coaches currently undergoing peer review.

As an emerging, multi-disciplinary field of research and practice, coaching has numerous theoretical underpinnings that explain its efficacy (Grant, 2014; Shoukry & Cox, 2018). The main theories informing the present study are goal and control theory. Goal attainment is considered the hallmark of coaching that sets it apart from other helping professions (Grant, 2012; 2014). Goal theory is an overall approach to motivation that emphasises the need to establish goals as intrinsic motivation. A relationship exists between goal difficulty, level of performance, and effort involved (Locke & Latham, 2002). Locke and Latham (1990) state five principles of goal setting: clarity (specific and clear), challenging (sufficiently difficult), commitment (buy-in from onset), and feedback (regular stock-taking on progress). In essence, goals are 'internal representations of desired states or outcomes' (Austin & Vancouver, 1996, p.388). Goal theory is relevant to coaching as a mechanism to facilitate self-regulation (Grant, 2006). Grant (2006) holds that coaching is at its core a process of goal-focused self-regulation where individuals set goals, develop and execute action plans, monitor progress and change either goals or action plans based on feedback and progress (Carver & Schneiner, 1998). There are various types of goals, for example proximal (short-term) and distal (long term) goals (Grant, 2006). Control theory compliments goal theory and suggests that the process of monitoring influences goal progress. Monitoring helps to translate

goals into actions and actions lead to progress (Harkin, Webb, Chang, Prestwich, Conner, Kellar, Benn and Sheeran, 2016)

Artificial intelligence and coaching

Artificial intelligence (AI) has gained renewed prominence in recent years, likely as a result of the announcement of the arrival of the Fourth Industrial Revolution in 2016 (Schwab, 2017). AI is considered by some as one of the most significant developments in human history with the potential to disrupt all aspects of human life (Acemoglu & Restrepo, 2018). AI is defined as “the broad collection of technologies, such as computer vision, language processing, robotics, robotic process automation and virtual agents that are able to mimic cognitive human functions” (Bughin & Hazan, 2017, p. 4).

A distinction can be made between artificial narrow intelligence (ANI), which refers to systems that can perform only a very specific task in a narrow context; artificial general intelligence (AGI), which refers to systems that are at least as intelligent as humans and can apply their learning in different contexts; and artificial super intelligence (ASI), which refers to systems that can outperform humans in most dimensions (Bostrom, 2014; Shanahan, 2015; Siau & Yang, 2017). AGI and especially ASI are currently still in their infancy with no clear indication of reaching maturity in the foreseeable future. ANI, however, is showing steady signs of progress with encouraging results in specific applications, such as speech recognition and self-driving cars (Panetta, 2018). Expert systems, relevant to the present study are considered a form of ANI and are defined as complex software programmes based on specialised knowledge, able to provide acceptable solutions to individual problems in a narrow topic area (Chen, Hsu, Liu and Yang, 2012; Telang, Kalia, Vukovic, Pandita and Singh, 2018).

Given the current limitation around the levels of general intelligence of AI, ANI expert systems should be employed when creating AI coaches (Terblanche, 2020). At present, chatbots (an instance of ANI) in particular is a promising technological platform for creating AI coaches (Terblanche, 2020). Chatbots are defined as computer programmes that interact with users via natural language either through text, voice, or both (Chung & Park, 2019). Chatbots can be embodied (have a physical presence) or disembodied (no physical presence) (Araujo, 2018). Considerations of the level of human-likeness (anthropomorphism) of an AI entity are important as it has been shown to influence aspects such as resilience (De Visser, Monfort, McKendrick, Smith, McKnight, Krueger and Parasuraman, 2016) and even gambling frequency (Riva, Saachi and Brambilla, 2015).

Artificial intelligence coaching efficacy

Research into the efficacy of AI in the coaching domain as defined in the present study is sparse. To gain insight into the potential efficacy of AI coaching, we turn to the related fields of psychology and healthcare where AI has been studied for longer. We start by reviewing two specific examples of AI application, one in psychology and one in healthcare before reviewing two relevant meta-studies.

A study by Fulmer, Joerin, Gentile, Lakerink and Rauws (2018) investigated the efficacy of a cognitive behavioral therapy (CBT) based AI agent to reduce self-identified symptoms of depression and anxiety in college students. They found that the AI reduced symptoms of depression (after two weeks) and anxiety (after four weeks) in the user group relative to a control group. They conclude that AI can serve as a cost effective and accessible therapeutic agent. In a RCT study from the healthcare domain, Greer, Ramo, Chang, Fu, Moskowitz and Haritatos (2019) investigated the feasibility of improving key psychosocial well-being outcomes in young adults treated for cancer by delivering positive psychology skills via a chatbot. They found that after four

weeks the experimental group had reduced anxiety compared to the control group, however there were no significant changes in depression, positive or negative emotions.

Two meta-studies provide a general overview of AI efficacy in psychology. The first is a meta-study by Lattie et al. (2019) on digital mental health interventions for depression, anxiety and psychological wellbeing among college students. They found that the majority of the 89 programs evaluated were either effective or partially effective in producing beneficial psychological changes, however they conclude that generally the research quality was problematic. The second meta-study by Gaffney, Mansell and Tai (2019) looked at the ability of conversational agents to treat mental health problems. They found that all 13 studies reviewed (eight were RCT) reported some reductions in psychological distress. Five of the RCT studies demonstrated significant reductions in psychological distress compared with control groups while three RCT studies failed to demonstrate superior effects. They conclude that efficacy and acceptability of conversational agent interventions for mental health problems are promising, but that more robust experimental designs are required to demonstrate efficacy and efficiency. It appears therefore that AI can be effective in treating psychological and health related issues, suggesting that AI could potentially also be effective in the coaching domain.

Summary and hypotheses

Given the preceding overview and discussion on coaching efficacy, general AI capabilities and current AI efficacy in fields related to coaching, we now summarise our position and state hypotheses for the present study.

We aimed to investigate the ability of an AI coach, Vici, to improve the goal attainment, psychological wellbeing, resilience and to reduce perceived stress of users. The review has shown that AI agents are able to improve psychological and health related outcomes relative to a control group. We therefore expect that our AI coach, Vici may have some efficacy. A limitation of Vici is that it was designed based on expert system and artificial narrow intelligence principles whereby its functionality is limited to a specific objective, that of assisting with goal-attainment. It is therefore reasonable to expect that:

H1: Using the AI Coach will provide a statistically significant improvement in *goal-attainment* for the experimental group relative to the control group.

