



The Coach in the Loop

Tacit Knowledge, Professional Judgement and the Risk of Deskilled Coaching

About this paper

This paper asks what happens to coaching judgement when AI tools begin to make, or propose, the decisions that coaches have always made for themselves. It is written for coach developers, heads of coaching, club and governing body leaders, and the designers of coaching software, all of whom will shape how chat-based AI enters everyday coaching over the next few seasons. Its focus is the coach and the long-run health of a coaching workforce. The athlete appears mainly through the coach's decisions.

Within the Athleet.AI Working Papers, it carries the strand's single argument into the coach's own development. Earlier work in the strand treats athletic progress as non-linear, human and developmental, and warns against tools that flatten it. The same argument applies one level up, because coaching expertise is itself a developmental achievement, built slowly through decisions made, consequences observed and adjustments tried. A tool that takes over those decisions may help the athlete in front of the coach this week while weakening the coach who will stand in front of athletes for the next 20 years.

Editorial balance. Roughly two fifths of the paper reports evidence from coaching science and from human factors research on automation, about a third is analysis that carries that evidence across into AI-assisted coaching (an analogy I mark as such), and the remainder sets out my own proposed principles, which are offered for testing and have not been tested.

HOW TO READ THE EVIDENCE TAGS

Every substantive claim carries a tag stating what kind of evidence supports it. Two tags are combined where a claim rests on both.

PEER-REVIEWED A primary study in a peer-reviewed journal.

REVIEW A systematic review, meta-analysis or narrative review.

CONSENSUS A governing body position, consensus statement or official guidance.

PREPRINT A working paper or preprint not yet peer reviewed.

INDUSTRY A company announcement, press report or industry source.

SCHOLARLY A scholarly book, edited volume or law journal article.

PRACTICE Practitioner opinion, including the author's own.

EVIDENCE GAP The available evidence does not answer the question.

ATHLEET.AI The author's own framework or analysis, proposed rather than proven.

CONTENTS

Contents

FRONT MATTER

A note from the author	5
Executive summary	6

PART ONE · EXPERTISE AND THE ARRIVAL OF THE TOOL

1 A plan on a rainy Tuesday	9
2 What coaching expertise is	11

PART TWO · WHERE AI ENTERS AND WHAT IT CHANGES

3 Four roles for AI in the coaching process	14
4 Augmentation and erosion	16
5 The relationship and the qualification	19

PART THREE · KEEPING THE COACH IN THE LOOP

6 Keeping the coach in the loop	22
---------------------------------	----

END MATTER

References	24
About the author	25

ILLUSTRATIVE MATERIAL

Every organisation, person and scenario described as illustrative in this paper is invented. Published research and official documents are cited and numbered, and invented material is never presented as evidence.

Research closed 24 February 2023.
Sources and events are as they stood at that date.

CITATION

Forrest, P. (2023) *The Coach in the Loop: Tacit Knowledge, Professional Judgement and the Risk of Deskilled Coaching*. Athleet.AI Working Papers, AI in Coaching and Athlete Development, Paper 2. Athleet.AI, 27 February 2023. Available at: athleet.ai.

Comments and corrections are welcome and will be acknowledged in later versions.

PAPERS IN THIS STRAND

- Paper one. *Predicting the Unpredictable*. Why Algorithmic Talent Identification Struggles With Non-Linear Athlete Development. 25 January 2023.
- **Paper two. *The Coach in the Loop. Tacit Knowledge, Professional Judgement and the Risk of Deskilled Coaching*. 27 February 2023.**
- Paper three. *The Measured Child*. Datafication, Consent and Autonomy Across the Athlete Pathway. 19 April 2023.
- Paper four. *Maturation, Bias and the Training Set*. Developmental Fairness in AI-Supported Athlete Development. 24 May 2023.
- Paper five. *Load, Readiness and False Precision*. The Limits of Algorithmic Training Prescription in Long-Term Endurance Development. 21 June 2023.
- Paper six. *Democratised or Stratified?* AI, Competitive Fairness and the

Grassroots Coaching Gap. 26 July 2023.

A NOTE FROM THE AUTHOR

Why the planning hour matters

I have spent a good many winter evenings writing training sessions that did not work. Some were too hard, some were pitched at an athlete who existed only in my head, and a few were simply badly ordered. Almost everything I now know about planning I learned from those failures, and from the slower business of watching an athlete respond over weeks and asking myself why the plan and the athlete had parted company.

That is the experience I kept returning to when chat-based AI arrived in the coaching conversations I am part of this winter. Coaches I respect were pasting questions into ChatGPT and getting back plausible, tidy, week-by-week programmes in seconds. My first reaction was admiration for the fluency. My second, which has lasted longer, was a question about the evening hours those coaches would no longer spend. Nobody tells you, when you start coaching, that the planning hour is where much of your judgement gets built, because it does not feel like learning at the time. It feels like homework.

I also advise organisations on how they adopt technology, and I have seen enough automation projects to know that the benefits arrive quickly and the costs arrive late, usually in the form of people who can no longer do the thing the system was bought to help with. Sport has little research on this yet. Aviation, process control and medicine have a great deal, and I think coaching should borrow from them before the habits set.

None of this makes me against the tools. I use them, and I expect to use them more. My concern is narrower, and it is about who is holding the tool and at what point in their own development. This paper is my attempt to set that concern out carefully enough that others can test it and, where it is wrong, correct it.

Paul Forrest
Founder, Athleet.AI
27 February 2023

TWO PAGES THAT CIRCULATE ON THEIR OWN

Executive summary

Chat-based AI can now draft a coherent training plan in seconds, and coaches are already using it. The risk is that the tool removes the decisions through which coaches become expert, so the design question is who is using it and at what stage of their own development, and that answer should limit what the tool is allowed to decide.

Coaching science describes expertise as contextual, relational and heavily tacit, built through repeated decisions whose consequences a coach sees in real athletes. Human factors research has shown for four decades that automation which takes over a task can leave its operator less able to perform or supervise that task. Put together, these literatures suggest that AI decision support will affect experienced and developing coaches in opposite ways. Sport, as far as I can tell, measures neither effect yet.

1

Expert coaching rests on knowledge that coaches struggle to put into words.

