Universiteit Leiden

Master Creative Intelligence and Technology

THE PROVOCATIVE COACH. Ego-Threatening versus Autonomy-Supportive Feedback in a Real-Time Running Coaching Application: A Pilot Study

Name:	Timothy Schuur
Student ID:	4383842
Date:	01/07/2026

1st supervisor:	Edwin van der Heide
2nd supervisor:	Bart Roelands

Master's Thesis in Creative Intelligence and Technology

Leiden Institute of Advanced Computer Science (LIACS)
Leiden University
Einsteinweg 55
2333 CC Leiden
The Netherlands

Contents

Abstract

Real-time digital coaching for endurance running has predominantly relied on supportive or neutral feedback styles. This pilot study examines a counterintuitive alternative: ego-threatening feedback, which directly challenges the athlete's competence relative to their goal. Drawing on Psychological Reactance Theory and the Dual Process framework, the thesis proposes that such cues may trigger an automatic fight impulse that reduces reliance on deliberate effort regulation, potentially improving pace under acute fatigue. General Self-Efficacy is theorised to moderate this effect.

A within-subjects, counterbalanced field study compared autonomy-supportive and ego-threatening auditory coaching cues across two 5-kilometre runs. Sixteen recreational runners were recruited; nine were excluded post-hoc due to insufficient pace challenge, structural pace deficits, or incomplete data, resulting in a sample of $n = 7$. Outcome variables included average running pace (GPS-derived), perceived cognitive processing, perceived psychological stress, and motivational reactance.

No hypothesis was statistically confirmed. Average pace did not differ significantly between conditions ($M_{\text{AS}} = 311.7$ s/km vs. $M_{\text{ET}} = 314.6$ s/km, $p = .203$, $r = .48$), with five of seven participants running faster under the autonomy-supportive condition. The cognitive processing scale failed under ego-threatening conditions ($\alpha = -3.525$), leaving H2 unresolved. Perceived psychological stress was directionally higher under ego-threatening feedback for all participants who showed any difference ($r = .70$), though this did not reach statistical significance ($p = .066$). Qualitative data suggest ego-threatening cues may trigger a brief acceleration that fades before accumulating into a lower overall pace.

The primary contribution of this study is methodological. Testing a novel and psychologically complex intervention in an uncontrolled field setting before its basic mechanism has been established under controlled conditions proved overambitious. Future work should prioritise a treadmill-based protocol that eliminates GPS noise and environmental confounds, with designs capable of capturing cue-level behavioural responses, strict counterbalancing, and a sample with verified pace-goal experience.

1. Introduction

Maintaining a consistent pace is often a challenge in endurance sports. Buman et al. (2008) report that 43–53% of recreational marathon runners experience a complete collapse in pacing. This phenomenon is known as "hitting the wall". As athletes try to keep up a challenging pace, they naturally experience increasing levels of exertion. Consequently, the athlete has to continuously regulate effort. During such a run, a constant battle occurs within the athlete. On the one hand, the athlete experiences the automatic urge to slow down due to fatigue, while on the other hand, they must make the deliberate decision to maintain their target pace guided by performance goals (Micklewright et al., 2017).

An increasing number of runners are using wearable technologies and digital coaching applications to help them manage this challenge (Stragier et al., 2016). Although many interventions can track physiological data and provide feedback, they often lack effectiveness when physical and cognitive demands increase during the run (Schmidt et al., 2015). As established in existing literature, the mental and physical demands of maintaining a challenging pace inherently limit an athlete's cognitive resources. This cognitive load makes the deliberate processing of detailed feedback more difficult (Marcora, Staiano, & Manning, 2009; Brick, MacIntyre, & Campbell, 2014).

It is known that autonomy-supportive coaching enhances long-term motivation and well-being (Mageau & Vallerand, 2003; Mossman et al., 2022). Nevertheless, its effectiveness as a real-time performance-improving digital intervention under the conditions of increased cognitive load during a run can be called into question. This highlights a gap in the current literature. Additionally, there is limited understanding of how alternative tones of feedback may function when an athlete's capacity for deliberate processing is reduced.

In order to address this gap, this thesis explores a counterintuitive approach to real-time digital coaching by examining the potential of ego-threatening feedback. In this thesis, we define ego-threatening feedback strictly as direct, competence-challenging feedback that confronts an athlete's current performance relative to their goal. Instead of giving an autonomy-supportive cue such as: "You're dropping pace. If it feels good and you want to push through, try to bring the pace back up", an ego-threatening cue directly challenges: "You're losing it. I did not expect you to fold this early. Pick up your pace now". With this approach, we aim to mitigate the need for deliberative rationalisation and elicit an immediate, corrective burst of effort. This thesis focuses specifically on recreational athletes with performance goals. More specifically, runners who have set a challenging target pace and possess sufficient general self-efficacy to engage competitively with feedback, but who are not elite athletes with professional coaching structures.

Additionally, giving this kind of feedback through a digital system and not a human may change how it is received. Non-human digital feedback may reduce concerns associated with social evaluation more effectively than human-delivered criticism (Dickerson & Kemeny, 2004). Thereby, the digital feedback potentially enables athletes to respond to more direct or confrontational cues without experiencing the same level of interpersonal pressure.

1.1 Research Aim and Outline

The primary objective of this thesis is to examine, in an exploratory pilot study, whether a digital coach can effectively assist runners in achieving a faster average pace above a challenging minimum through real-time ego-threatening feedback compared to autonomy-supportive feedback. To evaluate this intervention method, ego-threatening feedback will be compared to conventional autonomy-supportive feedback.

The impact of both feedback styles will be evaluated through three primary outcome variables: objective running performance (operationalised as average pace in seconds per kilometre), perceived cognitive processing, and perceived psychological stress. RPE serves as a validation measure, confirming that the individually set minimum pace induces a genuine effort state across conditions. General self-efficacy (GSE) and motivational reactance are included as explanatory variables to contextualise the psychological disposition and proposed mechanism through which feedback may operate. In this thesis we not only research whether ego-threatening feedback improves performance, but also at what cognitive and psychological cost. These metrics are selected to address exactly that.

The remainder of this thesis is structured as follows: Chapter 2 examines the literature through the Dual Process framework. Chapter 3 addresses the problem statement, research question, and hypotheses, building on the theoretical foundation that was set in Chapter 2. Chapter 4 covers the methodology. The results are presented in Chapter 5 and discussed in Chapter 6. Finally, Chapter 7 draws conclusions and reflects on the implications of this study.

2. Literature Review

2.1 The Regulation of Fatigue and Pacing in Endurance Running

In this section, we look at the physiological and cognitive processes that affect how a runner regulates pace and distributes effort over time under fatigue. Before we get into how the brain processes these fatigue signals, we first need to define what performance-based coaching is. The main goal of performance-based coaching is to actively improve an athlete's overall performance (Lyle & Cushion, 2017). In our case, this means helping runners keep up higher intensities and achieve a faster average pace. To understand how a digital intervention can facilitate this improvement successfully, we first have to examine the internal boundaries the athlete might face.

Generally, running is seen as a physical task, with its performance relying on physiological and muscular factors. However, The Central Governor Model (CGM) proposes that fatigue is partly a brain-derived emotion that works as a protecting mechanism (Noakes, 2012). The main function of this emotion is to protect homeostasis and prevent fatal biological failure (Noakes, 2012). The brain regulates this by modulating motor unit recruitment before actual biological damage occurs (Weir et al., 2006). At the core of this regulation is anticipatory pacing. The brain constantly evaluates internal physiological signals (e.g., core temperature or metabolic reserves) to compute a sustainable pace for the remainder of the run (Noakes, 2012; Weir et al., 2006). Critics, however, frequently challenge this theory and argue that the model is inherently unfalsifiable. This contributes to ongoing debate surrounding its validity.

To counter this deterministic view, the Psychobiological Model of endurance performance was developed (Marcora & Staiano, 2010). Based on the established Motivational Intensity Theory (Brehm & Self, 1989), this paradigm has gained significant popularity within sports science as the main alternative to the CGM. Contrary to the CGM, this model states that pacing and fatigue are regulated by a conscious decision-making process and not by an unconscious physiological safety switch. The psychobiological model suggests that an athlete's decision to stop exercising is based on a conscious assessment. This assessment occurs when an athlete's perceived effort outweighs the internal motivation to keep going (Inzlicht & Marcora, 2016).

Instead of viewing the Central Governor Model and the Psychobiological Model as mutually exclusive, recent literature suggests that this division works as a "conceptual quagmire" that limits our understanding of pacing (Micklewright et al., 2017). Alternatively, Micklewright et al. (2017) proposed applying the widely accepted Dual Process Theory from cognitive psychology that distinguishes automatic responses from deliberative thoughts (e.g., Kahneman, 2011) to the regulation of pacing. Within this proposed framework, fatigue and pacing are regulated by two distinct but interacting systems. First, an intuitive system (System 1) is thought to function automatically and can therefore be associated with the unconscious processing of physiological feedback such as the body's fatigue signals. Second, a deliberative system (System 2) enables conscious, goal-directed decision-making based on motivation (aligning with the Psychobiological Model).