A number of studies from the psychological and healthcare domains have demonstrated the ability of AI to improve psychological wellbeing, resilience and reduce stress. The AI entities used in these studies were specifically designed to incorporate CBT practice, known to positively affect a range of psychological outcomes. Although Vici was not designed to improve psychological state, there is a positive link between goal attainment and wellbeing. Research has shown that the feedback provided to people through achieving their goals could help them to overcome obstacles and ultimately could lead to enhanced wellbeing (Koestner, Lekes, Powers and Chicoine, 2002; Niemiec, Ryan and Decci, 2009; Sonnentag, 2002). Therefore, even though Vici only assists with goal attainment we still expect that:

H2: Using the AI Coach will provide a statistically significant improvement in *psychological wellbeing* for the experimental group relative to the control group.

In terms of the final two constructs, resilience and perceived stress we do not expect Vici to be of assistance because it was not designed to improve these facets and there is no clear link between goal attainment (Vici's primary focus) and resilience and perceived stress. We included these measures since our research is a replication study of a human coaching efficacy study that measured these aspects. We therefore state our final two hypotheses as follows:

H3: Using the AI Coach will not provide a statistically significant improvement in *resilience* for the experimental group relative to the control group.

H4: Using the AI Coach will not provide a statistically significant decrease in *perceived stress* for the experimental group relative to the control group.

Methodology

Participants

An email invitation to participate in this research was sent to undergraduate students studying at a business school in the United Kingdom. Their fields of study included business management, economics, marketing, tourism, events management and logistics. They were offered employability-scheme credits to encourage participation. In total, 268 students responded to the first baseline questionnaire, who were then randomly allocated into two different groups: one experimental group (n=134) receiving AI coaching and one non-intervention control group (n=134). After the initial baseline survey, 97 participants who had been allocated to the experiment engaged with the chatbot coach and responded to the second survey, while 121 participants in the control group responded to the second survey. A further 22 participants in the experimental group dropped out through the period of data collection and 27 of the control group dropped out, leaving a final sample of 168 participants (56% females) with 75 in the experimental group and 94 in the control group. Their average age was 22 years old (sd = 4.96) and the two groups did not significantly differ in either gender distribution or age distribution.

In order to capture a placebo effect a study design requires subjects to not be aware whether they receive treatment or not (Wampold & Imel, 2015). In that sense we were unable to investigate the placebo effect in our study as coached students were given access to Vici almost immediately after the start of the experiment. The students in the control group were also aware of not having access to the AI coach. To address outcomes expectancy (Colagiuri & Smith, 2012) the control group received a fact sheet that provided information about goal attainment, psychological wellbeing, resilience and perceived stress. They were also asked to think of and specify goals they wanted to achieve over the next 10 months.

In terms of sample size, we performed a power-analysis based on a Mixed Factorial MANOVA over eight time points between two groups using G*Power 3.1.9.7 (Faul, Erdfelder, Lang & Buchner, 2007). A meta-analysis indicated that coaching has a lower effect on variables such as psychological well-being compared to goal attainment (de Haan & Nilsson, 2021, unpublished). Thus, to be able to identify an effect on all measured scales, a power-analysis was conducted based on results from a previous psychological well-being study rather than previous goal-attainment effects. Previous research on the impact of coaching on psychological wellbeing showed an effect of $\eta^2 = .033$ (de Haan, Gray & Bonnywell, 2019) which was used as a baseline for the analysis. The analysis indicated that a total sample size of 126 participants is required to identify the given effect size with a power of .80 and alpha level of .05. The true sample size of this project (N = 169) provided a 90% probability of detecting the effect size which suggest that the sample size in this study is appropriate for answering the hypotheses.

Procedure

The data collection took place over eight time points from October 2019 until September 2020. T1 captured pre-study scores of demographics, independent and dependent variables. T2 to T7 was one month apart and required participants in both groups to complete the same survey measuring the independent and dependent variables at the end of the month. The experimental group had to

use the AI coach at least once before completing the monthly survey. There was a three-month lapse between T7 and the final data collection point T8. During this period experimental group participants were free to use the AI coach optionally. Both control and experimental groups completed the same surveys, and these surveys were sent to both groups at the same time. The participants were also asked to write down two goals that they were aiming to work on over the duration of the study. Each student had a unique log-in for their survey which allowed them to see their original goals every time they logged in and responded to the survey. One of the researchers sent out reminders to the students to complete the survey if they had not engaged with the data collection.

AI coach

Vici, the AI coach used in this study is a custom development, text-based chatbot designed according to the DAIC (Designing AI Coaches) framework (Terblanche, 2020) and deployed on the Telegram instant messaging platform. The DAIC framework merges human coaching efficacy aspects of a strong coach-coachee relationship with chatbot design best practices. It recommends that, due to the existing limitation of ANI, the AI coach should be designed to fulfil one specific coach task only, in line with expert system design principles (Terblanche, 2020). This recommendation was used to inform the design of Vici to have the sole purpose of being a goal-attainment chatbot coach. The choice of goal attainment as focus for Vici was based on the fact that, while coaching has numerous positive outcomes, goal attainment is arguably what distinguishes coaching from other similar helping professions (Grant, 2012; 2014). A further recommendation of the DAIC framework is that the AI coach must be designed based on proved coaching theories. As such, Vici uses goal-attainment theory (Grant, 2014; Latham & Locke, 2007) and the GROW coaching model (Grant, 2011).

Vici's sole coaching objective is assisting people with goal achievement. Vici helps users to identify goals, specify actions to reach the goals, monitor the progress of goals and actions, and to adjust any if necessary. Vici also helps users to distinguish between proximal and distal goals (Latham & Locke, 2007) and keeps track of both. Vici operationalises goal-attainment theory through two types of coaching conversations: initial goal-setting and progress-tracking. The initial goal-setting conversation follows the GROW coaching model (Grant, 2011). Developed by Graham Alexander in 1984 (West & Milan, 2001) and popularised by Whitmore (2003) the GROW Coaching model consists of goal-identification, reality checking, identification of options available to reach the goal, and the will and commitment to reach the goal. The progress-tracking conversation records, monitors and assesses the progress made towards proximal and distal goals and optionally allows goals to be changed. Vici was available 24/7 to the experimental group.

Measures

In this study, we report all measures, manipulations and exclusions.

Goal Attainment

The authors adapted goal attainment to be suitable for the purpose of this article based on the scale from Grant et al. (2009). The participants in this study were asked to write down two goals they would work on over the time and they were instructed that the goals should be something challenging that is either new or something that has been difficult in the past. The participants then rated the two self-assigned goals separately in terms of both success and difficulty. The success-score was measured from 0% (no achievement at all) up to a 100% (total success), while the difficulty scale was measured on a 7-point Likert scale from 'very easy' to 'very difficult'. Goal attainment scores were then computed by multiplying the success score with the difficulty score for each goal and then creating a mean score between the two goals to create a composite scale.