Nash and Collins argue that expert coaches rely on tacit knowledge built from experience, and that much of it resists articulation.¹ **PEER-REVIEWED** Côté and Gilbert define effectiveness as the integrated use of professional, interpersonal and intrapersonal knowledge in a specific context.² **PEER-REVIEWED** A text model sees only the written part of coaching.

3

Automation research predicts that handing over decisions can erode the skill needed to overrule them.

Bainbridge set out how automating a task leaves operators out of practice for the moments when they must take over.⁵ **PEER-REVIEWED** Experimental work links automated decision aids to lower situation awareness and slower recovery when the aid fails.⁶ **PEER-REVIEWED** No study has yet tested this in coaching. The transfer therefore rests on analogy.

EVIDENCE GAP

5

Chat-based advice is moving into tools coaches already carry.

ChatGPT opened as a free research preview on 30 November 2022, and in the week of 6 February 2023 Microsoft and Google both announced chat-based search.^{8,9,10} **INDUSTRY** The sports and exercise medicine literature has already begun to debate what such tools mean for its own work.¹¹ **PRACTICE**

2

Coaching judgement is learned through decisions and their consequences.

Research on intuitive expertise indicates that reliable intuition forms only where the environment is regular enough to learn and where practice brings feedback.³ **PEER-REVIEWED** Work on professional judgement in coaching places the decision itself, made in interaction with learners and conditions, at the centre of practice.⁴ **PEER-REVIEWED**

4

The same tool may help an experienced coach and hinder a developing one.

An expert can use a drafted plan as a prompt and reject what does not fit, while a novice may have no basis for rejection. Complacency and automation bias have been found in experts as well as novices, so experience reduces the risk without removing it.⁷

REVIEW **PRACTICE**

6

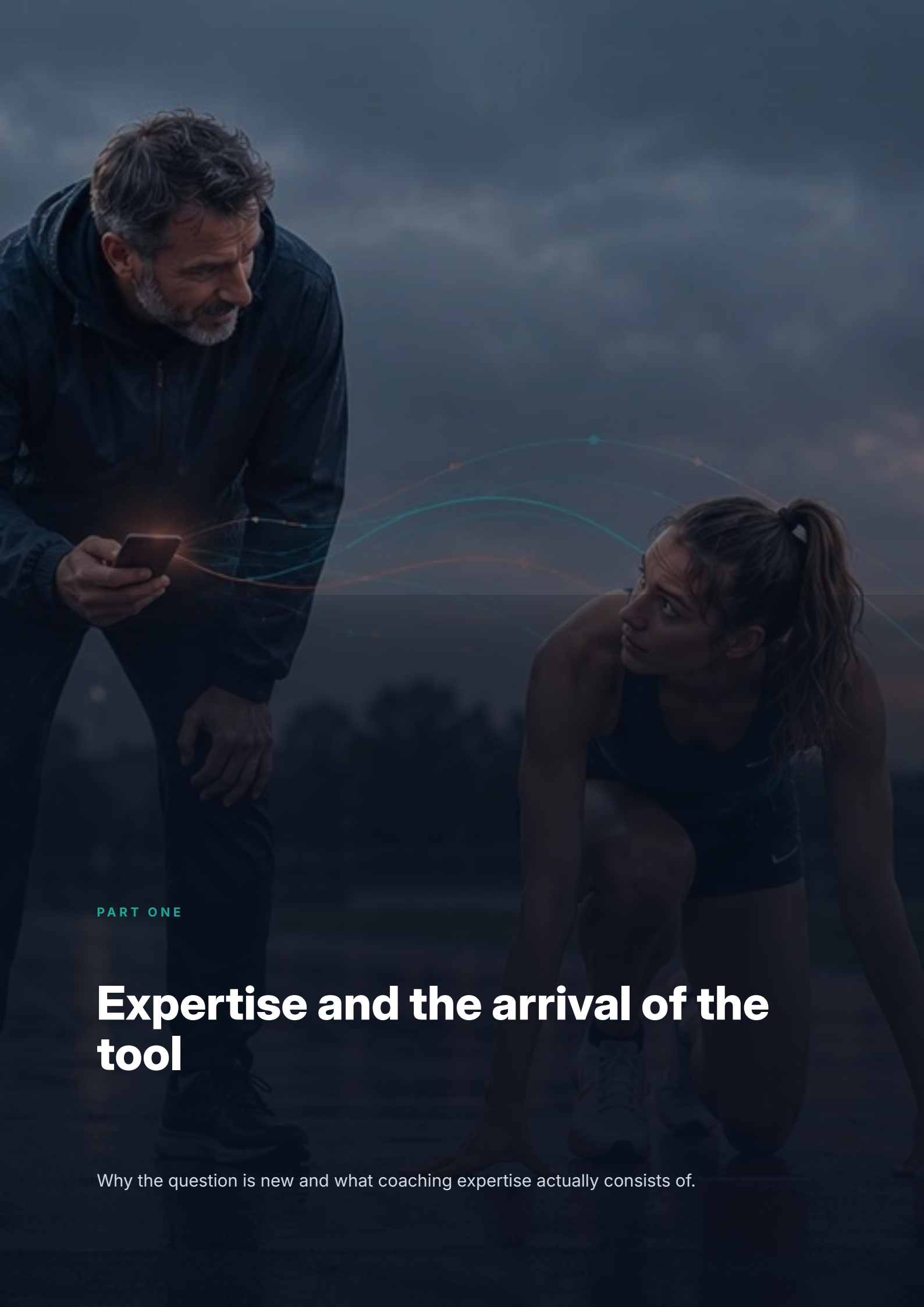
Coach education has not yet decided where AI belongs.

Coaching in England relies heavily on volunteers, the workforce the national coaching plan sets out to support, and they have the least time and the most to gain from a quick plan.¹² **CONSENSUS** I have found no published guidance on AI in coach qualifications.

EVIDENCE GAP**What follows for a club or governing body**

Clubs and governing bodies do not need to ban AI planning tools, and they would struggle to enforce a ban if they tried. They do need to decide, deliberately and soon, how such tools fit their coach development pathways. I propose that tools used by coaches in formation should show their reasoning, invite challenge and leave the final decision with the coach, and that the amount a tool decides should rise with the judgement the coach has shown.