A key limitation is that the process of making decisions during intense exercise is affected by both mental and physiological demands. This may lead to more reliance on the automatic processing of System 1 (Micklewright et al., 2017). Sakalidis et al. (2022) suggest that external stimuli, such as coaching feedback, can improve performance by reducing the focus on internal signs of fatigue. They function as social affordances that instantly trigger performance-sustaining behaviours. By engaging with these external cues, the athlete can potentially reduce reliance on System 2's deliberative work of fighting the automatic stop signals. This may reduce the internal cognitive load and help the athlete to maintain goal-directed performance without the exhaustive mental effort that is normally required to override the physiological safety signals of System 1. However, the specific cognitive processes that cause this effect are still being studied.

2.2 Psychological Factors Influencing Running Performance

The Dual Process framework explains the cognitive mechanics of pacing. However, it does not address the underlying motivational drive that dictates why athletes sustain or stop effort. For that reason, this section examines psychological factors, including motivation, self-efficacy, and reactance. These factors determine whether the goal-directed process of System 2 remains engaged during demanding endurance exercise or not.

A recent review by Alkasasbeh and Akroush (2025) emphasises that individual sports (e.g., running) rely primarily on intrinsic motivation to sustain effort. Morris et al. (2022) highlight that this intrinsic motivation is a proactive process. It is driven by autonomy, the need for achievement, and the inherent value of the activity itself. Therefore, setting personal goals is the main driver of intrinsic motivation.

However, alongside the critical role of intrinsic motivation for sustained effort, the dynamic nature of running requires quick psychological interventions. While intrinsic motivation is proactive, Morris et al. (2022) characterise extrinsic motivation as a reactive mechanism which is influenced by external stimuli and immediate cost-benefit analyses. Athletes often face difficulties when they rely solely on their internal willpower during intense physical activity or when their motivation temporarily decreases.

In these vulnerable instances, external reactive stimuli can bridge this motivational gap. Liu et al. (2022) revealed that immediate external rewards and stimuli can boost an individual's intrinsic motivation. As a result, this increases the task persistence. Within the context of running, these reactive triggers can potentially overcome acute reductions in drive experienced during fatigue. These triggers, by artificially supporting the deliberative processes of System 2, offer the psychological support that allows the athlete to maintain their intended pace (McCormick, Meijen, & Marcora, 2015).

An external trigger does not automatically translate into sustained internal drive. Self-efficacy is what mediates the interaction between internal drive and outside support. It can be conceptualised as a cognitive filter that influences an athlete's processing and response to external feedback. According to Bandura (1997), self-efficacy is a fundamental determinant of effort, persistence, and performance. To fully grasp how this mechanism operates under fatigue, a distinction can be made between General Self-Efficacy (GSE) and Specific Self-Efficacy (SSE). GSE is a stable motivational characteristic that reflects a generalised belief in one's capabilities, while SSE is a variable, task-specific confidence in one's ability to sustain the current effort (Chen, Gully, & Eden, 2001). Both factors influence how much effort, persistence, and resilience an athlete is willing to invest when confronted with demanding circumstances. When runners are physically exhausted, they often experience a temporary decrease in their SSE. This leads to negative or self-limiting thoughts about their capacity to complete the exercise while maintaining the same level of effort. To avoid this, the athlete must actively fight and restructure these negative beliefs. This is a psychological strategy known as cognitive reframing. Tóth, Resperger, and Tóth (2024) show that effectively reframing these thoughts reduces performance anxiety and builds mental resilience. This is necessary for athletes to perform at their best and to keep the mental drive that is needed to control the brain's automatic fatigue signals (System 1).

During prolonged runs, as physical exhaustion sets in, this internal regulation becomes less effective, and the athlete becomes more reliant on external input to sustain effort. Yet, as Bandura (1997) notes, the effectiveness of external persuasion depends strongly on its perceived credibility. Positive encouragement operates by targeting SSE directly by attempting to restore the athlete's task-specific confidence. When that confidence is severely depleted by fatigue, however, optimistic cues sharply conflict with the athlete's internal physical reality and are likely to be discounted as unrealistic or untrue.

Interestingly, the dynamic between external feedback and the athlete's internal state becomes specifically complex when we look at negative or doubtful external stimuli. Unlike positive encouragement, ego-threatening feedback does not attempt to restore

SSE and bypasses it. When an athlete is confronted with feedback that questions their capabilities, it challenges not their task-specific confidence but their identity as a capable athlete. This can trigger a strong automatic counterreaction.

According to the Psychological Reactance Theory (Brehm, 1966), individuals experience a motivational spike to recover their threatened autonomy and prove their competence when their freedom or capability is challenged. Because this defensive response manifests as an immediate emotional reflex, it can be partly conceptualised as operating within more automatic (System 1-like) processes. This creates a direct functional link to pace regulation: within the Dual Process framework, pacing breaks down when System 2 can no longer suppress the automatic "stop" signals of System 1. A reactance-driven "fight" impulse can operate at a similarly automatic level and potentially counteracts these fatigue signals directly. If so, this reduces the regulatory burden on System 2 rather than adding to it.

Yet the conversion of this defensive urge into enduring effort is not guaranteed. GSE may be the decisive factor here. The Theory of Challenge and Threat States in Athletes (TCTSA; Meijen et al., 2020) provides a useful lens here: it proposes that athletes appraise situational demands against their available coping resources. When resources meet or exceed demands, a performance-enhancing challenge state arises; when demands outweigh resources, a maladaptive threat state occurs. Whether a reactance response tips into sustained effort or disengagement therefore depends largely on whether the athlete has sufficient coping resources. Athletes with higher GSE possess more stable identity-level resources to draw on. Therefore, they are more likely to appraise a competence challenge as motivating instead of threatening. This distinction becomes critical under the depleted cognitive conditions of intense fatigue, as discussed in Section 2.3, and provides the theoretical basis for treating GSE as a moderator of ego-threatening feedback's effectiveness.

2.3 The Impact of Cognitive Load

Given that motivation and self-efficacy influence an athlete's willingness to maintain effort, it is important to recognise that these processes operate under limited cognitive resources during prolonged exertion. This means that it is necessary to understand the role of cognitive load when designing a digital coaching intervention.

As noted in Section 2.1, System 2 processing occurs under extreme physiological and cognitive demands during intense exertion. In contrast to the mental relaxation associated with recreational jogging at a leisurely speed (Dietrich, 2006), endurance running with specific performance-based goals presents a distinct experience. Maintaining a set pace requires continuous conscious deliberation and active assessment of physical signals (Brick, MacIntyre, & Campbell, 2014). This challenging self-regulation relies heavily on "inhibitory control", which is a key executive function responsible for overriding strong internal impulses and automatic reflexes (Diamond, 2013). The importance of this specific executive function in endurance sports was shown by Cona et al. (2015), who found that faster ultra-marathon runners had significantly better inhibitory control than slower runners. Yet, inhibitory control comes at the cost of demanding extensive mental focus and working memory capacity (Diamond, 2013).

By consistently using this cognitive ability to override the body's protective reflexes, the athlete depletes their mental energy. During prolonged intensive running, the deliberative System 2 often lacks the mental energy required to successfully execute conscious cognitive reframing. When athletes are exhausted and have a reduced cognitive capacity for this reappraisal, their response to external cues becomes more reliant on automatic evaluations.

As outlined in Section 2.2, the TCTSA framework predicts that adding cognitive demands to an already depleted athlete tips the balance from a challenge state into a threat state (Meijen et al., 2020). Imposing deliberative feedback under these conditions does precisely that. It increases the demand side of the appraisal at the moment coping resources are least available. If the added mental effort of processing this feedback overwhelms the athlete's already depleted resources, the risk of entering a threat state significantly increases. Unlike SSE that fluctuates with momentary performance and fatigue, GSE is a stable belief in one's general capabilities that is relatively resistant to situational depletion (Chen, Gully, & Eden, 2001). It is therefore not subject to the same exhaustion as working memory or inhibitory control. It functions as a fixed resource rather than a consumable one. When GSE is high, it provides a stable psychological buffer. The athlete then possibly automatically appraises ego-threatening feedback as a challenge and responds with compensatory effort (Wright, Agtarap, & Mlynski, 2015). When GSE is low, ego-threatening feedback is unlikely to trigger reactance at all. Rather than reading the cue as a challenge to overcome, the depleted athlete is more likely to interpret it as confirmation of failure and activate shame rather than the fight response (Conroy, Willow, & Metzler, 2002). Under these conditions, the feedback accelerates disengagement rather than preventing it (Steindl et al., 2015).

Marcora, Staiano, and Manning (2009) demonstrated how cognitive demands affect endurance performance. In their experiment, subjects who were given a mentally challenging task before an intense cycling task reached physical exhaustion significantly faster (640 ± 316 s) than subjects in the control group (754 ± 339 s). Additionally, the mentally exhausted athletes experienced a significantly higher Rating of Perceived Exertion (RPE). Due to the physical task being experienced as cognitively harder from the start, they reached their maximal perceived effort and disengaged prematurely. For an external coaching intervention, this underscores a fundamental requirement for its design. The system's triggers must offload the processes of System 2, instead of adding additional cognitive load through complex or difficult-to-process feedback.

2.4 The Digital Coach's Medium

Because minimising these cognitive burdens is vital for maintaining running performance, the medium through which a digital coach provides feedback must be decided on carefully. Running is an activity with a high demand for visual attention. As noted by Eriksson, Halvorsen, and Gullstrand (2011), visual feedback occupies the runner's vision (e.g., by forcing them to look at a screen). This directly conflicts with the fact that a runner uses visual cues to monitor their surroundings. It is established in cognitive psychology that splitting attention between two tasks that use the same sensory channel significantly increases cognitive load (Wickens, 2008). Additionally, depending on visual displays can impact running mechanics (Van Hooren et al., 2020). These distractions disrupt a steady cadence and therefore lead to micro-accelerations and continuous pace corrections. Bood et al. (2013) discovered that an inconsistent cadence leads to energy depletion and a reduction in running economy. This directly impacts the athlete's running performance.