Psychological Wellbeing

The authors used the long version of the Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS) (Stewart-Brown, Tennant, Tennant, Platt, Parkinson and Weich, 2009) to measure the psychological wellbeing of the participants. This 14-item scale measures their perceived psychological wellbeing using a 5-point Likert scale, from 'None of the time' to 'All of the time'. An example item is: "I've been feeling relaxed". The internal consistency analysis of the WEMWBS showed a Cronbach's Alpha of 0.93 which indicates a high internal consistency.

Resilience

Brief Resilience Scale (BRS) is a 6-item scale, measured by a 7-point Likert Scale from 'Strongly disagree' to 'Strongly agree' (Smith, Dalen, Wiggins, Tooley, Christopher and Bernard, 2008). The BRS includes items such as "I tend to bounce back quickly after hard times". The internal consistency analysis of the BRS showed a Cronbach's Alpha of 0.77 which indicates an acceptable internal consistency.

Perceived Stress

This project measured the participants' stress level by using the Perceived Stress Scale (PSS) (Cohen, Kamarck and Mermelstein, 1983). This is a ten-item scale that measures how often stressful events occurred the last month using a 5-point Likert scale from 'Never' to 'Often'. An example of an item was: "How often have you been able to control irritations in your life?" The internal consistency analysis of the PSS showed a Cronbach's Alpha of 0.85 which indicates a high internal consistency.

Note that Shapiro-Wilk tests were used on all scales to assess the normal distribution of the scales due to the different nature of the measures. The tests indicated that none of the scales, for either of the groups deviated from a normal distribution.

Results

The data from the surveys was first screened to determine any potential relationships between the dependent variables. Pearson correlation on the whole sample indicated that Psychological Wellbeing, Resilience and Stress showed a medium-sized correlation with each other, ranging from $r = .50$ to $r = -.64$ (correlations with stress being negative). However, Goal Attainment was not related to any of the other dependent variables. See Table 1 for correlation matrices for the different dependent variables. These correlations indicate that Goal Attainment is being measured independently of the other dependent variables, while Psychological Wellbeing, Resilience and Stress were analysed in one MANOVA to avoid family errors that may occur with multiple testing of related variables.

Table 1: Correlation matrices for the different dependent variables

	1.	2.	3.	4.
1. Goal attainment		.10	-.01	.09
2. Resilience		$\alpha = .77$	-.56***	.48***
3. Perceived stress			$\alpha = .85$	-.64***
4. Psychological wellbeing				$\alpha = .93$

Note: Table shows Pearson's correlations *** = $p < .001$.
 α = Cronbach's alpha of the measurement

H1: Using the AI Coach will provide a statistically significant improvement in goal-attainment for the experimental group relative to the control group.

A Mixed Factorial ANOVA was conducted using the Goal Attainment Scale between the two groups in order to answer *Hypothesis 1*. The analysis indicated that there was a statistical significant interaction of time and group, $f(7, 973) = 4.44$, $p < 0.001$, $\eta p^2 = 0.031$. This significant interaction allowed follow-up analysis over time, per group, and comparisons between the groups on each time point. Both groups indicated an initial increase in the Goal Attainment for the first three months and this increase was statistically significant for each group at three months (exp, $p = 0.005$, control, $p = 0.01$) according to Bonferroni corrected Pairwise comparisons. Moreover, the groups did not significantly differ between each other at this time point, $p = 0.82$, $d = 0.03$. However, the AI coaching chatbot had an effect after using Vici for three months where the goal attainment for the experimental group kept on significantly increasing while the control group faded off. At the end of the project, three months after the AI coaching had finished, the difference between the groups was now statistically different after Bonferroni adjustments for multiple comparisons: $p = 0.005$, $d = 0.60$, as illustrated in Figure 1. See Table 2 for all means and standard deviations over time.

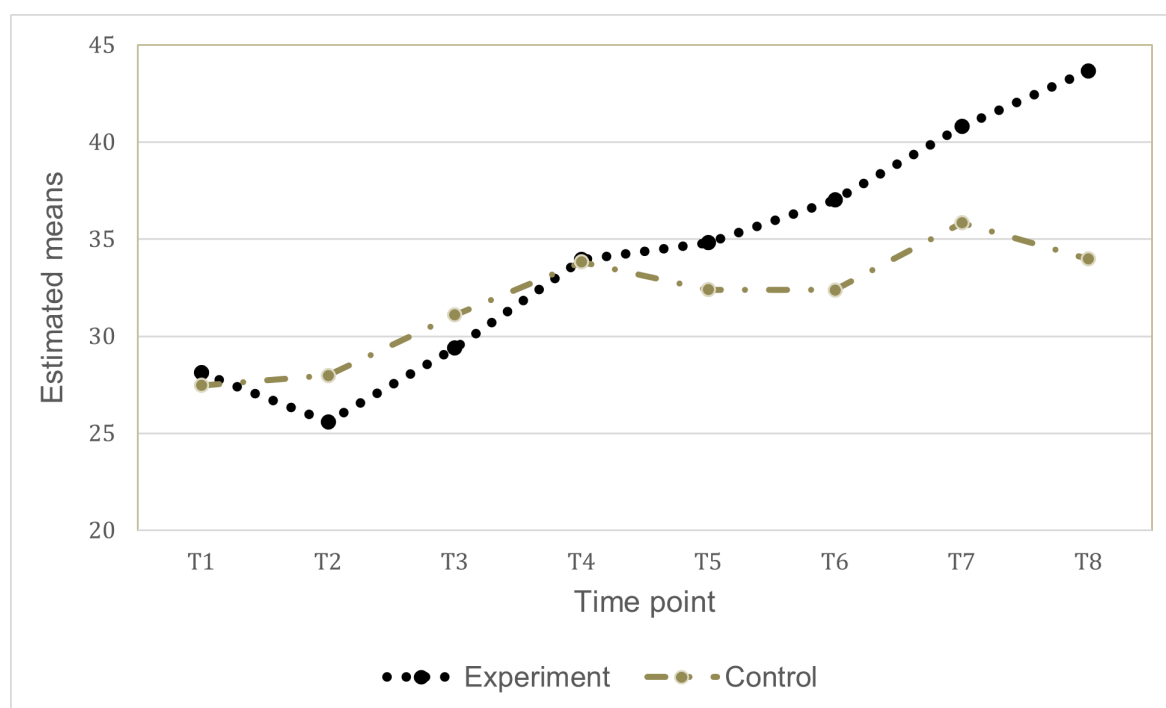
Table 2: All means and standard deviations over time