Governing bodies should add AI use to coach education as a taught and assessed topic, and should keep AI-generated plans out of the assessed planning tasks through which early judgement is formed. Clubs should ask a plain question of every AI-assisted plan a volunteer uses with young athletes, namely who checked it and on what basis, and should record the answer.



PART ONE

Expertise and the arrival of the tool

Why the question is new and what coaching expertise actually consists of.

CHAPTER ONE

A plan on a rainy Tuesday

A chatbot can now write a sensible-looking training plan faster than a coach can find a pen. This chapter sets out why that changes the question sport should be asking about AI.

The public debate about AI in sport has mostly concerned what systems can measure and predict, from load and injury risk to talent and tactics. That debate matters, but it tends to treat the coach as the recipient of a better number. Chat-based AI changes the picture, because as well as supplying information it drafts the plan, which is the coach's own product. It does so in fluent prose that invites acceptance.

1.1 The plan arrives

ILLUSTRATIVE SCENARIO

The Tuesday session

A volunteer coach at a community athletics club (an invented composite, not a real person or club) has 40 minutes before a Tuesday session with a mixed group of athletes aged 13 and 14. It is raining, the usual lead coach is away, and the plan she was sent has not arrived. She types a request into a chatbot on her phone asking for a wet-weather sprint and conditioning session for young teenagers. Within seconds she has a warm-up, three drills, a set of hill repetitions and a cool-down, each with timings and coaching points. The session looks professional. She runs it. Two athletes who returned from growth-related knee pain last month complete the hill repetitions with the rest, because nothing in the plan or in her own experience prompted her to adjust it. Nobody is hurt that night. She saves the chat and uses the same approach the following week.

Nothing in the scenario is dramatic, and that is the point of it. The coach did something sensible under pressure, and the tool performed as designed. What did not happen is harder to see. She did not plan the session herself, so she did not rehearse the questions a planning coach learns to ask about who is in the group, what they did last week and which of them needs something different. The plan carried no trace of its assumptions, and she had no prompt to supply the knowledge that only she held about those two athletes.

1.2 Why the question is new

OpenAI released ChatGPT as a free research preview on 30 November 2022, and its own release note warned that the model sometimes gives answers that sound plausible but are wrong.⁸ **INDUSTRY** In the week of 6 February 2023, Microsoft opened a limited preview of a chat-enabled Bing and Google announced Bard for trusted test-

ers, with wider access promised in the coming weeks.^{9,10} **INDUSTRY** Taken together, these announcements suggest that chat-based advice is heading into the search tools that coaches already use between sessions. The coach's phone is where it will arrive.

In the weeks since, I have seen coaches and athletes share AI-generated training plans openly on social media and in coaching groups, often with approval and sometimes with amusement at the errors. **PRACTICE** The research community in sport is moving too. An editorial in *BMJ Open Sport and Exercise Medicine*, published on 16 February 2023, used ChatGPT to generate manuscript text and tested whether an AI text detector could identify it, as a way into the question of what such tools mean for the field.¹¹

PRACTICE That debate concerns authorship and research integrity. The question for coaching is different in kind, since it concerns what happens to the person whose job the tool partly performs.

1.3 What this paper argues

Coaching expertise is built by making decisions, seeing their consequences in a real athlete, and adjusting. An AI tool that makes the planning decision short-circuits that loop. For an experienced coach, AI can extend attention across a large group and free time for the relational work that athletes value. For a developing coach, the same tool can remove the practice that builds expertise in the first place. My claim is that the design and governance of these tools should turn on the user's stage of development, and that the amount a tool is allowed to decide should be set accordingly.

The argument proceeds in three parts. The first sets out what coaching science says expertise is. The second maps where AI can enter the coaching process and weighs the case for augmentation against the case for erosion, including what changes in the coach-athlete relationship and in coach education. The third proposes design principles and closes with open questions, several of which the available evidence cannot yet answer.

CHAPTER TWO

What coaching expertise is

Coaching science has spent two decades describing expertise as tacit, contextual and relational. Each of those properties bears directly on what an AI tool can and cannot take over.

Any claim that AI will augment or erode coaching expertise depends on an account of what that expertise consists of. Coaching science offers a reasonably consistent one, and it is less flattering to text-generating tools than the fluency of their output suggests.

2.1 Knowledge that resists being written down

Nash and Collins asked whether expert coaching is better understood as science or as art, and argued that much of what expert coaches know is tacit, meaning that it guides their decisions without being easily put into words.¹ **PEER-REVIEWED** Coaches in this account often describe their decisions as intuitive or as a feel for the situation, and the authors suggest that this intuition rests on accumulated experience. The practical consequence they draw is that coach education cannot rely on transmitting explicit rules alone.

Kahneman and Klein, writing from two traditions that had long disagreed about expert intuition, reached a shared position on when intuitive judgement can be trusted.³ **PEER-REVIEWED** They concluded that skilled intuition develops where the environment has enough regularity to be learned and where the person has had prolonged practice with timely feedback. Coaching partly meets those conditions. The same athlete responds to the same kind of session in broadly similar ways across a season, and the coach sees the response, although the feedback is often delayed, noisy and confounded by growth, school, sleep and mood.

Two things follow for AI. A language model learns from text, and the tacit part of coaching is by definition the part that is thinly represented in text. More importantly for this paper, if coaching intuition is formed through practice with feedback, then whatever removes the practice also removes the formation. The first point limits what the tool can know. The second concerns what the coach may never come to know.

2.2 Effectiveness as a situated achievement

Côté and Gilbert proposed an integrative definition of coaching effectiveness that has become a common reference point in coach education.² **PEER-REVIEWED** In their account, effective coaching is the consistent application of professional, interpersonal and intrapersonal knowledge to improve athletes' competence, confidence, connection and character, within a specific coaching context. The definition places context inside the concept. A session that is effective for an under-12 development group may be ineffective for a senior performance squad, and a plan that is sound in general can be poor for a particular group on a particular night.

TACIT KNOWLEDGE

Tacit knowledge is knowledge a practitioner uses without being able to state it fully. In coaching it includes reading an athlete's mood, sensing when a session has gone flat and knowing which instruction a particular child will hear.