To avoid these risks in performance loss, an eyes-free interface is required. While haptic feedback provides an alternative eyes-free intervention, it is heavily underexplored in real-time performance guidance while running (Van Hooren et al., 2020). In general, haptic feedback provides relatively simple signals that cannot convey complex instructions without the assistance of another perceptual channel. Additionally, subtle vibrations are difficult to interpret and are easily masked by the mechanical impacts of running itself.

Therefore, audio feedback appears to be the most effective medium in this context. Unlike haptic signals, it facilitates an eyes-free interface that can deliver clear and more complex cues. The theoretical advantage of audio lies in minimising the cognitive load. According to the Multiple Resource Theory, the brain can process auditory and visual information in parallel without these channels interfering (Wickens, 2008). The experiments by Eriksson, Halvorsen, and Gullstrand (2011) confirm this auditory advantage. The researchers found that runners who used auditory cues could adjust their running style more effectively than those using visual cues, while still paying full attention to their surroundings. Furthermore, they noted that visual feedback can easily be ignored. In contrast, auditory feedback imposes itself on the athlete. This ensures that the intervention is received even during fatigue.

2.5 Performance-Based Digital Coaching

Now that we have established that these external auditory signals can decrease the cognitive burden, it is important to find what type of feedback a digital system must provide to regulate and improve performance. While the auditory channel provides the optimal path for offloading System 2 during fatigue, today's sport technology utilises this potential rarely. Stragier et al. (2016) characterise current running-app market leaders as 'Online Fitness Communities', designed around enjoyment, social validation, and peer reinforcement. For athletes who often strive to improve their running performance, the motivational orientation of many applications is not enough. While these design choices increase engagement, they position feedback primarily as social affirmation rather than real-time performance regulation.

Even platforms that do shift their focus towards performance metrics struggle to provide the correct real-time feedback needed for an improvement in performance. Modern wearables and applications collect enormous amounts of data during the run. However, there is a fundamental difference between the logging of data and the actual coaching of an application. Schmidt et al. (2015) state that a strict distinction must be made between 'Fitness Trackers' and 'Digital Personal Coaches'. Part of the role of a digital personal coach is to motivate the user through the process and help them achieve their goals. Current performance-focused systems function primarily as trackers. They register physiological and biomedical variables (such as heart rate, cadence, and pace) and visualise this afterwards, while often lacking the capacity to translate this data to real-time actionable interventions.

For a real-time digital coach, the intervention must match the principles of motor learning closely. As Giraldo-Pedroza et al. (2020) highlighted, every form of real-time biofeedback is actively competing with the primary motor task for working memory resources. When a system is too complex or gives continuous feedback, it limits the available attention of the athlete, which can worsen running mechanics. Another risk factor of continuous feedback provision is that it could lead to feedback dependency (Van Hooren et al., 2020). In this scenario, the athlete does not learn to interpret physiological signals without relying on the intervention. While current applications utilise continuous visual feedback, Van Hooren et al. (2020) argue that effective real-time coaching must utilise bandwidth feedback. The system is then silent while the athlete performs optimally and intervenes only when surpassing specific biomedical or physiological thresholds. The clinical effectiveness of this bandwidth approach was recently tested in a large in-field randomised controlled trial (Van Hooren, Plasqui, & Meijer, 2024). Participants received real-time auditory instructions that targeted spatiotemporal metrics and relative speed (e.g., "try to run a little faster") when deviating from individually set target zones. While the study found some evidence for reduced injury rates among participants who consistently used the feedback, neither group showed a significant improvement in running performance (Van Hooren, Plasqui, & Meijer, 2024).

Two potential explanations for this null performance effect are directly relevant to the present thesis. First, the cues were informational and biomechanical in nature, instructing the athlete what to adjust (Van Hooren, Plasqui, & Meijer, 2024). It is possible that such instructions require the athlete to consciously integrate the cue with ongoing pacing decisions, thereby adding to rather than reducing the deliberative System 2 burden. Whether cues that are designed to trigger a more automatic affective response would produce a different outcome remains untested. Second, performance was operationalised as change in personal best time over a six-month training period, a long-term outcome shaped by training volume and recovery. This differs fundamentally from in-session pace regulation during a single high-intensity effort, which is the outcome the present study examines. The limited available evidence suggests that purely informational pace feedback does not significantly alter in-session running pace (Puleo & Abraham, 2018), though this finding is based on a small treadmill study using visual display feedback rather than auditory motivational cues. Whether the latter can influence in-session pace regulation under field conditions therefore remains insufficiently tested, and constitutes one of the contributions this pilot study aims to make.

Even when the content and timing of the feedback are biomedically accurate, the manner in which these are translated (the tone of voice) is crucial for the eventual effectiveness. A digital coach must build a working alliance with its user in order to realise behavioural adherence. Contrary to the assumptions that technology lacks empathy, Barger (2025) found that AI-coaches are highly capable of forming this bond. In fact, users can build a working alliance with an AI-coach similar to a working alliance with a human professional, where often the digital interaction is experienced as an unexpectedly safe and non-judgemental environment. Barger (2025) links this to a reduction of impression management behaviours. Due to the user interacting with a machine, there is less motivation to maintain a certain social image, which results in higher psychological safety. For an athlete pushing their physical boundaries, this could be an advantage. Potentially, this opens the door to a more direct, confrontational, or even ego-threatening effective coaching approach. Due to the feedback originating from a machine rather than a human peer, the runner is likely to experience less social shame, frustration, or fear of being judged. This allows the athlete to accept strict, corrective interventions that might otherwise cause stress or damage the relationship with a human coach. From a psychological perspective, corrective or negative feedback from a human coach can cause a social-evaluation threat (Dickerson & Kemeny, 2004). Dickerson, Mycek, and Zaldivar (2008) demonstrated that it is specifically this evaluative judgment component and not mere social presence that drives the cortisol and shame response. The stress reaction, in other words, requires a perceived conscious evaluator.

When the feedback originates from a non-human digital system, this evaluative pressure may be reduced. Barger (2025) found that users exhibit reduced impression management behaviours when interacting with AI-coaches, experiencing the interaction as unexpectedly safe and non-judgemental. Yet according to the Computers Are Social Actors (CASA) paradigm, individuals automatically respond to voice-based and interactive technologies as if they were human, even when they know they are interacting with a machine (Nass & Moon, 2000). These are unconscious, automatic reactions rather than conscious beliefs that the computer is human. A digital coach may therefore occupy a unique psychological position: it may trigger automatic social responses, while reducing the shame and ego depletion that would accompany such feedback from a human coach.

However, with the evaluation of speech-based digital coaches, Kannampallil et al. (2022) found a new risk. Users frequently reject the interaction or perceive it as non-human due to insufficient variation in tone and emotion. A monotone voice limits the emotional engagement of the interaction.

2.6 Autonomy-Supportive vs. Ego-Involving Feedback

To create a digital coaching intervention that can be programmed to use either a strict or an understanding tone, a clear psychological framework is essential. In sports science, these distinct conversational styles are typically examined through the lens of Self-Determination Theory (SDT) (Deci & Ryan, 1985). Within this theory, we can separate coaching styles into autonomy-supportive and controlling coaching (Mageau & Vallerand, 2003). From the perspective of Achievement Goal Theory (AGT), motivational environments are conceptualised as either task-involving or ego-involving (Nicholls, 1989; Ames, 1992). Although these frameworks come from different theoretical backgrounds, the controlling behaviours in SDT and ego-involving climates in AGT often overlap. To determine which style is the best fit for real-time audio cues under conditions of cognitive fatigue, we must look at the strengths and weaknesses of each approach in light of the constraints established in Sections 2.2 and 2.3.

Autonomy-supportive feedback provides athletes with choices, acknowledges their emotional states, and avoids controlling language. This feedback approach has been shown to enhance intrinsic motivation and psychological well-being as demonstrated by Mageau and Vallerand (2003). Furthermore, a comprehensive meta-analysis ($N = 38,844$) confirmed that coach autonomy support is a strong predictor of motivation and sustained athletic engagement across various populations (Mossman et al., 2022). While this evidence is strong for long-term athletic growth and commitment, autonomous self-regulation requires the athlete to actively process, accept, and integrate external input with their personal goals and values (Deci & Ryan, 2000). Integrating external input relies heavily on System 2 processing. As established in Section 2.3, deliberative processing is severely constrained under acute fatigue and this is precisely the condition under which autonomy-supportive feedback demands the most cognitive resources. This could make it less effective as a real-time performance intervention during acute fatigue.

In contrast, ego-threatening feedback links performance directly to perceived competence and uses evaluative or confrontational cues. For the purposes of this thesis, ego-threatening feedback is distinguished from simply negative or insulting feedback by a single criterion: it must be competence-challenging. It confronts the athlete's current performance relative to their goal and implies they are capable of better, without attacking personal character. This distinction matters because competence challenges are theorised to activate reactance, whereas unconditional negative judgements risk activating shame which is a fundamentally different and counterproductive psychological response (Conroy, Willow, & Metzler, 2002).