Measure	Group	T1	T2	T3	T4	T5	T6	T7	T8
Goal attainment	Exp	28.11 (14.68)	25.60 (13.16)	29.39 (14.22)	33.95 (14.83)	34.82 (14.38)	37.02 (18.09)	40.81 (19.7)	43.68 (20.73)
	Contr	27.47 (15.32)	27.99 (16.25)	31.09 (16.37)	33.82 (17.80)	32.41 (18.15)	32.38 (18.72)	35.85 (18.94)	33.97 (19.39)
Resilience	Exp	3.34 (0.69)	3.20 (0.8)	3.24 (0.69)	3.31 (0.67)	3.34 (0.67)	3.28 (0.68)	3.24 (0.73)	3.21 (0.74)
	Contr	3.40 (0.62)	3.27 (0.69)	3.33 (0.74)	3.35 (0.70)	3.34 (0.70)	3.31 (0.76)	3.23 (0.73)	3.34 (0.73)
Psychological wellbeing	Exp	47.75 (8.47)	48.68 (8.87)	48.40 (8.45)	47.69 (8.72)	46.60 (8.35)	45.39 (10.10)	46.91 (10.02)	47.32 (8.80)
	Contr	47.13 (9.55)	48.56 (9.40)	47.44 (10.93)	46.97 (10.83)	46.99 (12.17)	46.08 (12.02)	45.48 (12.60)	47.57 (11.13)
Perceived stress	Exp	27.26 (6.99)	27.85 (6.54)	27.87 (6.66)	28.15 (7.09)	29.82 (6.04)	30.08 (7.12)	29.65 (6.89)	27.66 (7.44)
	Contr	28.29 (7.06)	28.54 (6.05)	28.54 (7.24)	29.09 (7.03)	29.62 (8.16)	30.27 (7.32)	30.25 (7.21)	29.03 (7.45)

Note: The table shows means for each time point and group (standard deviations within brackets)

We analysed the usage of the AI application in terms of how many times the AI coaching chatbot was used to identify any potential within-group differences. We were able to identify a significant difference in development of Goal Attainment (and not of any of the other dependent variables) when splitting the frequency of usage into two equal groups based on their median usage (6 AI coaching sessions), $t(73) = -2.24$, $p = 0.028$, $d = 0.52$. The lower usage group had an average increase on Goal Attainment of 17.62 (sd = 32.50) compared to 37.62 (sd = 34.16) in the higher usage group.

Figure 1: Goal attainment



These findings indicate that the AI coaching chatbot did not seem to have an immediate effect on goal attainment. However, using the chatbot over three months' time resulted in a significant effect, which was sustained three months after the end of the research project. The experimental group showed an increase of 55% on their goal attainment compared to 24% in the control group, the findings further indicate that, using the coaching application more frequently, led to a higher increase in the goal attainment from T1 to T8. Thus, the findings from these analyses support *Hypothesis H1*.

The dependent variables Psychological Wellbeing, Resilience and Perceived Stress were all analysed in a Mixed Factorial MANOVA since they showed moderate correlations with each other. We conducted the Mixed Factorial MANOVA to identify if there were developments between the two groups of the three different but correlated variables. However, the analysis could not identify a significant interaction of time and group, $\Lambda = 0.91$, $f(21, 119) = 0.56$, $p = .94$, $\eta p^2 = 0.09$. Furthermore, univariate tests of each dependent variable could not either identify any unique interaction of group and time on any of the three variables (resilience, $p = 0.80$; psychological wellbeing, $p = 0.89$; and stress, $p = 0.91$).

The insignificant results of the Mixed Factorial ANOVA suggest that the AI coaching chatbot in the study did not have an impact on Psychological Wellbeing, Resilience or Perceived Stress. Therefore, *Hypothesis H2 is rejected and Hypotheses H3 and H4 are supported*.

Given the positive effect the AI coaching chatbot had on goal attainment, we conducted further qualitative analysis on the nature of the goals in both groups. Two of the authors independently analysed the first goal to assess the theme of the goal, the type of outcome (concrete or vague goal) and whether the goal was proximal (<6 months) or distal (>6 months). The inter-rater reliability of the analysis indicated a very high similarity between the reviewers on all three categories, with Cohen's kappa of $\kappa = .96$, $p < .001$ for the theme of the goal, $\kappa = .90$ for the outcome and $\kappa = .87$, $p < .001$ for the timeline of the goal.

The themes of the goals that were identified relate to the participants' studies (42%), self-development (21%), career related (18%), health and well-being (17%), other (2 %), financial

related (1%) and family related goals (0.5%). Most of the goals were concrete and measurable (65%), for example, "To gain overall mark of 75% in study year 1" and 55% of the goals were long term focused (> 6 months).

The proportion of the type of goals, the type of outcome and the timeline was tested between the groups to identify if there were any association of group and the goals. However, chi-square tests did not identify any significant associations of the goals which indicates that the goals were similarly assigned between both groups.

Discussion

This RCT longitudinal study investigated the ability of Vici, a goal-attainment AI chatbot coach to improve the goal attainment, psychological wellbeing, resilience and reduce the perceived stress of people. Coaching as a helping profession is growing in stature with convincing evidence of its efficacy in various domains, including the four constructs investigated in this study. Understanding to what extent an AI coach could improve these aspects could help to democratize coaching by making its benefits available to a much wider audience.

There are three main findings of the present study that require in-depth discussion: (1) Vici was able to assist participants with increased goal attainment relative to the control group; (2) Vici users only outperformed the control group in goal attainment after three months (T4); (3) Vici was not able to help participants with improved resilience, psychological wellbeing or reduced perceived stress.

To explain our first finding we turn to control and goal theory. Control theory provides several clues to Vici's performance. Firstly, control theory suggests that the process of monitoring influences goal progress. Monitoring helps to translate goals into actions and actions lead to progress (Harkin, Webb, Chang, Prestwich, Conner, Kellar, Benn and Sheeran, 2016). Monitoring is one of the core features of Vici. After the initial goal-setting conversation, all follow-up interactions with Vici are dedicated to continued monitoring of goals and actions progress, hence providing one explanation of Vici's goal efficacy.

Secondly, the frequency of progress monitoring has a positive effect on goal attainment (see for example Acharya, Elci, Sereika, Styn and Burke, 2011; Coughlin, Gullion, Brantley, Stevens, Bauck, Champagne and Appel, 2013; Sherwood, Crain, Martinson, Anderson, Hayes, Anderson, Senso and Jeffery, 2013). The results from the present study indicate that the issue of frequency is also true for AI-based monitoring. Participants who interacted with Vici more often achieved higher goal attainment.