Lyle's earlier framework treats coaching as a process, a planned and coordinated set of activities extending over time, and distinguishes the demands of participation coaching from those of performance coaching.¹³ **SCHOLARLY** The process view matters here because it treats a plan as one moment in a longer cycle of observation, intervention, monitoring and revision. A tool that produces the plan in isolation produces the most visible product of that cycle while leaving the cycle itself unsupported. The cycle is where the learning happens.

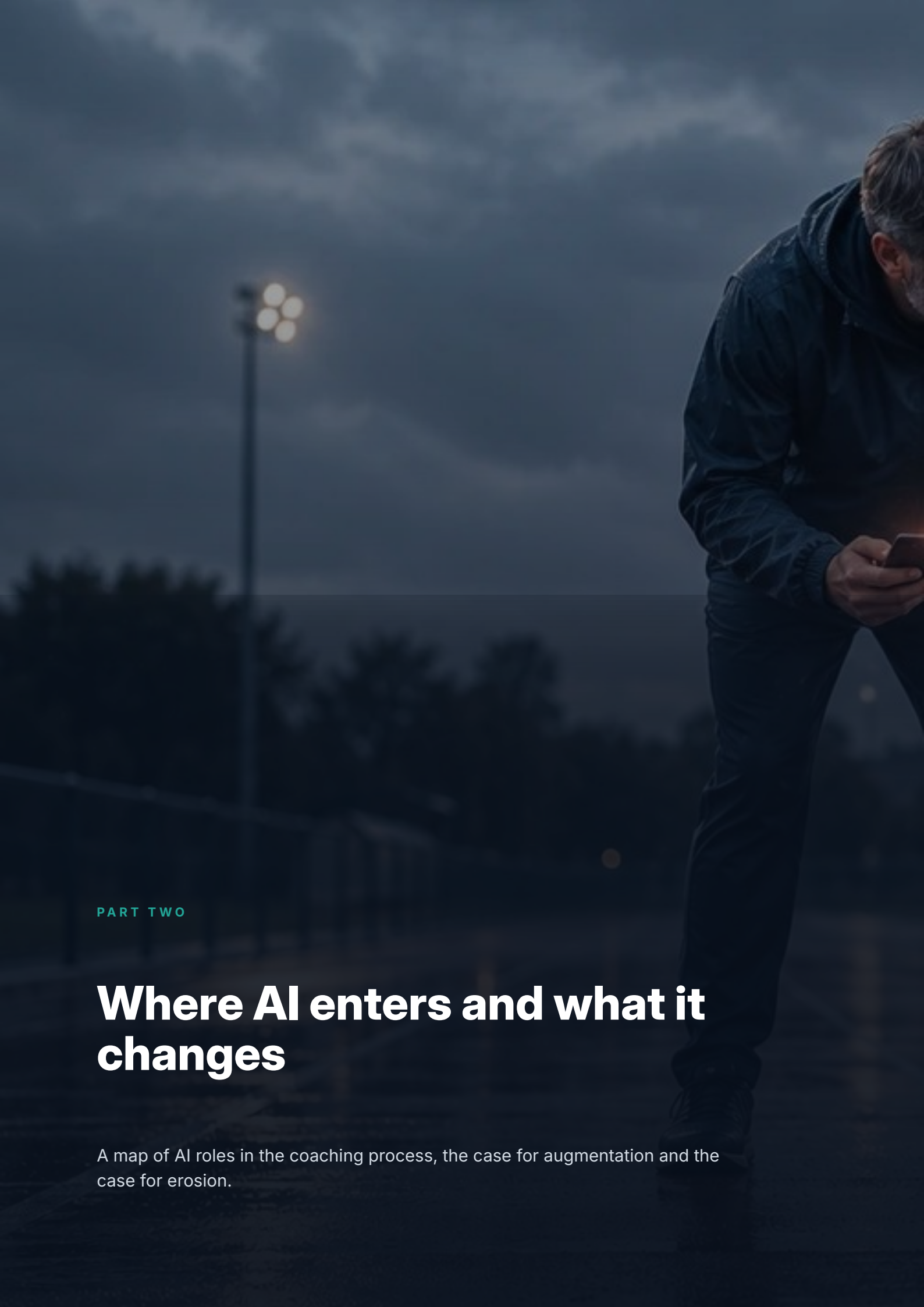
Several of the sources reviewed in this chapter are conceptual papers, which argue from experience and from other literatures, and the empirical base under them is thinner than their influence in coach education might suggest. **EVIDENCE GAP** I rely on them because they are the best account available and because their central claims are consistent with each other and with the wider research on expertise. Readers who doubt it will find what follows weaker.

2.3 Judgement and decision making

A third strand treats coaching as professional judgement and decision making. Martindale and Collins, working first in applied sport psychology, argued that a practitioner's decisions should be judged against an explicitly stated intention for impact, so that the reason behind an intervention can be examined as well as the intervention itself.¹⁴ **PEER-REVIEWED** Collins and Collins carried the approach into adventure sports coaching and described judgement that emerges through interaction with learners and with a changing environment, combining deliberate planning with adjustment in the moment.⁴ **PEER-REVIEWED**

Abraham and Collins had earlier made the case that coaching science should study how coaches actually decide, drawing on decision-making research, and should treat coaching behaviour as the visible output of those decisions.¹⁵ **PEER-REVIEWED** That emphasis on decision-making gives this paper its central concern. If expertise lives in the quality of decisions, and if AI tools increasingly make or pre-empt those decisions, then the question of what AI does to coaching is at root a question about who is deciding.

Coaching expertise is formed in the gap between a plan and an athlete. The literature reviewed here converges on a picture of expertise built through repeated decisions, taken in context, whose consequences the coach observes and learns from. Any tool that closes that gap on the coach's behalf deserves scrutiny for what it removes from the coach's learning as well as for what it adds to the athlete's session.



PART TWO

Where AI enters and what it changes

A map of AI roles in the coaching process, the case for augmentation and the case for erosion.

CHAPTER THREE

Four roles for AI in the coaching process

AI can enter coaching as a reference, an adviser, a co-planner or a substitute. The same software can play any of these roles, depending on how a coach uses it.

Discussion of AI in coaching tends to treat the tool as a single thing. In practice the effect depends on which decisions the tool touches and how far it goes in making them. Human factors research, which has studied automation in cockpits, control rooms and hospital wards for several decades, offers a useful way of separating these questions from one another.