Instead of making the athlete think about the meaning of the feedback, these cues might trigger an immediate, affective response. Steel, Bishop, and Taylor (2021) demonstrated that participants in a controlling motivational condition performed better on a physical persistence task than autonomously motivated participants. This suggests that controlled motivation can drive greater short-term physical effort than autonomy-supportive alternatives. As discussed in Section 2.2, such provocative feedback may evoke Psychological Reactance (Brehm, 1966) that possibly results in more automatic processes that partially bypass effortful System 2 regulation. However, this approach carries considerable risks. Steel, Bishop, and Taylor (2021) found a notable correlation between controlled motivation and a marked increase in cortisol secretion.

That said, ego-threatening feedback does not automatically produce a cognitive advantage. A critical theoretical distinction must be made between automatic attentional capture and deliberative self-protection here. According to Attentional Control Theory (ACT; Eysenck et al., 2007), a perceived threat automatically pulls attentional resources toward the threatening stimulus via the stimulus-driven system (System 1). In the context of endurance running, this thesis proposes that this mechanism can be exploited as a cognitive distraction. When an athlete's attention is overwhelmingly dominated by internal pain signals, an ego-threatening cue uses this same stimulus-driven reflex to forcibly redirect focus away from fatigue.

Yet ACT also warns that if the threat triggers worry or self-focused rumination, processing efficiency deteriorates and cognitive load increases. Furthermore, Lazarus and Folkman (1984) add that if a perceived threat is processed beyond this initial reflex, it inevitably triggers a secondary evaluation of whether one has the psychological resources to cope with it. This secondary appraisal is itself a highly deliberative (System 2) process. If the athlete shifts from the initial automatic reactance into this conscious rumination or worries about failure (possibly due to low GSE), cognitive load increases rather than reducing it. Therefore, safely integrating ego-threatening feedback into the digital coaching model requires specific psychological safeguards, adding to the structural constraints (such as eyes-free or bandwidth delivery) established in earlier sections. First, to prevent the shift from automatic reactance to deliberative secondary appraisal, the intervention must structurally minimise the temporal window for ruminative processing. Second, it must fundamentally mitigate the social-evaluative threat (Dickerson & Kemeny, 2004) that most reliably triggers this self-protective deliberation.

In sum, autonomy-supportive feedback excels at building long-term motivation but places the greatest deliberative demands on athletes precisely when cognitive resources are most depleted. Ego-threatening feedback offers a potential shortcut through automatic reactance, but its effectiveness depends critically on the athlete's available coping resources and the design of the cue itself. How these contrasting dynamics play out as a real-time digital intervention under conditions of acute fatigue remains empirically unresolved and forms the central problem this thesis addresses. Figure 1 provides a schematic overview of the proposed mechanisms through which ego-threatening and autonomy-supportive feedback influence pace regulation under fatigue.

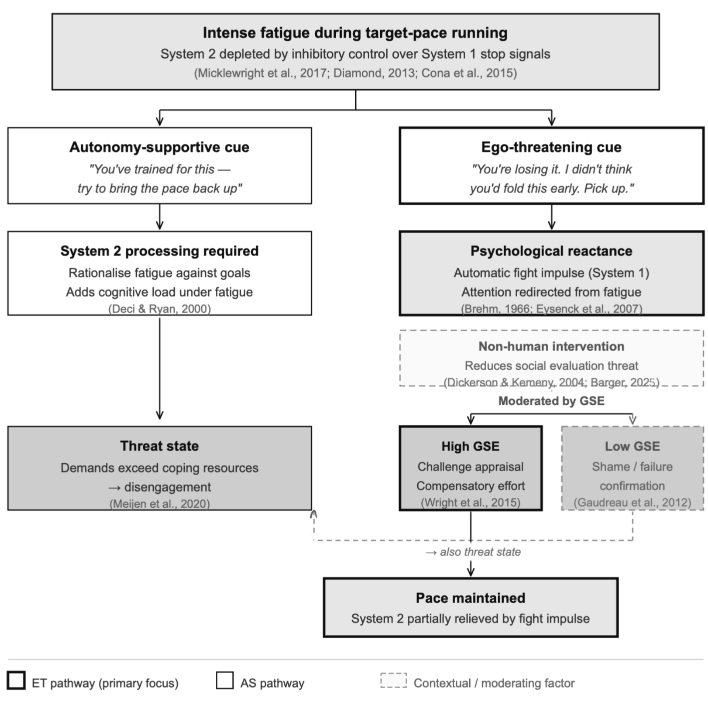

Figure 1
Theoretical framework visualising the proposed mechanisms through which ego-threatening and autonomy-supportive feedback influence pace regulation under fatigue.

3. Problem Statement and Research Question

The literature reviewed in Chapter 2 highlights a conflict in how to design digital coaching systems for endurance running. To keep up a challenging pace, the athlete requires the use of cognitive resources for deliberate effort regulation (System 2). At the same time, autonomy-supportive feedback fosters long-term motivation, but its reliance on conscious processing may overwhelm the athlete during these moments of high cognitive load.

Alternatively, ego-threatening feedback could trigger an immediate, affective response (Psychological Reactance) that potentially reduces the reliance on effortful System 2 regulation. Nevertheless, this approach presents a critical trade-off. It may enhance the athlete's short-term performance by increasing motivational intensity, but it could also elevate the athlete's psychological stress, increase cognitive load, or trigger disengagement if the athlete lacks the necessary coping resources. It remains unclear how these contrasting coaching styles function as real-time interventions in performance-based runs.

3.1 Research Gap and Contribution

Despite extensive research on motivation and coaching styles in sport, three key gaps are present in the existing literature. First, existing studies predominantly link coaching styles to long-term outcomes (e.g., well-being or adherence), resulting in a limited empirical understanding of their moment-to-moment effectiveness as performance-boosting interventions. Second, while external feedback generally enhances performance, it is unclear whether provocative feedback styles successfully reduce the reliance on effortful cognitive control (System 2) during high-demand situations (McCormick, Meijen, & Marcora, 2015). Third, existing studies on provocative or ego-involving feedback have predominantly used human coaches or confederates as the source of feedback, meaning the social-evaluative dimension of the threat cannot be disentangled from the motivational effect (Dickerson & Kemeny, 2004). Conversely, research on digital coaching has predominantly focused on supportive or neutral feedback styles. No study has examined whether the theorised cognitive advantage of ego-threatening feedback is amplified, maintained, or reduced when delivered by a non-human digital system.

This thesis addresses these gaps by piloting a within-subjects field comparison of autonomy-supportive and ego-threatening auditory feedback. By focusing on the challenge of achieving a faster average pace under increasing physical fatigue, this research examines how the different feedback styles influence pacing behaviour under constrained cognitive resources.

To assess the impact of both feedback styles, this study measures three outcome variables. Objective running performance tests whether ego-threatening feedback produces a measurable performance benefit. Perceived cognitive processing operationalises the System 1/System 2 distinction that existing research has left unexplored. Perceived psychological stress captures the potential cost of ego-threatening feedback and directly addresses the performance-risk trade-off central to this thesis. Rating of Perceived Exertion (RPE) is included as a standard secondary measure to contextualise perceived effort across conditions (Borg, 1982).

3.2 Research Question

Based on the identified problem and research gap, this thesis is guided by the following research question:

> To what extent does ego-threatening auditory feedback, delivered by a digital coach, lead to a faster average running pace compared to autonomy-supportive feedback during high-intensity running, and what are its effects on perceived cognitive processing effort and perceived psychological stress?

3.3 Hypotheses

Based on the theoretical framework outlined in Chapter 2, the following hypotheses are proposed:

Given the exploratory nature of this pilot study, H1–H4 are treated as directional predictions instead of formal confirmatory hypotheses.

H1 Athletes receiving ego-threatening feedback are expected to achieve a faster average pace than those receiving autonomy-supportive feedback during a high-intensity 5-kilometre run.

H2 Ego-threatening feedback will be perceived as requiring less deliberate cognitive processing (System 2) compared to autonomy-supportive feedback.

H3 Ego-threatening feedback is expected to be associated with higher perceived psychological stress than autonomy-supportive feedback (Steel, Bishop, & Taylor, 2021).

H4 General Self-Efficacy (GSE) will positively moderate the effect of ego-threatening feedback on average running pace, with athletes higher in GSE showing greater pace improvement in response to ego-threatening cues (Bandura, 1997; Wright, Agtarap, & Mlynski, 2015).

4. Methodology

This chapter describes the methodology designed to answer the research question and test the hypotheses formulated in Chapter 3. It covers the core measurement approach (4.1), the experimental design (4.2), the intervention and system architecture (4.3), the participant sample (4.4), the operationalisation of variables (4.5), the procedure (4.6), and the data analysis plan (4.7).

4.1 Core Measurement Approach

The primary dependent variable in this study is *objective running performance*. In this thesis we define this metric as the average pace achieved across a 5-kilometre run in seconds per kilometre.

Pace is selected over alternative performance indicators (e.g., heart rate, cadence, or finish time) for three reasons. First, it is the most direct expression of running output and therefore the most direct behavioural outcome of the effort-regulation process central to this thesis. Second, it is the metric recreational runners themselves use to evaluate performance. Third, it is measurable at high resolution and objectively via GPS without imposing additional burden on the athlete.