Thirdly, Harkin et al.'s (2016) meta-study found that the physical recording of goals increases goal attainment. One of Vici's core purposes is to help people capture their goals by physically typing it into the chatbot system. It seems therefore that Vici's ability to monitor goal setting and attainment can be explained through the dynamics of control theory.

This result can also be interpreted in terms of goal theory. Goal attainment is moderated by ability, goal commitment, feedback on goal progress, goal complexity, and external factors such as resources required (Latham & Locke, 2007). In the goal-setting and progress conversations, Vici asks explicit questions covering these aspects including helping users to assess how realistic their goals are, how committed they are, what resources they need and have available. Vici also helps users break down their goals into long-term (distal) and short-term (proximal) goals. Goal-setting theory holds that feedback from a proximal goal helps shape a person's idea of what is realistic and keeps the goal attainment process congruent with the distal goal (Latham & Seijts, 1999). The

setting of proximal goals also leads to more regular feedback, which aids overall goal achievement as discussed earlier and hence explaining Vici's ability to improve goal achievement.

There appears to be compelling evidence from control theory and goal theory that explains why Vici was able to improve goal attainment. This insight empirically supports the conceptual recommendation of the Designing AI Coaches (DAIC) framework (used to create Vici) that AI coaches must be designed using sound theoretical models (Terblanche, 2020).

The second important finding is that Vici users only outperformed the control group in goal attainment after three months (T4). In fact, goal attainment at T2 when users had their first interaction with Vici was lower for the Vici group than the control group. This phenomenon could possibly be explained in terms of control and goal theories. Participants in the control group did not any active feedback. After setting their initial goals, they also experienced initial improvements in goal attainment for the first three months through self-sustained effort. However due to lack of feedback and support their progress eventually decreased, in line with control and goal theory predictions (Acharya et al., 2011; Coughlin et al., 2013; Sherwood et al., 2013). Vici users on the other hand sustained their goal progression after three months due to monitoring and feedback provided by the chatbot. Vici's questioning about the relevance and appropriateness of actions aimed at achieving goals may also have helped users to optimize their action plans. The initial decrease in goal attainment at T2 for Vici users could be because Vici helped users to gain a more realistic sense of their goals compared to the control group. Knowing practically what is required to achieve a goal may cause an initial sense of slower progress. In the long term however, clear and realistic goals lead to a higher level of goal attainment as the results indicate (Latham & Locke, 2007). This finding on the role of the number of AI coaching session on coaching outcomes is also significant since in human coaching there is limited evidence of the impact of the number of coaching sessions (Theeboom et al., 2014). In AI coaching it appears that more sessions are more beneficial, however this claim needs to be investigated further.

The third important finding of our study is that Vici was not able to assist people with psychological wellbeing, resilience or perceived stress. We expected Vici to be successful in improving psychological wellbeing since goal theory suggests that goal setting and achievement have positive influences on a person's wellbeing (Koestner et al., 2002; Niemiec et al., 2009; Sonnentag, 2002). In human coaching, even though the focus may be on goal attainment, the very presence of a human as coach has implications for the scope, focus and outcomes of the intervention. Good human coaches are aware of the importance of maintaining a strong coach-coachee relationship (De Haan et al., 2016). This supportive relational dimension could positively influence aspects such as wellbeing. The fact that Vici was not able to engage on a human relationship level may therefore explain why no spill-over effects of the goal attainment focus was observed. Vici's sole focus on goal-attainment and lack of human-like empathy and contextual awareness may also explain why there was no improvement on resilience and reduction in perceived stress. There is evidence from the psychology domain that AI entities can in fact improve wellbeing, resilience and reduce stress (Gaffney, Mansell & Tai, 2019; Lattie et al., 2019). These AI entities were based on theoretical models from psychology such as CBT, similar to how Vici was created based on goal theory, with the associated outcomes.

Theoretical contributions

This study contributes on a theoretical level to our understanding of the link between coaching and goal theory as well as AI theory. Goal theory states that goal attainment is enhanced by setting clear and challenging (but not too difficult) goals, commitment to the process and obtaining regular feedback (Locke & Latham, 1990). The relevance of goal theory in human coaching has been established (Grant 2006). The present study extends our understanding of goal theory in coaching by showing that a machine algorithm that implements goal theory accurately delivers positive coaching outcomes, thus extending the application and reach of goal theory in coaching.

AI theory holds that the narrow application of AI in the form of expert systems produce applications that can perform one function well (Bostrom, 2014; Shanahan, 2015; Siau & Yang, 2017). The AI coach in the present study was designed based on these principles of narrow AI and expert systems as captured in the DAIC framework that suggests creating AI coaches with a focused approach (Terblanche, 2020). The positive efficacy of Coach Vici therefore supports the theoretical understanding of the limitations of current AI technology and confirms our theoretical understanding of the application of AI in coaching.

Practical implications

Vici's ability to facilitate goal attainment is encouraging in terms of the prospect of democratising coaching. Coaching has proven benefits for people, yet many strata of society do not have access to this service due to limited availability of skilled coaches and associated high costs. As of August 2020, the statistics from the largest such body for coaching in the world, the International Coach Federation (ICF), shows that of their 42 786 members, fewer than 2 000 are practicing in Africa (ICF, 2020b). The average cost of organisational coaching in Africa is approximately 100 USD per hour (Terblanche et al., 2021). This combination of low availability and high cost places human coaching well beyond the reach of most people in Africa. The once-off cost of creating Vici, its scalability, 24/7 availability and now proven efficacy makes a compelling case for employing AI coaching to democratise this helping service for people who are currently unable to access it.

Even in organisations that pay for coaching services, coaching is often reserved for more senior staff due to cost. AI coaching could therefore also democratise coaching within organisations by providing a basic coaching service to more junior employees. Employees who show commitment to the AI coaching process could then be offered a more expensive human coach at a later stage since they are then more likely to maximise the investment. This mechanism could help improve return on investment in coaching by organisations.

Limitations and future research

This study has two notable limitations. Firstly, participants were undergraduate students, which suggests that the results may not generalise to other populations. While students may well benefit from using an AI coach, there are other domains such as the workplace where AI coaching could be used. It is therefore imperative that AI coaching should be investigated beyond the current limited context of higher education. Secondly, we only measured self-scores by participants, which opens the possibility of self-score bias. Self reporting measure are a common occurrence in coaching research and often criticised. Objective measures would enhance the validity of the present study. These limitations are offset to some degree by the research design (longitudinal, RCT) and the relatively large sample size.

In terms of future research, two suggestions are put forward. Firstly, the efficacy of the hybrid between human and AI coaches must be studied. Is the combination of a human and AI coach more efficacious than a human coach alone and what are the contributing factors? Combinations such as blended asynchronous and synchronous models should be researched where a coach is available for remaining issues if the AI coach was inadequate. Secondly, AI coaches who focus on various outcomes other than goal attainment, such as wellbeing, must be created and their efficacy studied. It is suggested that all these studies employ longitudinal RCT designs.