3.1 Borrowing a map from automation research

Parasuraman, Sheridan and Wickens proposed that automation can apply to four broad functions, namely acquiring information, analysing it, selecting a decision or action, and implementing that action, and that each function can be automated to a different degree.¹⁶ **PEER-REVIEWED** Their model was built for systems such as air traffic control, and its central insight transfers readily. Automating information gathering has different consequences from automating the choice itself, and designers can choose a level for each function separately.

The coaching process described by Lyle maps onto these functions without much strain.¹³ **SCHOLARLY** A coach gathers information about athletes, interprets it, chooses a plan and then delivers it in person. Chat-based tools currently act mostly on the second and third functions, and they are least able to act on the fourth, since delivery remains a human activity in a physical place. That leaves the choice of plan, which is the function most closely tied to judgement, as the one most exposed.

3.2 Reference, adviser, co-planner, substitute

I distinguish four roles that an AI tool can play in the planning decision. **ATHLEET.AI** They sit on a gradient of how much of the decision the tool takes, and the gradient is set as much by the coach's behaviour as by the software.

Role	What the tool does	Where the decision sits	Main risk
Reference	Answers factual questions about exercises, loads or rules	Wholly with the coach	Plausible but wrong answers
Adviser	Comments on a plan the coach has drafted	With the coach, informed by the tool	Over-deference to fluent critique
Co-planner	Drafts options that the coach adapts	Shared, with the coach editing	Anchoring on the first draft

Role	What the tool does	Where the decision sits	Main risk
Substitute	Produces the plan that is delivered unchanged	With the tool in practice	Loss of judgement and of accountability

TABLE 1 Four roles for AI in the planning decision, ordered by how much of the decision the tool takes. Source: author's framework, drawing on Parasuraman, Sheridan and Wickens (2000).

The volunteer in the opening scenario used a general chatbot as a substitute, although nothing in the software required her to. A more experienced coach might have used the identical output as a co-planner, striking out the hill repetitions for the two returning athletes within a minute of reading it. The software is the same in both cases. What differs is the knowledge each user brings to it.

The role can also shift within a single coach from week to week. A coach who begins by asking a tool to critique a draft may, on a busy week, accept its redraft unchanged and slide from adviser to substitute without noticing. That drift is the practical form of the complacency described in the next chapter, and it is hard to see from the inside because each individual step feels reasonable. **PRACTICE**

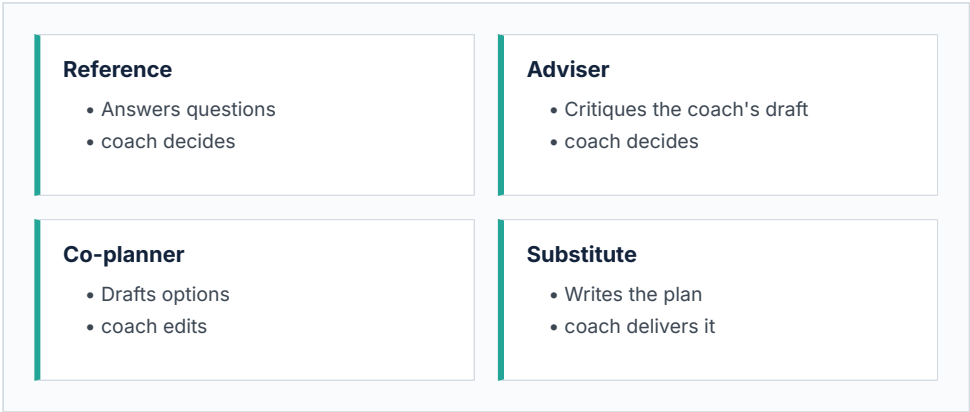


FIGURE 1 The same tool can occupy any of four roles, depending on its user. Source: author's framework.

This role-based view has a practical merit for clubs. It directs attention to how tools are used, which is where both the benefit and the risk sit.

CHAPTER FOUR

Augmentation and erosion

For an experienced coach, AI can extend attention and free time for the relationship. For a developing coach, it can remove the practice that builds judgement.

The two cases are often presented as opposing forecasts, with one camp predicting better coaching and the other worse. I think both are true at once, for different coaches.

4.1 The augmentation case

Coaches of large groups face a problem of attention. A volunteer running a session for 30 children cannot individualise much, and a performance coach with a full squad has limited hours for planning and monitoring. An AI tool that drafts sessions, summarises training notes or suggests progressions could return time to the coach.

PRACTICE

Where that time goes matters. Jowett's model describes the coach-athlete relationship in terms of closeness, commitment and complementarity, together with co-orientation, meaning the degree to which coach and athlete share an understanding of each other.¹⁷ **SCHOLARLY** None of these can be delegated to software. If AI takes over some of the written planning and the coach spends the saved time talking with athletes, watching them closely and adjusting the session in person, the relationship may benefit. That is my inference, and I know of no study in sport that has tested it.

EVIDENCE GAP

The augmentation case is strongest for the experienced coach. Such a coach already holds the tacit knowledge the tool lacks, and can read a drafted plan quickly, see where it misfits the group and repair it, so the tool works as an extension of existing judgement. Kahneman and Klein's account suggests why this coach can be trusted to overrule the tool, since their intuitions have been formed through long practice with feedback in a reasonably regular environment.³ **PEER-REVIEWED**

The adviser role may suit experienced coaches best of all. A coach who drafts a block and then asks a tool to find weaknesses in it, such as a missing recovery day or an abrupt rise in volume, gets a second reader at no cost and keeps the decision entirely. Used that way, the tool plays something like the part of a colleague looking over a plan, a part that many volunteer and part-time coaches rarely have anyone to fill. I think this is the most promising use in the near term, although I hold that view on the basis of practice and argument alone. **PRACTICE**

4.2 The erosion case

Bainbridge's short paper on the ironies of automation made an argument that has held up for four decades.⁵ **PEER-REVIEWED** Designers automate the tasks they can, which leaves the human operator with the tasks that could not be automated plus the

job of monitoring the system and taking over when it fails. Yet an operator who rarely performs the task loses the skill, and an operator who has only ever monitored may never acquire it. The irony is that automation makes the human contribution harder at the moment it becomes necessary.