The study does not track heart rate or biomechanical variables. Adding sensors or unfamiliar devices would add confounds and reduce ecological validity. This is a deliberate trade-off. Physiological data would enrich the analysis, but the added complexity is not justified given the scope of thesis-level research.

4.2 Experimental Design

The study employs a randomised, counterbalanced within-subjects design. Each participant completes two 5-kilometre runs under two contrasting conditions:

- **Condition A: Autonomy-Supportive Feedback (AS) —** cues that acknowledge effort, offer choices, and use non-controlling language.
- **Condition B: Ego-Threatening Feedback (ET) —** cues that directly challenge the athlete's competence relative to their goal, using confrontational and directive language.

A within-subjects design is used because individual differences in fitness, running economy, and psychological resilience are the primary threat to internal validity. Having each participant serve as their own control holds these factors constant and allows the analysis to isolate the effect of feedback style with a smaller sample than a between-subjects design would require.

Two potential confounds are addressed. Order effects are controlled through counterbalancing. Half of the participants complete Condition A first, the other half start with Condition B, with assignment randomised. Physiological carryover is controlled by a mandatory washout period of at least three days between runs. This is the time required to ensure physical recovery and prevent the effort of the previous run from affecting the second test (Bishop, Jones, & Woods, 2008). Even though psychological carryover cannot be completely eliminated, its impact as a limitation is mitigated by counterbalancing, which ensures the results remain free from systematic bias.

Importantly, the informational content of the feedback is identical across conditions. This applies across all cue types. Corrective cues communicate that pace has dropped below the individually set challenging minimum; 500-metre check-in cues confirm that pace is on track at each 500-metre marker; and recovery cues confirm that pace has returned to the minimum. In every case, only the tone and framing differ across conditions. The autonomy-supportive condition frames this as an invitation ("if you feel ready"), while the ego-threatening condition frames it as a competence challenge ("I did not expect you to fold this early"). Furthermore, all audio cues across both conditions are generated using the same TTS voice and baseline intonation settings, so that differences in vocal characteristics (e.g., pitch, speed, or timbre) do not confound the manipulation. This ensures that any observed differences can be attributed to the linguistic framing of the cues, not to informational content or vocal delivery.

The general awareness of being observed (participants know their data is reviewed) is held constant across conditions and therefore cannot account for between-condition differences. To further reduce equipment-related confounds the same smartphone (iPhone 15 Pro Max) and running armband were used for each run across both conditions. This ensures that potential differences in GPS signal quality or device performance cannot account for between-condition differences in pace measurement.

4.3 Intervention Design and System Architecture

4.3.1 Design Principles. To ensure the intervention aligns with the constraints established in Chapter 2, the system is designed around four fundamental principles. It utilises an audio-first interface featuring believably human-like speech (Sections 2.4 and 2.5). The entire coaching interaction is delivered via headphones and is designed as an eyes-free experience. The visual display only needs to be used for initiating a session and for opening the post-run questionnaire.

It relies on bandwidth feedback (Section 2.5), extended with proactive periodic cues (see Section 4.3.2). Pure bandwidth feedback intervenes only upon threshold deviation (Van Hooren et al., 2020). The addition of periodic cues departs from this, but ensures that all participants receive the audio cues, even when performance thresholds are not met during the run. Furthermore, periodic cues are necessary as they provide pace confirmation that substitutes for visual monitoring. The cues are brief and direct to minimise deliberative processing demands (Sections 2.3 and 2.6). Finally, the feedback is delivered by a non-human digital system, which, as argued in Section 2.5, may preserve automatic social responses while reducing the social-evaluation threat that comes with confrontational human feedback.

4.3.2 Technical Implementation. A mobile application was developed (co-designed with the AI language model Claude by Anthropic) for iOS and Android. All design decisions, including trigger thresholds, cue structure, and smoothing parameters, were made by the researcher based on pre-study testing. Claude was used to implement these specifications as functional code. The participant's individually determined challenging minimum pace is entered before each run. The app calculates current pace locally by computing the distance between successive GPS coordinates using the haversine formula, converted to seconds per kilometre. This position-based approach is smoothed with an exponential moving average (EMA) of $\alpha = 0.2$. This value was settled on through trial-and-error during pre-study testing. A lower value reduces GPS noise but delays detection; a higher value responds faster but may trigger on noise rather than genuine pace drops. Utilising an EMA is a standard signal-processing technique for real-time noise reduction while maintaining responsiveness to genuine pace changes. Local processing avoids dependency on internet connectivity and minimises latency. GPS outliers exceeding twice the current smoothed pace estimate are discarded before applying the EMA. This prevents false measurement spikes from triggering false threshold crossings.

The coaching system operates through four distinct trigger types. Each trigger delivers condition-appropriate audio cues.

1. Corrective threshold. If the athlete's pace decreases more than 5% below their predetermined minimal pace for a sustained duration of 15 seconds, an audio cue is played. To allow the GPS signal and EMA to stabilise at the start of a run, corrective threshold detection is suspended for the first 150 metres. Additionally, if the athlete's pace returns within the threshold band before the 15-second window elapses, the timer resets without a cue being delivered. This threshold accounts for pace fluctuations that occur naturally during outdoor running (e.g., due to slowdowns around turns, slight inclines, or wind). Not having this margin would overload the athlete with feedback that does not necessarily reflect a decline in effort. The 15-second duration serves a dual purpose. First, the system only triggers a cue when sustained pace drops occur and filters out momentary dips. Second, it mitigates the risk of GPS measurement noise triggering a false cue, as random signal fluctuations are unlikely to be sustained over 15 consecutive seconds. The residual risk that some cues reflect measurement error rather than true effort lapses is acknowledged as a limitation of GPS-based field assessment. Both values were determined through outdoor testing during development. No established standard for these thresholds exists in the literature; they are informed by the bandwidth feedback principles described by Van Hooren et al. (2020). Each condition has eight distinct corrective cues (Table 1), played sequentially and cyclically through headphones.

2. 500-metre check-in cues. These cues are delivered at nine distance points throughout the run: 0.5 km, 1 km, 1.5 km, 2 km, 2.5, 3 km, 3.5 km, 4 km, and 4.5 km. These cues serve two functions. First, they confirm pace status in a fully eyes-free manner, eliminating any need for the participant to glance at the screen. Second, they ensure that all participants receive condition-appropriate coaching throughout the run, including participants who never drop below threshold and would otherwise receive no coaching at all. Note that while check-in cues ensure exposure to both feedback styles, they do not substitute for corrective cues as the primary performance-dependent manipulation. The informational content is identical across conditions (pace confirmation), only the framing differs (Appendix A, Table 8).

3. Recovery confirmation cues. These cues are delivered after three consecutive GPS measurements (3–6 seconds) confirm the athlete's pace has returned to the minimum pace zone. This multi-sample criterion guards against single GPS noise readings that may spuriously appear as recovery. This cue is necessary because, without it, the athlete receives no audio signal confirming successful recovery. This is essential in order to provide a fully eyes-free digital performance coaching intervention. Without this confirmation, the intervention leaves the athlete uncertain about when to stop pushing. In Condition A, recovery is acknowledged positively; in Condition B, it is acknowledged with a brief directive to maintain pace. A single cue per condition is used (Appendix A, Table 10). All audio cues are generated with an energetic ElevenLabs Text-to-Speech voice and stored locally in the app. This engine is chosen for its human-like intonation. Monotone or robotic delivery would risk disengagement and diminished emotional impact (Kannampallil et al., 2022). Local storage ensures instantaneous playback without network dependency.

4. Start and finish cues. The system provides start and finish cues appropriate to the active condition (Appendix A, Table 7). The start cue is played after a 5-second countdown from the moment the athlete presses the start button on the application's visual display. After these 5 seconds, the run begins and the application starts collecting running data. This countdown allows the participant to put their phone away (e.g., in a running belt, armband, or pocket). When the system registers that 5 kilometres have been covered, it automatically triggers the final cue and ends the run. These start and finish cues are crucial for this eyes-free auditory running intervention.

4.3.3 Data Logging. The application logs GPS-based pace at 1-second intervals for time-in-zone analysis. This resolution is sufficient for that metric while keeping the dataset manageable. The average pace across the 5-km run is calculated as the elapsed time between the Run Started and Run Stopped log entries divided by the GPS-measured distance. This timestamp-based calculation eliminates the influence of GPS speed measurement noise on the primary outcome variable. Threshold crossings, cue deliveries (corrective and check-in cues), and recovery events are logged at the exact moment they occur, as the feedback algorithm reacts to every GPS update (approximately every 1 second at running pace). A confirmation cue for returning above the minimum pace, requires three consecutive GPS measurements at or below the minimum pace. All data is stored locally during the run and exported afterwards for analysis.

4.4 Participants

Participants (N = 16) are recruited via local running clubs and online running communities. The target population is recreational athletes with performance goals, i.e., runners who set a challenging target pace and are invested enough in their performance to engage meaningfully with feedback. Elite athletes with professional coaching structures and purely recreational joggers without pace ambitions are excluded.

Inclusion criteria require participants to (1) run at least once per week and (2) be free of current physical injuries. All participants are adults (18+). They are aged between 22 and 60, of which 37.5% are female and 62.5% male. No age restrictions are imposed beyond the requirement that participants are adults (18+). While age may influence how athletes respond to confrontational feedback, the within-subjects design controls for this variable.

The study was approved by the CrIT Ethics Committee of LIACS, Leiden University. All participants provided written informed consent before taking part.