Conclusion

Coaching offers numerous proven benefits to society; however, not everyone can afford a coach. Our study shows that Vici, an AI chatbot coach was able to successfully assist users with a particular aspect of coaching, increased goal attainment, in line with its design. This evidence of the efficacy of AI coaching is a major step towards democratising coaching and offering the benefits of this service to people who could not normally afford it.

References

- Acemoglu, D., & Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6), 1488–1542. DOI: [10.1257/aer.20160696](https://doi.org/10.1257/aer.20160696).
- Acharya, S. D., Elci, O. U., Sereika, S. M., Styn, M. A., & Burke, L. E. (2011). Using a personal digital assistant for self-monitoring influences diet quality in comparison to a standard paper record among overweight/obese adults. *Journal of the Academy of Nutrition and Dietetics*, 111(4), 583–588. DOI: [10.1016/j.jada.2011.01.009](https://doi.org/10.1016/j.jada.2011.01.009).
- Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*, 85, 183–189. DOI: [10.1016/j.chb.2018.03.051](https://doi.org/10.1016/j.chb.2018.03.051).
- 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](https://doi.org/10.1016/j.leaqua.2017.11.004).
- Austin, J. T., & Vancouver, J. B. (1996). Goal constructs in psychology: Structure, process, and content. *Psychological bulletin*, 120(3), 338. DOI: [10.1037/0033-2909.120.3.338](https://doi.org/10.1037/0033-2909.120.3.338).
- Bachkirova, T., Cox, E., & Clutterbuck, D. (2014). Introduction. In E. Cox, T. Bachkirova & D. Clutterbuck (Eds.), *The Complete Handbook of Coaching* (pp. 1–20). Sage.
- Blackman, A., Moscardo, G., & Gray, D. E. (2016). Challenges for the theory and practice of business coaching: A systematic review of empirical evidence. *Human Resource Development Review*, 15(4), 459–486. DOI: [10.1177/2F1534484316673177](https://doi.org/10.1177/2F1534484316673177).
- Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.
- Carver, C.S. & Scheier, M.F. (1998). *On the self regulation of behaviour*. Cambridge, UK: Cambridge University Press.
- Chen, Y., Hsu, C., Liu, L., & Yang, S. (2012). Constructing a nutrition diagnosis expert system. *Expert Systems with Applications*, 39(2), 2132–2156. DOI: [10.1016/j.eswa.2011.07.069](https://doi.org/10.1016/j.eswa.2011.07.069).
- Chung, K., & Park, R.C. (2019). Chatbot-based healthcare service with a knowledge base for cloud computing. *Cluster Computing*, 22(1), S1925–S1937. DOI: [10.1007/s10586-018-2334-5](https://doi.org/10.1007/s10586-018-2334-5).
- Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 24(4), 385–396. DOI: [10.2307/2136404](https://doi.org/10.2307/2136404).
- Colagiuri, B., & Smith, C. A. (2012). A systematic review of the effect of expectancy on treatment responses to acupuncture. *Evidence-based complementary and alternative medicine*, 2012. DOI: [10.1155/2012/857804](https://doi.org/10.1155/2012/857804).
- Coughlin, J. W., Gullion, C. M., Brantley, P. J., Stevens, V. J., Bauck, A., Champagne, C. M., Appel, L. J. (2013). Behavioral mediators of treatment effects in the weight loss maintenance trial. *Annals of Behavioral Medicine*, 46(3), 369–381. DOI: [10.1007/s12160-013-9517-3](https://doi.org/10.1007/s12160-013-9517-3).
- De Haan, E. (2021). *What Works in Executive Coaching: Understanding outcomes through quantitative research and practice-based evidence*. Routledge.
- De Haan, E., Nilsson, V.O. (in press). Does executive coaching work? A meta-analysis based only on randomized controlled trials. *Journal of Occupational and Organisational Psychology*.
- De Haan, E., Grant, A. M., Burger, Y., & Eriksson, P.O. (2016). A large-scale study of executive and workplace coaching: The relative contributions of relationship, personality match, and self-efficacy. *Consulting Psychology Journal: Practice and Research*, 68(3), 189–207. DOI: [10.1037/cpb0000058](https://doi.org/10.1037/cpb0000058).
- De Haan, E., Gray, D. E., & Bonneywell, S. (2019). Executive coaching outcome research in a field setting: A near-randomized controlled trial study in a global healthcare corporation. *Academy of Management Learning & Education*, 18(4), 581–605. DOI: [10.5465/ame.2018.0158](https://doi.org/10.5465/ame.2018.0158).
- De Visser, E. J., Monfort, S. S., McKendrick, R., Smith, M. A., McKnight, P. E., Krueger, F., & Parasuraman, R. (2016). Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology: Applied*, 22(3), 331–349. DOI: [10.1037/xap0000092](https://doi.org/10.1037/xap0000092).