Later experimental work gave the argument a mechanism. Endsley and Kiris studied operators working with an expert system on a navigation task and found that reduced situation awareness under automation accompanied slower decisions when the system failed, attributing the loss mainly to a shift from active to passive processing.⁶

PEER-REVIEWED They also reported that the level of control left to the operator moderated the loss, which is encouraging for design because it implies the effect can be reduced. Parasuraman and Manzey reviewed the related phenomena of complacency and automation bias, in which people under-check automated output or follow it against other evidence, and concluded that these effects appear in experts as well as novices and are not reliably removed by training alone.⁷ **REVIEW**

Carried into coaching, the argument is uncomfortable. A coach who routinely accepts AI plans is placed in the monitoring role Bainbridge described. When the plan is wrong for a particular athlete, as it was for the two returning athletes in the scenario, the coach must notice and intervene. That demands the planning judgement the tool has been doing for them. **PRACTICE**

BY ANALOGY

None of these studies involved coaches. The transfer from control rooms and cockpits to a sports field is my own inference, and it should be tested before it is treated as a finding.

4.3 The developing coach

The erosion case bears hardest on coaches who are still forming their judgement. An experienced coach who hands planning to a tool may lose some fluency over time, but has a store of experience to draw on. A new coach who plans with AI from the start may never build that store, because the practice with feedback that Kahneman and Klein identify as the condition for skilled intuition has been bypassed.³ **PEER-REVIEWED** The coach still sees the athlete respond, but has not made the decision whose consequence they are watching, and so has little to learn from the outcome. My own view is that this is the most serious long-run risk AI poses to coaching, and it is also the least visible, since it would show only years later in a workforce that plans competently with a tool and poorly without one.

1	Decide	The coach chooses a session for a known group.
2	Deliver	The coach runs it and watches the athletes respond.
3	Compare	The coach sets the response against what they expected.
4	Adjust	The coach revises their model of the athletes and the next plan.
5	Interrupted	If a tool made the decision, the comparison has little to teach.

FIGURE 2 How coaching judgement is built, and where an AI plan can interrupt the loop. Source: author's synthesis of Nash and Collins (2006) and Kahneman and Klein (2009).

The UK context makes this pressing. Sport England's coaching plan for 2017 to 2021 addresses a coaching workforce in which volunteers play a large part, and it sets out to make coaching more responsive to participants and more accessible to new coaches.¹² **CONSENSUS** Volunteers have the least time to plan, which makes a quick AI plan attractive, and many have little formal coach education to fall back on when the plan is wrong. Whether that combination leads to wide substitute use is an empirical question I cannot answer from published data. **EVIDENCE GAP**

CHAPTER FIVE

The relationship and the qualification

AI planning changes what athletes may believe about their coach and what coach education must teach. Neither question has been studied in sport.

The previous chapter considered the coach's own expertise. This one turns to two things that surround it, the athlete's trust and the qualification system that forms new coaches, and then to the harder question of responsibility.

5.1 What the athlete believes about the plan

Jowett's co-orientation dimension concerns how far coach and athlete understand each other's views and feelings.¹⁷ **SCHOLARLY** A plan is partly a message within that relationship. When an athlete receives a session, they may read it as evidence that the coach has thought about them in particular. If the athlete suspects the session came from a chatbot, that reading may weaken, and with it some of the commitment the model describes. The athlete may also wonder whether the coach is able to explain the session, and a coach who cannot explain a plan they did not write is in an awkward position.

I have found no study in sport of how athletes respond to learning that a plan was machine-generated. **EVIDENCE GAP** The question matters most in youth sport, where the relationship does developmental work that goes beyond the training effect of any session. I would welcome evidence either way. My working assumption, which I hold loosely, is that disclosure paired with visible adaptation (the coach showing what they changed and why) would preserve trust better than either concealment or unedited use.

5.2 Coach education and the timing of AI

Abraham and Collins argued that coach education should develop decision-making, since behaviour follows from it.¹⁵ **PEER-REVIEWED** If that is right, then the assessed planning tasks in early qualifications are where judgement is first formed and first examined. An AI-generated plan submitted for such a task would defeat its purpose, much as a calculator defeats a test of mental arithmetic, because the thing being built is the capacity the tool replaces.

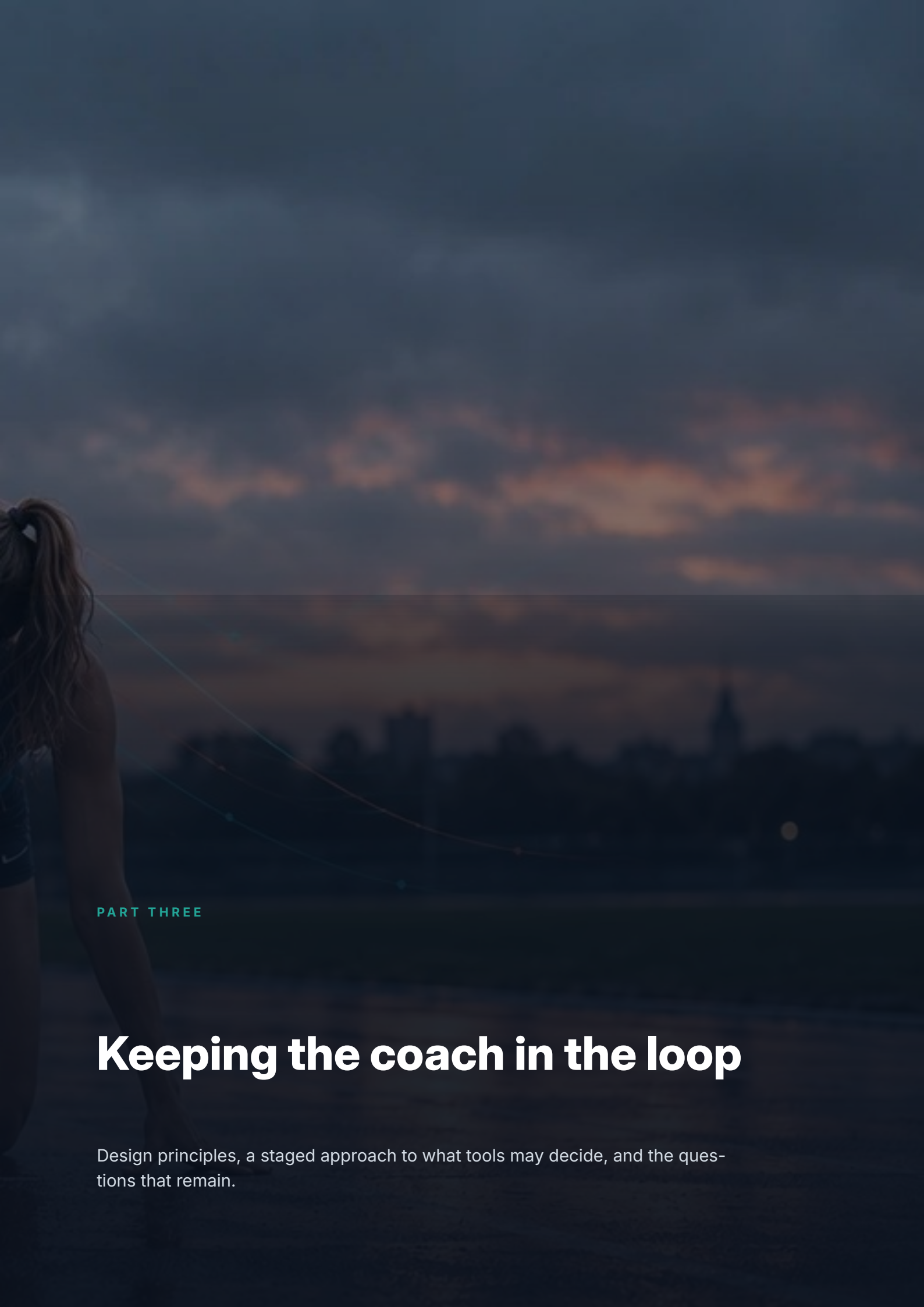
AI still has a place in coach education. Coaches will use these tools whether or not qualifications mention them, and they need to learn how to question AI output, spot its errors and adapt it to a group. My proposal is a sequence. Early qualifications should keep AI out of assessed planning while teaching its limitations explicitly, and later qualifications should include AI as a co-planner under observation, where the assessor can see how the coach challenges and edits the draft. **ATHLEET.AI** I have found no governing body guidance on this as of February 2023. **EVIDENCE GAP**