Table 1
Corrective threshold cues for both experimental conditions. Triggered when pace drops more than 5% below minimum for 15 sustained seconds. Played sequentially and cyclically (1→8→1).

Autonomy-Supportive	Ego-Threatening
"Your pace is dropping a bit! If your body feels okay, see if you can slowly bring it back up to your minimum."	"You are losing your pace! I did not expect you to fold this early. Pick it up now."
"I notice your pace is slowing down! Whenever you are ready, try to find that rhythm again."	"Your pace is dropping too much! Stop making excuses and get back to your minimum."
"It looks like you are losing some speed! No pressure, but see if you can gently pick up the pace."	"You are falling behind! Is this really the best you can do right now? Speed up."
"You are slipping below your minimum! Check in with yourself and push a little harder if you have the energy."	"Unacceptable pace drop! Do not give up now, force yourself back to your minimum."
"You are slightly off your minimum pace! Remember to listen to your body, but try to maintain your goal if you can."	"You are losing speed! I thought you set a challenging goal. Prove you can handle it."
"Pace is drifting! Let's see if we can comfortably return to your minimum pace."	"Pace is dropping! Stop being weak. Push through and get back to your minimum."
"I am seeing a small drop in speed! You are doing great, just try to ease back into your minimum pace."	"You are slacking off! Stop feeling sorry for yourself and accelerate."
"You are a bit below your minimum pace! Feel free to adjust, but try to push back up when you feel ready."	"Pace is below minimum! Are you seriously quitting on yourself? Move faster."

4.5 Variables and Operationalisation

The variables are selected to capture the performance-risk trade-off central to this thesis. Namely, whether ego-threatening feedback improves performance, through what cognitive mechanism, and at what psychological cost.

4.5.1 Primary Outcome Variables. **Objective Running Performance (H1)** is operationalised using the app's GPS data. The primary metric is average pace across the 5-kilometre run (in seconds per kilometre). Because pace is measured in seconds per kilometre, a lower value means faster running. This metric is interpretable for all participants regardless of whether they cross the minimum pace threshold. This makes it robust to individual differences in how challenging the minimum pace proves to be.

Three secondary metrics complement the primary outcome: (a) the percentage of total run duration spent at or above the minimum pace, indicating how consistently participants maintained their individually set minimum; (b) the number of corrective cues triggered per run, reflecting how often performance dropped to a threshold-crossing level; and (c) the cue-response pace profile, capturing the immediate behavioural response to each corrective cue. For each cue, pace deviation relative to the participant's target pace is extracted across a window of 200 metres before and 600 metres after the cue moment, and averaged across all cues per condition. If ego-threatening feedback activates a more automatic fight response as theorised, this should manifest as a faster and stronger pace recovery in the metres immediately following the cue compared to autonomy-supportive feedback. This measure provides a behavioural operationalisation of the reactance mechanism proposed to underlie H1 and H2. This measure serves as the objective behavioural complement to the self-reported Motivational Reactance measure (Section 4.5.2). Together they capture the reactance mechanism from two angles: whether it was subjectively experienced, and whether it translated into an immediate change in running behaviour.

Perceived Cognitive Processing (H2) captures whether the feedback was processed deliberately (System 2) or more automatically (System 1). It is measured post-run using three items on a 5-point Likert scale: "I had to consciously think about the meaning of the feedback," "My reaction to the feedback felt automatic" (reverse-coded), and "The feedback interrupted my focus while running." This variable addresses the second component of the research question: whether ego-threatening feedback reduces the perceived cognitive effort required to process the cue. Directly measuring System 1/System 2 activation during a field run would require invasive neuroimaging (e.g., EEG), which is infeasible during outdoor running. Self-report is the most ecologically valid approximation available, though retrospective recall bias is a limitation.

Perceived Psychological Stress (H3) is measured using a validated single-item measure (Elo, Leppänen, & Jahkola, 2003): "How stressful did you find the feedback you received during the run?" on a 5-point scale (1 = 'Not at all', 5 = 'Very much'). This addresses the third component of the research question: what effect ego-threatening feedback has on perceived psychological stress. A single item is used because the post-run questionnaire must be completed within minutes to minimise memory decay, and questionnaire fatigue would reduce response quality.

4.5.2 Explanatory Variables. **Motivational Reactance** is measured post-run via two items on a 5-point Likert scale: "The feedback made me want to prove it wrong" and "The feedback motivated me to push harder immediately." These capture the competitive "fight" impulse predicted by reactance theory (Brehm, 1966). If ego-threatening feedback does improve average running pace (H1), elevated reactance scores would support the theoretical explanation that this improvement operates through the reactance mechanism proposed in Section 2.2. The cue-response pace profile (secondary outcome c, Section 4.5.1) serves as the objective behavioural counterpart to this self-report measure.

General Self-Efficacy (GSE) is assessed at baseline using the 10-item General Self-Efficacy Scale (Schwarzer & Jerusalem, 1995). GSE is theorised to moderate the effect of ego-threatening feedback (H4). Given sample size constraints, H4 findings are interpreted as exploratory.

4.5.3 Validation and Control Variables. **Manipulation Check:** After each run, participants rate "I experienced the feedback as supportive" and "I experienced the feedback as confrontational" (5-point Likert). Without this check, a null result would be ambiguous. It could mean that the feedback styles have no differential effect as well as suggesting that participants did not perceive the manipulation as intended.

Rating of Perceived Exertion (RPE): Measured post-run using the Borg CR10 scale (Borg, 1982). This measure serves a dual purpose. First, it confirms that the individualised target pace induces a state of high physical exertion across all participants. Second, a difference in RPE between conditions would itself be informative. If ego-threatening feedback triggers a reactance-driven fight impulse that counteracts internal fatigue signals (Section 2.2), athletes may report lower perceived exertion despite comparable physical output. Such a pattern would provide supplementary evidence for the proposed cognitive mechanism.

Two open-ended questions were included in the post-run questionnaire to provide contextual insight into participants' subjective experience of the coaching intervention: (1) "Describe what went through your mind when a corrective cue played, and what you did next"; (2) "Was there anything about the coaching that bothered or helped you? What exactly?" The first question probes the cognitive and behavioural response to the cue, directly relevant to the proposed reactance mechanism. The second captures usability and perception differences between conditions.

4.5.4 Post-hoc Exclusion Criteria. Post-hoc exclusion criteria are applied to ensure the validity of the between-condition comparison. Participants are excluded from the primary analysis if: (1) their chosen target pace was not sufficiently challenging, operationalised as receiving zero corrective cues across both conditions. Although 500-metre check-in cues were delivered regardless of pace, corrective cues serve as the main performance-based manipulation. A participant who consistently runs faster than their minimum pace is unlikely to experience their ego as threatened (Nicholls, 1989), as the feedback confronts a performance gap that does not exist for them. This is the context the ego-threatening condition is designed to address. The ecological validity of the between-condition comparison is therefore reduced. Additionally, without corrective cues the cue-response pace profile (secondary outcome c) cannot be computed. This makes inclusion of such participants analytically incomplete; (2) they demonstrated an inability to sustain the target pace across one or both conditions, evidenced by either extended walking intervals or a structural pace deficit throughout the run without meaningful recovery. This indicates that the chosen target pace was too fast for the participant on that day, making the pacing data incomparable across conditions; or (3) they failed to complete both experimental conditions or provided incomplete data for a run. These criteria are assessed from the logged GPS data prior to any statistical analysis.

4.6 Procedure

The experiment spans two testing days per participant.

Pre-experiment. Participants are briefed on the general purpose (examining digital coaching during running) without being informed about the two feedback styles. This partial blinding prevents expectation effects. They sign informed consent and complete the baseline questionnaire (demographics and GSE scale). Participants are instructed on how to install the app. An individually challenging minimum is jointly determined based on their fitness level and recent training history. The same determined pace is used for both runs. This ensures comparable physical demand across conditions.

Run 1. First, the participants are requested to do a 5-minute self-paced warm-up. Next, the run is configured within the app with the assigned condition (counterbalancing) and the participant's headphones are connected. Music and other audio are prohibited to isolate the intervention. Participants complete a 5 km run around the Karpendonkse Plas in Eindhoven. This location is chosen for its open area, preventing large buildings or trees from interfering with the GPS signal. Within three minutes of finishing, participants complete the post-run questionnaire via Qualtrics, to which they are directed within the app.

The two open-ended questions (Section 4.5.3) were by oversight omitted from the original Qualtrics survey. Responses were therefore collected verbally for all participants immediately after each run and recorded in written notes by the researcher at the time of delivery. This procedure introduces potential interviewer bias, as participants may have responded differently in a face-to-face setting than in a written format (Podsakoff et al., 2003). This limitation is considered when interpreting the qualitative findings in Section 5.10.

Run 2. After a minimum three-day washout, participants return to the same location and complete the second 5 km run on the same route, now with the alternate feedback condition. The procedure is identical to Run 1. After completing the final questionnaire, participants are fully debriefed about the two feedback styles, the partial blinding, and the study's goals.

Objective pacing data is matched using anonymised participant identifiers and exported from the app to the post-run Qualtrics questionnaire.