- Diller, S. J., Muehlberger, C., Braumandl, I., & Jonas, E. (2020). Supporting students with coaching or training depending on their basic psychological needs. *International Journal of Mentoring and Coaching in Education*, 10(1), 84–100. DOI: [10.1108/IJMCE-08-2020-0050](https://doi.org/10.1108/IJMCE-08-2020-0050).
- Duijts, S., Kant, I., van den Brandt, P., & Swaen, G. (2007). The compatibility between characteristics of employees at risk for sickness absence and components of a preventive coaching intervention. *International Journal of Evidence Based Coaching and Mentoring*, 5(1), 19–29. Available at: <https://radar.brookes.ac.uk/radar/items/ad870ed4-ac0c-4506-b178-911089345231/1/>.
- Ellis-Brush, K. (2021). Augmenting Coaching Practice through digital methods. *International Journal of Evidence Based Coaching & Mentoring*, S15. DOI: [10.24384/er2p-4857](https://doi.org/10.24384/er2p-4857).
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191 DOI: [10.3758/bf03193146](https://doi.org/10.3758/bf03193146).
- Fontes, A., & Dello Russo, S. (2021). An experimental field study on the effects of coaching: The mediating role of psychological capital. *Applied Psychology*, 70(2), 459–488. DOI: [10.1111/apps.12260](https://doi.org/10.1111/apps.12260).
- Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., & Rauws, M. (2018). Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial. *JMIR mental health*, 5(4), e64. DOI: [10.2196/mental.9782](https://doi.org/10.2196/mental.9782).
- Gaffney, H., Mansell, W., & Tai, S. (2019). Conversational agents in the treatment of mental health problems: mixed-method systematic review. *JMIR mental health*, 6(10), e14166. DOI: [10.2196/14166](https://doi.org/10.2196/14166).
- Geissler, H., Hasenbein, M., Kanatouri, S., & Wegener, R. (2014). E-coaching: Conceptual and empirical findings of a virtual coaching programme. *International Journal of Evidence Based Coaching and Mentoring*, 12(2), 165–187 Available at: <https://radar.brookes.ac.uk/radar/items/585eb4f9-19ce-49e1-b600-509fde1e18c0/1/>.
- Govindji, R., & Linley, P. A. (2007). Strengths use, self-concordance and well-being: Implications for strengths coaching and coaching psychologists. *International Coaching Psychology Review*, 2(2), 143–153.
- Grant, A. M. (2006). An Integrative Goal-Focused Approach to Executive Coaching. In D. R. Stober & A. M. Grant (Eds.), *Evidence based coaching handbook: Putting best practices to work for your clients* (p. 153–192). John Wiley & Sons Inc.
- Grant, A. M. (2011). Is it time to REGROW the GROW model? Issues related to teaching coaching session structures. *The Coaching Psychologist*, 7(2), 118–126. The British Psychological Society. Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.453.6410&rep=rep1&type=pdf>.
- Grant, A. M. (2012). An integrated model of goal-focused coaching: An evidence-based framework for teaching and practice. *International Coaching Psychology Review*, 7(2), 146–165.
- Grant, A. M. (2014). Autonomy support, relationship satisfaction and goal focus in the coach-coachee relationship: Which best predicts coaching success? *Coaching: An International Journal of Theory, Research and Practice*, 7(1), 18–38. DOI: [10.1080/17521882.2013.850106](https://doi.org/10.1080/17521882.2013.850106).
- Grant, A. M., Curtayne, L. & Burton, G. (2009). Executive coaching enhances goal attainment, resilience and workplace well-being: a randomized controlled study. *The Journal of Positive Psychology*, 4(5), 396–407. DOI: [10.1080/17439760902992456](https://doi.org/10.1080/17439760902992456).
- Grant, A. M., Passmore, J., Cavanagh, M. J., & Parker, H. M. (2010). The State of Play in Coaching Today: A Comprehensive Review of the Field. *International Review of Industrial and Organisational Psychology*, 25, 125–167. DOI: [10.1002/9780470661628.ch4](https://doi.org/10.1002/9780470661628.ch4).
- Graßmann, C., & Schermuly, C. C. (2020). Coaching With Artificial Intelligence: Concepts and Capabilities. *Human Resource Development Review*, 20(1), 106–126. DOI: [10.1177/1534484320982891](https://doi.org/10.1177/1534484320982891).
- Green, S., Grant, A. M., & Rynsaardt, J. (2020). Chapter 13: Evidence-based life coaching for senior high school students: Building hardiness and hope. In J. Passmore & D. Tee (Eds.), *Coaching Researched: A Coaching Psychology Reader* (pp. 257–268). John Wiley & Sons. DOI: [10.1002/9781119656913.ch13](https://doi.org/10.1002/9781119656913.ch13).
- Greer, S., Ramo, D., Chang, Y. J., Fu, M., Moskowitz, J., & Haritatos, J. (2019). Use of the chatbot “vivibot” to deliver positive psychology skills and promote well-being among young people after cancer treatment: randomized controlled feasibility trial. *JMIR mHealth and uHealth*, 7(10), e15018. DOI: [10.2196/15018](https://doi.org/10.2196/15018).
- Grover, S., & Furnham, A. (2016). Coaching as a developmental intervention in organisations: A systematic review of its effectiveness and the mechanisms underlying it. *PLoS one*, 11(7), e0159137. DOI: [10.1371/journal.pone.0159137](https://doi.org/10.1371/journal.pone.0159137).
- Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, 61(4), 5–14. DOI: [10.1177/0008125619864925](https://doi.org/10.1177/0008125619864925).
- Harkin, B., Webb, T. L., Chang, B. P. I., Prestwich, A., Conner, M., Kellar, I., Benn, Y., & Sheeran, P. (2016). Does monitoring goal progress promote goal attainment? A meta-analysis of the experimental evidence. *Psychological Bulletin*, 142(2), 198–229. DOI: [10.1037/bul0000025](https://doi.org/10.1037/bul0000025).
- International Coach Federation (ICF). (2020a). *2020 ICF Global Coaching Study*. Available at: <https://coachingfederation.org/research/global-coaching-study>.