Reflection offers a bridge between the two stages. A developing coach could be asked to write a plan unaided, then generate one with a tool, and then explain in a short reflective note where the two differ and which they would run with a named group. The exercise uses the tool to expose the coach's reasoning to scrutiny, and it gives the assessor evidence of judgement that a finished plan alone cannot supply. **ATHLEET.AI**

5.3 Responsibility when the plan harms

Suppose the scenario had ended differently, and one of the returning athletes had aggravated the knee problem. The question of who was responsible has legal dimensions that this paper does not address and that clubs should take to qualified advisers. The governance question is simpler and belongs to the club. A club that allows volunteers to use AI plans with young athletes should know that it does so, should expect someone with relevant knowledge to check those plans, and should be able to show afterwards who checked them. **ATHLEET.AI**

The OpenAI release note acknowledged that its model can give confident answers that are wrong.⁸ **INDUSTRY** A plan that looks professional carries that same risk, and fluent presentation may make the risk harder to notice, which is the pattern Parasuraman and Manzey describe under automation bias.⁷ **REVIEW** The club's safeguard is the coach's judgement, and the club therefore has a direct interest in whether that judgement is being formed or bypassed. Whether insurers or governing bodies will begin asking about AI use in session planning remains an open question.

A person with long hair in a ponytail is seen from the side, looking out at a city skyline at sunset. The sky is dark with some orange and pink clouds. The city skyline is visible in the distance, with some buildings and a bridge. The person is wearing a dark top.

PART THREE

Keeping the coach in the loop

Design principles, a staged approach to what tools may decide, and the questions that remain.

CHAPTER SIX

Keeping the coach in the loop

I propose four design principles and a staged approach in which the share of the decision a tool takes grows with the judgement the coach has shown.

The preceding chapters point towards a design stance. Tools used in coaching should support the formation of judgement as well as its exercise, and the level of automation should fit the user. What follows is my own framework, offered for testing.

ATHLEET.AI

6.1 Four design principles

The first principle is that a tool should show its reasoning. A plan that states its assumptions about age, training history, recent injury and session aim gives the coach something to check against what they know. A plan that states none gives them nothing, and the opening scenario turned on precisely that absence. ATHLEET.AI

The second is that a tool should invite challenge. It can ask the coach what it does not know, prompt them to name any athlete who needs a different load and offer alternatives in place of a single answer. Endsley and Kiris found that keeping the operator more involved moderated the loss of awareness, which gives this principle some support by analogy.⁶ PEER-REVIEWED ATHLEET.AI

The third is that the coach remains the owner of the decision. The tool drafts or advises, and a named coach accepts, edits or rejects, with a record of what was changed. Ownership is partly a matter of accountability and partly a matter of learning, since a coach who has consciously accepted a plan has made a decision they can later learn from. ATHLEET.AI

The fourth is that what the tool decides should depend on the coach's stage. Parasuraman, Sheridan and Wickens showed that the level of automation can be set separately for each function.¹⁶ PEER-REVIEWED A coaching tool could apply the same idea to users, offering reference and adviser roles to coaches in formation and co-planning to coaches whose judgement has been observed and assessed. ATHLEET.AI

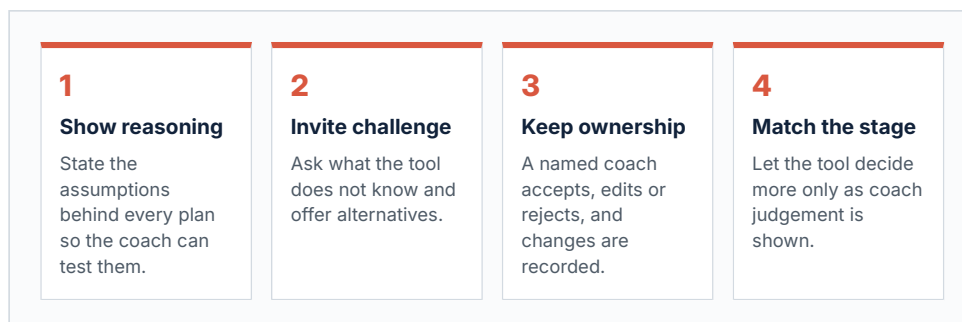


FIGURE 3 Four design principles for AI tools used in coaching. Source: author's framework.