4.7 Data Analysis

Statistical analyses are conducted using IBM SPSS Statistics. All primary and secondary outcome variables are analysed using Wilcoxon signed-rank tests, comparing each participant's score in Condition A against their score in Condition B. The Wilcoxon signed-rank test is chosen over the paired samples t-test because the anticipated sample size is too small to reliably assess normality, and because several outcome variables are ordinal. Effect sizes are reported as $r = \frac{Z}{\sqrt{N}}$. The significance threshold is set at $\alpha = .05$. Given the small sample size, statistical tests are interpreted with caution. Effect sizes are prioritised over significance values; a non-significant result is attributed to insufficient power rather than the absence of an effect.

The primary outcome variable, average pace, is computed per run as the elapsed time between the Run Started and Run Stopped log entries divided by the GPS-measured distance. This timestamp-based calculation is independent of the 6-second pace samples and eliminates the influence of GPS speed measurement noise on the primary outcome variable.

Secondary behavioural outcomes are derived directly from the event log. The Pace Sample entries are used to calculate the percentage of the run spent at or above the minimum pace. The number of corrective cues triggered per run is counted from the Audio Cue Triggered events. Internal consistency of the perceived cognitive processing scale (H2) is assessed using Cronbach's α. Perceived psychological stress (H3) and RPE are analysed as single-item scores. Motivational reactance is analysed as the mean of two items. Manipulation check responses are analysed descriptively and with a Wilcoxon signed-rank test to confirm that the ego-threatening condition was perceived as more confrontational and less supportive than the autonomy-supportive condition.

The moderating role of general self-efficacy on the relationship between condition and average pace (H4) is examined using the PROCESS macro Model 1. Given the anticipated sample size, this analysis is treated as exploratory and interpreted with caution. Open-ended responses are analysed using inductive thematic analysis to identify recurring patterns in participants' cognitive and behavioural reactions to the coaching cues.

The cue-response pace profile is not subjected to inferential testing, as it carries no standalone hypothesis. Instead, it is examined descriptively alongside the Motivational Reactance scores to provide a behavioural counterpart to that self-report measure: where Motivational Reactance captures whether the fight impulse was subjectively experienced, the cue-response profile shows whether it translated into an immediate change in running behaviour. Group-level averages are plotted per condition to visualise differences in the speed and magnitude of post-cue pace recovery.

5. Results

This chapter reports the results of the within-subjects pilot study comparing autonomy-supportive and ego-threatening auditory coaching feedback across three primary outcome domains: running performance (H1), perceived cognitive processing (H2), and perceived psychological stress (H3). The moderating role of General Self-Efficacy (H4) is evaluated without inferential testing due to the sample size constraints. Motivational reactance and Rating of Perceived Exertion (RPE) are reported as secondary measures. Qualitative open-ended responses are analysed to contextualise the quantitative findings.

Before presenting the results, two structural limitations of the present dataset must be explicitly noted, as they fundamentally constrain every inferential interpretation in this chapter. First, the post-hoc exclusion of 9 of 16 recruited participants reduced the analytic sample to $n = 7$. This rate is far beyond what was anticipated at study design and renders all inferential tests statistically underpowered. Second, due to logistical constraints during field testing and the subsequent participant exclusions, a severe order imbalance occurred: six of the seven retained participants completed the autonomy-supportive condition (A) before the ego-threatening condition (B). The within-subjects design was intended to control for individual differences through counterbalancing, but with only one participant in the B→A order (P03), this control has effectively failed. These two issues are the primary threats to internal validity in this study and are further discussed in Chapter 6. All statistical results in this chapter must be interpreted with these constraints in mind.

5.1 Participant Sample and Exclusions

Of the 16 participants recruited, 9 were excluded based on the post-hoc criteria detailed in Section 4.5.4. Five of those 9 participants received zero corrective cues in at least one condition, indicating that their individually set target pace was not sufficiently challenging to trigger the corrective threshold across an entire 5 km run. One participant demonstrated a structural pace deficit across both conditions — meaning their average pace persistently fell short of the minimum target, making a meaningful between-condition comparison impossible. The remaining three participants were unable to complete the second trial due to adverse weather conditions and scheduling constraints.

The final analytic sample comprised $n = 7$ participants (5 men, 2 women; age range 23–57; $M = 31.7$, $Mdn = 25$). Scores on the General Self-Efficacy scale (possible range 10–40) varied between 28 and 36 ($M = 32.7$, $SD = 2.6$). This reflects a relatively high baseline self-efficacy within the sample.

Table 2
Individual sample characteristics.

Participant	Age	Gender	Frequency	Target pace	GSE	Order
P01	24	M	1×/week	6:30/km	28	A→B
P02	23	F	1×/week	4:24/km	31	A→B
P03	27	M	1×/week	4:35/km	34	B→A
P04	25	M	2–3×/week	5:00/km	36	A→B
P05	57	M	2–3×/week	6:38/km	33	A→B
P06	41	F	2–3×/week	5:05/km	34	A→B
P07	25	M	2–3×/week	4:20/km	33	A→B

5.2 Manipulation Check and Validation Measures

Confrontational ratings. Participants perceived the ego-threatening condition as more confrontational ($M = 4.14$, $SD = 0.69$) than the autonomy-supportive condition ($M = 2.57$, $SD = 1.13$). The Wilcoxon signed-rank test for all seven participants yielded $T = 2.0$, $Z = -2.05$, $p = .040$, $r = .77$. This confirms that the ego-threatening condition was significantly more confrontational than the autonomy-supportive condition.

Supportiveness ratings. Supportiveness ratings were lower in the ego-threatening condition ($M = 3.14$, $SD = 0.90$) compared to the autonomy-supportive condition ($M = 4.00$, *SD* = 1.00). After excluding P06 (who reported identical ratings), the Wilcoxon signed-rank test for the remaining six participants yielded $T = 4.5$, $Z = -1.30$, $p = .194$, $r = .49$. Although the majority of participants experienced the ego-threatening condition as less supportive, the result did not reach statistical significance, largely due to the contrary response pattern of one participant (P07).

Anomaly — P06. P06 reported identical supportiveness ratings for both conditions. Regarding her confrontational ratings, she perceived the autonomy-supportive condition as more confrontational (5) than the ego-threatening condition (4). Her open-ended response clarified: "I was primarily confronted with my own flaws. I found that quite intense and didn't expect it." This indicates that even the autonomy-supportive cues were perceived as direct by at least one participant, a finding that warrants further consideration regarding the subjective nature of feedback reception. P06's data are retained in the primary analysis.

Anomaly — P07. P07 showed a response pattern contrary to the group trend, perceiving the ego-threatening condition as more supportive (4) than the autonomy-supportive condition (2). Qualitative feedback indicated a preference for directive coaching: "The stricter form of feedback suits me better; I feel like I need this approach." This suggests that perceived supportiveness can be mediated by individual coaching preferences.

RPE. RPE scores (Borg CR10) were broadly comparable across conditions ($M_{AS} = 6.14$, $SD = 2.04$; $M_{ET} = 6.29$, $SD = 1.25$). A Wilcoxon signed-rank test yielded $T = 4$, $Z = -.38$, $p = .705$, $r = .14$. ET feedback did not yield lower perceived exertion than AS.

Individual scores for all psychological measures are presented in Table 3.

Table 3
Individual manipulation check and psychological outcome scores per condition (all items 5-point Likert).

Participant	Sup A	Sup B	Conf A	Conf B	Cog A	Cog B	Str A	Str B	Rea A	Rea B	RPE A	RPE B
P01	5	4	3	4	1.67	2.33	1	2	3.5	4.0	8	7
P02	4	2	2	5	2.00	2.33	1	3	3.5	2.0	4	6
P03	4	2	2	3	1.00	2.00	1	4	4.5	2.5	4	6
P04	5	3	2	4	3.67	2.00	2	2	2.0	4.5	8	6
P05	4	3	2	4	1.33	2.67	2	3	4.0	4.0	7	7
P06	4	4	5	4	2.33	2.00	1	1	4.0	4.0	4	4
P07	2	4	2	5	3.67	1.67	2	2	2.0	4.0	8	8
M	4.00	3.14	2.57	4.14	2.24	2.14	1.43	2.43	3.36	3.57	6.14	6.29
SD	1.00	0.90	1.13	0.69	1.07	0.32	0.53	0.98	0.99	0.93	2.04	1.25

Note. Sup = supportiveness; Conf = confrontational; Cog = perceived cognitive processing (mean of 3 items); Str = psychological stress; Rea = motivational reactance (mean of 2 items). A = AS condition; B = ET condition.

5.3 H1: Primary Outcome — Running Performance

H1 predicted that participants receiving ego-threatening feedback would achieve a faster average pace than those receiving autonomy-supportive feedback. Individual pace data per condition are presented in Table 4.

Contrary to H1, the majority of participants ($n = 5$) performed faster in the autonomy-supportive condition. P06 showed the largest AS advantage (11 s/km), while P04 showed the largest ET benefit (8 s/km). The group-level mean paces were $M = 311.7$ s/km ($SD = 55.2$) for the autonomy-supportive condition and $M = 314.6$ s/km ($SD = 55.9$) for the ego-threatening condition. A Wilcoxon signed-rank test yielded $T = 6.5$, $Z = -1.27$, $p = .203$, $r = .48$. Consequently, H1 is not supported.

Table 4
Average pace per participant per condition.