- International Coach Federation (ICF). (2020b). *2020 ICF Global Coaching Study: Executive summary*. Available at: https://coachfederation.org/app/uploads/2020/09/FINAL_ICF_GCS2020_ExecutiveSummary.pdf.
- Jones, R. J., Woods, S. A., & Guillaume, Y. R. (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. DOI: [10.1111/joop.12119](https://doi.org/10.1111/joop.12119).
- Jones, R. J., Woods, S. A., & Zhou, Y. (2021). The effects of coachee personality and goal orientation on performance improvement following coaching: A controlled field experiment. *Applied Psychology*, 70(2), 420–458. DOI: [10.1111/apps.12218](https://doi.org/10.1111/apps.12218).
- Junker, S., Pömmmer, M., & Traut-Mattausch, E. (2020). The impact of cognitive-behavioural stress management coaching on changes in cognitive appraisal and the stress response: A field experiment. *Coaching: An International Journal of Theory, Research and Practice*, 1–18. DOI: [10.1080/17521882.2020.1831563](https://doi.org/10.1080/17521882.2020.1831563).
- Kamphorst, B.A. (2017). E-coaching systems: What they are, and what they aren't. *Personal and Ubiquitous Computing*, 21(4), 625–632. DOI: [10.1007/s00779-017-1020-6](https://doi.org/10.1007/s00779-017-1020-6).
- Koestner, R., Lekes, N., Powers, T.A., & Chicoine, E. (2002). Attaining personal goals: Self-concordance plus implementation intentions equals success. *Journal of Personality and Social Psychology*, 83(1), 231–244. DOI: [10.1037/0022-3514.83.1.231](https://doi.org/10.1037/0022-3514.83.1.231).
- Kombarakaran, F. A., Yang, J. A., Baker, M. N., & Fernandes, P. B. (2008). Executive coaching: It works! *Consulting Psychology Journal: Practice and Research*, 60(1), 78–90. DOI: [10.1037/1065-9293.60.1.78](https://doi.org/10.1037/1065-9293.60.1.78).
- Latham, G. P., & Locke, E. A. (2007). New Developments in and Directions for Goal-Setting Research. *European Psychologist*, 12(4), 290–300. DOI: [10.1027/1016-9040.12.4.290](https://doi.org/10.1027/1016-9040.12.4.290).
- Latham, G. P., & Seijts, G. H. (1999). The effects of proximal and distal goals on performance on a moderately complex task. *Journal of Organizational Behavior*, 20(4), 421–429. DOI: [10.1002/job.70](https://doi.org/10.1002/job.70).
- Lattie, E. G., Adkins, E. C., Winquist, N., Stiles-Shields, C., Wafford, Q. E., & Graham, A. K. (2019). Digital mental health interventions for depression, anxiety, and enhancement of psychological well-being among college students: systematic review. *Journal of medical Internet research*, 21(7), e12869. DOI: [10.2196/12869](https://doi.org/10.2196/12869).
- Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Prentice-Hall, Inc.
- Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American psychologist*, 57(9), 705. DOI: [10.1037/0003-066X.57.9.705](https://doi.org/10.1037/0003-066X.57.9.705).
- Mackie, D. (2007). Evaluating the effectiveness of executive coaching: Where are we now and where do we need to be? *Australian Psychologist* 42(4), 310–318. DOI: [10.1080/00050060701648217](https://doi.org/10.1080/00050060701648217).
- Niemiec, C. P., Ryan, R. M., & Deci, E. L. (2009). The path taken: Consequences of attaining intrinsic and extrinsic aspirations in post-college life. *Journal of research in personality*, 43(3), 291–306. DOI: [10.1016/j.jrp.2008.09.001](https://doi.org/10.1016/j.jrp.2008.09.001).
- Panetta, K. (2018, August 16). *Widespread artificial intelligence, biohacking, new platforms and immersive experiences dominate this year's Gartner Hype Cycle*. Gartner. Available at: <https://www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018/>.
- Riva, P., Sacchi, S., & Brambilla, M. (2015). Humanizing machines: Anthropomorphization of slot machines increases gambling. *Journal of Experimental Psychology: Applied*, 21(4), 313–325. DOI: [10.1037/xap0000057](https://doi.org/10.1037/xap0000057).
- Schwab, K. (2017). *The Fourth Industrial Revolution: What it Means and How to Respond*. Crown Publishing Group.
- Shanahan, M. (2015). *The Technological Singularity*. MIT Press.
- Sheldon, K. M., Kasser, T., Smith, K., & Share, T. (2002). Personal goals and psychological growth: Testing an intervention to enhance goal attainment and personality integration. *Journal of personality*, 70(1), 5–31. DOI: [10.1111/1467-6494.00176](https://doi.org/10.1111/1467-6494.00176).
- Sherwood, N. E., Crain, A. L., Martinson, B. C., Anderson, C. P., Hayes, M. G., Anderson, J. D., Senso, M. M., & Jeffery, R. W. (2013). Enhancing long-term weight loss maintenance: 2 year results from the Keep It Off randomized controlled trial. *Preventive Medicine*, 56(3–4), 171–177. DOI: [10.1016/j.ypmed.2012.12.014](https://doi.org/10.1016/j.ypmed.2012.12.014).
- Shoukry, H., & Cox, E. (2018). Coaching as a social process. *Management Learning*, 49(4), 413–428. DOI: [10.1177/2F1350507618762600](https://doi.org/10.1177/2F1350507618762600).
- Siau, K. L. & Yang, Y. (2017). Impact of Artificial Intelligence, Robotics, and Machine Learning on Sales and Marketing. *Midwest United States for Information Systems Conference Proceedings (MWAIIS 2017)*, 48. Available at: <http://aisel.aisnet.org/mwais2017/48>.
- Sin, N. L., & Lyubomirsky, S. (2009). Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: A practice-friendly meta-analysis. *Journal of clinical psychology*, 65(5), 467–487. DOI: [10.1002/jclp.20593](https://doi.org/10.1002/jclp.20593).
- Smith, B. W., Dalen, J., Wiggins, K., Tooley, E., Christopher, P., & Bernard, J. (2008). The Brief Resilience Scale: Assessing the Ability to Bounce Back. *International Journal of Behavioral Medicine*, 15(3), 194–200. DOI: [10.1080/10705500802222972](https://doi.org/10.1080/10705500802222972).

- Smither, J. W. (2011). Can psychotherapy research serve as a guide for research about executive coaching? An agenda for the next decade. *Journal of Business and Psychology*, 26(2), 135–145. DOI: [10.1007/s10869-011-9216-7](https://doi.org/10.1007/s10869-011-9216-7).
- Sonnentag, S. (2002). Performance, well-being, and self-regulation. In S. Sonnentag (Ed.), *Psychological management of individual performance* (pp. 405–424). Wiley.
- Spence, G. B., & Grant, A. M. (2007). Professional and peer life coaching and the enhancement of goal striving and well-being: An exploratory study. *The Journal of Positive Psychology*, 2(3), 185–194. DOI: [10.1080/17439760701228896](https://doi.org/10.1080/17439760701228896).
- Stewart-Brown, S., Tennant, A., Tennant, R., Platt, S., Parkinson, J., & Weich, S. (2009). Internal construct validity of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS): A Rasch analysis using data from the Scottish Health Education Population Survey. *Health and Quality of Life Outcomes*, 7(15). DOI: [10.1186/1477-7525-7-15](https://doi.org/10.1186/1477-7525-7-15).
- Telang, P. R., Kalia, A. K., Vukovic, M., Pandita, R., & Singh, M. P. (2018). A Conceptual Framework for Engineering Chatbots. *IEEE Internet Computing*, 22(6), 54–59. DOI: [10.1109/MIC.2018.2877827](https://doi.org/10.1109/MIC.2018.2877827).
- Terblanche, N. (2020). A design framework to create Artificial Intelligence Coaches. *International Journal of Evidence Based Coaching & Mentoring*, 18(2), 152–165. DOI: [10.24384/b7gs-3h05](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. DOI: [10.22316/poc/05.1.06](https://doi.org/10.22316/poc/05.1.06).
- Terblanche, N., Passmore, J., & Myburgh, J. (2021). African organisational coaching practice: Exploring approaches used, and the factors influencing coaches' fees. *South African Journal of Business Management*, 52(1), a2395. DOI: [10.4102/sajbm.v52i1.2395](https://doi.org/10.4102/sajbm.v52i1.2395).
- Theeboom, T., Beersma, B., & van Vianen, A. E. (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. DOI: [10.1080/17439760.2013.837499](https://doi.org/10.1080/17439760.2013.837499).
- Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work*. Routledge.
- West, L. & Milan, M. (2001), *The Reflecting Glass: professional coaching for leadership development*, Basingstoke: Palgrave.
- Whitmore, J. (2003) *Coaching for Performance*, London: Nicholas Brealey Publishing