6.2 A staged approach for clubs and governing bodies

Principles need a structure to live in. I recommend that clubs and governing bodies tie permitted AI roles to existing coach development stages. **ATHLEET.AI** A coach in their first qualification would use AI as a reference and, under a mentor's eye, as an adviser on plans they have drafted themselves. A qualified coach with some seasons of experience could use AI as a co-planner, with the expectation that they can explain every change they made to a draft. Senior coaches would decide for themselves, since their judgement is the safeguard the whole arrangement relies on, though Parasuraman and Manzey's finding that experts are also prone to automation bias argues for periodic planning without the tool even at this level.⁷ **REVIEW**

No stage in this scheme permits the substitute role with young athletes. I hold that line firmly, on grounds of safeguarding and of coach development together.

Mentoring could carry much of the weight here. A mentor who reviews an AI-assisted plan with a new coach, asking what they would change and why, turns the tool's output into a teaching object. Whether that works as well as planning from a blank page is unknown, and it is the kind of question a small controlled study in a coach education programme could answer. **EVIDENCE GAP**

WHY NOT A BAN

A ban on AI planning would be hard to enforce and would push use out of sight. Tying permitted roles to coach stages keeps use visible and gives coach developers something to observe.

6.3 Conclusion and open questions

Chat-based AI has arrived in coaching faster than coaching science has been able to study it. The existing literatures on coaching expertise and on automation, read together, suggest a specific risk, one that attention to the quality of AI plans can obscure. A tool that makes the planning decision may produce a good session tonight and a weaker coach in five years, and the effect will fall hardest on developing and volunteer coaches, who most need the time the tool saves. The response I propose keeps the coach as the decision owner, asks tools to show their reasoning and invite challenge, and ties what a tool may decide to the judgement its user has shown.

ATHLEET.AI

Several questions remain open, and I set them out as a list because each needs its own study.

- Should AI tools behave differently for a newly qualified coach and a Level 4 coach, and who would verify which is which?
- Can a tool be designed to teach judgement, for example by withholding its plan until the coach has drafted one?
- Do athletes, and the parents of young athletes, trust a coach less when they learn a plan was AI-generated, and does visible adaptation change that?
- Who is responsible, in governance terms, when a volunteer follows an AI plan that injures a young athlete, and what records would a club need to show it acted reasonably?

The last of these may be the one that moves fastest, since it is the one most likely to be tested by an incident before it is tested by research. I would prefer the research to come first, although the pace of the past three months gives little reason to expect it.

END MATTER

References

1. Nash, C. and Collins, D. (2006) Tacit knowledge in expert coaching: science or art? *Quest* 58(4), 465-477.
2. Côté, J. and Gilbert, W. (2009) An integrative definition of coaching effectiveness and expertise. *International Journal of Sports Science and Coaching* 4(3), 307-323.
3. Kahneman, D. and Klein, G. (2009) Conditions for intuitive expertise: a failure to disagree. *American Psychologist* 64(6), 515-526.
4. Collins, L. and Collins, D. (2016) Professional judgement and decision-making in adventure sports coaching: the role of interaction. *Journal of Sports Sciences* 34(13), 1231-1239.
5. Bainbridge, L. (1983) Ironies of automation. *Automatica* 19(6), 775-779.
6. Endsley, M.R. and Kiris, E.O. (1995) The out-of-the-loop performance problem and level of control in automation. *Human Factors* 37(2), 381-394.
7. Parasuraman, R. and Manzey, D.H. (2010) Complacency and bias in human use of automation: an attentional integration. *Human Factors* 52(3), 381-410.
8. OpenAI (2022) ChatGPT: optimizing language models for dialogue. *OpenAI Blog*, 30 November. Available at: openai.com/blog/chatgpt.
9. Mehdi, Y. (2023) Reinventing search with a new AI-powered Microsoft Bing and Edge, your copilot for the web. *Official Microsoft Blog*, 7 February. Available at: blogs.microsoft.com.
10. Pichai, S. (2023) An important next step on our AI journey. *The Keyword* (Google), 6 February. Available at: blog.google.
11. Anderson, N., Belavy, D.L., Perle, S.M., Hendricks, S., Hespanhol, L., Verhagen, E. and Memon, A.R. (2023) AI did not write this manuscript, or did it? Can we trick the AI text detector into generated texts? The potential future of ChatGPT and AI in Sports and Exercise Medicine manuscript generation. *BMJ Open Sport and Exercise Medicine* 9(1), e001568.
12. Sport England (2016) *Coaching in an Active Nation: The Coaching Plan for England 2017-21*. London: Sport England.
13. Lyle, J. (2002) *Sports Coaching Concepts: A Framework for Coaches' Behaviour*. London: Routledge.
14. Martindale, A. and Collins, D. (2005) Professional judgment and decision making: the role of intention for impact. *The Sport Psychologist* 19(3), 303-317.
15. Abraham, A. and Collins, D. (2011) Taking the next step: ways forward for coaching science. *Quest* 63(4), 366-384.
16. Parasuraman, R., Sheridan, T.B. and Wickens, C.D. (2000) A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans* 30(3), 286-297.
17. Jowett, S. (2007) Interdependence analysis and the 3+1Cs in the coach-athlete relationship. In Jowett, S. and Lavallee, D. (eds) *Social Psychology in Sport*. Champaign, IL: Human Kinetics, 15-28.

END MATTER

About the author

Paul Forrest is the founder of Athleet.AI. He advises boards and leadership teams on the adoption and governance of artificial intelligence, drawing on a background in business process reengineering and on hands-on work building machine learning systems, including small, domain-specific language models. He has served as a special adviser to UK government on the Global Entrepreneur Programme and works as a board member and business adviser.

He is also a qualified athletics coach who coaches sprint, hurdle and endurance athletes as a volunteer, and much of his thinking about how people learn, decide and improve comes from the track.

This paper is part of the Athleet.AI Working Papers, AI in Coaching and Athlete Development. The views expressed are the author's own.

athleet.ai

The Coach in the Loop

Athleet.AI Working Papers, AI in Coaching and Athlete Development
Paper two of six · 27 February 2023