Participant	Target (s/km)	Pace AS (s/km)	Pace ET (s/km)	Difference (ET − AS)	Direction
P01	390	378 (6:18)	382 (6:22)	+4	AS faster
P02	264	270 (4:30)	269 (4:29)	−1	ET faster
P03	275	274 (4:34)	275 (4:35)	+1	AS faster
P04	300	301 (5:01)	293 (4:53)	−8	ET faster
P05	398	398 (6:38)	402 (6:42)	+4	AS faster
P06	305	305 (5:05)	316 (5:16)	+11	AS faster
P07	260	256 (4:16)	265 (4:25)	+9	AS faster
M (*SD*)	—	311.7 (55.2)	314.6 (55.9)	+2.9	—

Note. Lower values indicate faster pace. Difference = ego-threatening minus autonomy-supportive in seconds per kilometre; positive values indicate slower performance under ego-threatening. Pace values are derived from elapsed run time divided by GPS-measured distance (timestamp method), rounded to the nearest second.

5.4 Secondary Behavioural Outcomes

Time at or above minimum pace. The proportion of run time spent at or above the minimum pace was higher under AS for six of seven participants ($M_{AS} = 53.7\%$, $SD = 16.9$; $M_{ET} = 47.4\%$, $SD = 16.1$). Only P04 spent more time on target under ET (63.6% vs. 44.5%). A Wilcoxon signed-rank test yielded $T = 6.0$, $Z = -1.35$, $p = .176$, $r = .51$, indicating no significant difference. The group means differ by 6.3 percentage points in favour of AS, directionally consistent with the pace findings.

Corrective cues triggered. Corrective cues were more frequent in the ego-threatening condition ($M_{AS} = 3.29$, $M_{ET} = 3.86$). More corrective cues indicate more threshold crossings, meaning participants drifted below their minimum pace more often under ego-threatening feedback. This is directionally consistent with both the slower group pace

Table 5
Secondary behavioural outcomes per participant per condition.

Participant	% on target AS	% on target ET	Cues AS	Cues ET
P01	71.9%	66.7%	1	3
P02	32.7%	28.3%	5	5
P03	48.7%	48.0%	7	6
P04	44.5%	63.6%	4	4
P05	68.0%	59.2%	2	4
P06	37.4%	34.2%	2	3
P07	72.9%	32.0%	2	2
M (*SD*)	53.7 (16.9)	47.4 (16.1)	3.29 (2.14)	3.57 (1.90)

Note. % on target = proportion of pace samples at or below (i.e., at or faster than) the minimum target pace.

and the lower percentage of time on target under ET, and counter to H1. This measure was not statistically tested given the small sample and count nature of the variable.

5.5 H2: Perceived Cognitive Processing

Internal consistency of the three-item Perceived Cognitive Processing scale was poor overall (Cronbach's $\alpha = .30$, $N = 14$ observations), with item-total correlations ranging from .03 to .47. Per-condition reliability analyses revealed a striking asymmetry: internal consistency was acceptable under AS ($\alpha = .740$), with item-total correlations ranging from .35 to .84, but yielded a strongly negative value under ET ($\alpha = -3.525$), with all item-total correlations negative (range: $-.77$ to $-.48$). SPSS flagged this as a violation of reliability model assumptions due to negative average covariance among items.

Individual composite scores and group means are presented in Table 3. The group means showed a small numerical difference across conditions (M_{AS} = 2.24, *SD* = 1.07; M_{ET} = 2.14, *SD* = 0.33). Four participants reported higher perceived cognitive processing under ET and three under AS, with no tied pairs. A Wilcoxon signed-rank test revealed no significant difference (*T* = 13.50, *Z* = $-.09$, *p* = .933, *r* = .03). H2 is not supported.

5.6 H3: Perceived Psychological Stress

H3 predicted that ET feedback would be associated with higher perceived psychological stress than AS feedback. Individual scores and group means are presented in Table 3. Perceived psychological stress was higher under ET than AS for four participants, with three tied and none reporting lower stress under ET (M_{AS} = 1.43, *SD* = 0.53; M_{ET} = 2.43, *SD* = 0.98). A Wilcoxon signed-rank test on the four non-tied pairs yielded *T* = 0, *Z* = -1.84, *p* = .066, *r* = .70. The result does not reach α = .05, attributable to the reduced effective sample after excluding tied pairs. The effect size is large and the direction perfectly consistent: zero participants reported higher stress under AS. H3 receives directional support but does not reach statistical significance.

5.7 Motivational Reactance

Reactance scores showed no consistent directional pattern and no group-level difference across conditions (M_{AS} = 3.36, SD = 0.99; M_{ET} = 3.57, SD = 0.93). Three participants reported higher reactance under ET, two under AS, and two reported equal scores. A Wilcoxon signed-rank test on the five non-tied pairs yielded T = 5.50, $Z = -.54$, p = .588, r = .20, confirming no significant difference.

Cue-response pace profile. A descriptive group-level cue-response profile was generated to visualise the immediate behavioural reaction to the feedback (Figure 2). This profile maps the average pace deviation in the metres immediately preceding and following a corrective cue across both conditions (200 m before, 600 m after). It serves as the objective behavioural complement to the Motivational Reactance self-report above: where reactance captures whether the fight impulse was subjectively experienced, the cue-response profile captures whether it translated into an immediate pace change. Visually, the group-level averages (bold lines) suggest a recovery towards the target pace in both conditions. However, the underlying individual data (faint background lines) exhibit extreme variance. Because group-level averages are highly sensitive to individual outliers at this sample size, the current dataset is insufficient to draw reliable conclusions regarding this metric, and no inferential test is applied.

5.8 H4: General Self-Efficacy as Moderator (Exploratory)

H4 predicted that GSE would positively moderate the effect of ET feedback on pace, with higher-GSE athletes showing greater pace improvement under ET. With n = 7, the planned PROCESS macro moderation analysis (Model 1) is statistically uninformative. Any effect estimate would carry confidence intervals too wide to be interpretable. The analysis is therefore not reported. H4 cannot be evaluated with the current sample.

One descriptive observation is noted, with the explicit caveat that it is illustrative rather than evidential. The participant with the highest GSE in the sample (P04, GSE = 36) showed the largest performance benefit under ET (-8 s/km), directionally consistent with H4. The remaining six participants show no monotonic relationship between GSE and the ET–AS pace difference. No inference is warranted. H4 remains an open empirical question that requires a purpose-powered study to address.

5.9 Qualitative Findings: Open-Ended Responses

Open-ended responses were analysed using inductive thematic analysis. All participant responses were originally given in Dutch and have been translated into English by the researcher. As noted in Section 4.6, responses were collected verbally and recorded in written notes, which introduces potential interviewer bias (Podsakoff et al., 2003); this should be considered when interpreting the findings below. Participants were asked two questions after each run: (1) Describe what went through your mind when the

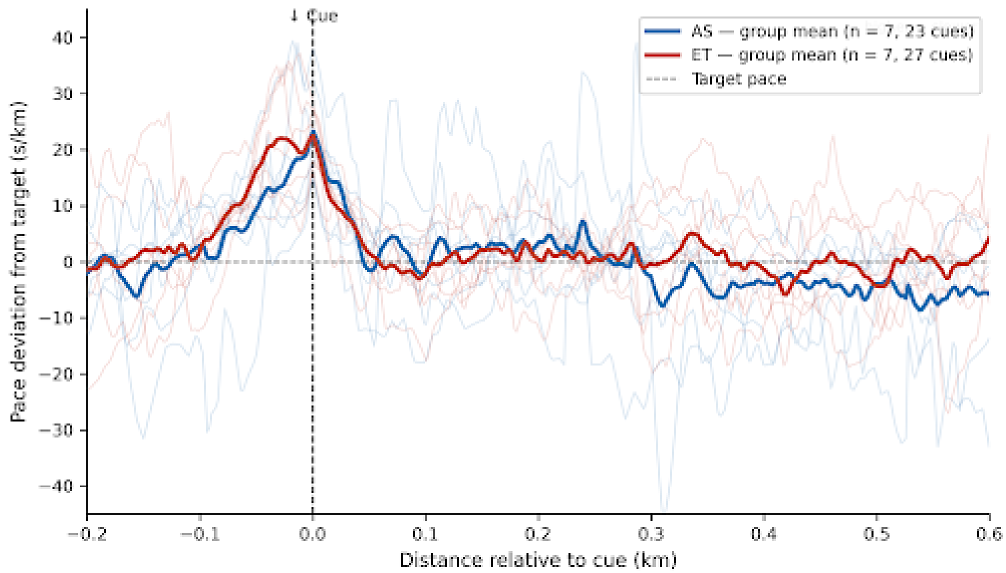

Figure 2
Cue-response pace profile: average pace deviation from individual target pace in the 200 m preceding and 600 m following a corrective cue, across both conditions. Positive values indicate running slower than target. Faint lines represent individual participant means; bold lines represent the group mean across all cues (AS: n = 7, 23 cues; ET: n = 7, 27 cues).

corrective cue played, and what you did next"; and (2) Was there anything about the coaching that bothered or helped you? What exactly?" The first question targeted the immediate cognitive and behavioural response to the corrective cues; the second captured broader usability and perception differences between conditions. Two themes were identified.

Theme 1: ET Cues Elicited a Corrective Behavioural Response Through Distinct Motivational Routes. Six of seven participants described some form of behavioural or motivational response following an ET cue: three reported explicit acceleration (P04, P05, P06), one described active pace correction (P02), one described avoidance-driven pace maintenance (P01), and one reported feeling motivated to push harder (P07). Only P03 described a purely disruptive effect. Three motivational routes were identified.

Two participants described an avoidance-driven response, correcting or maintaining pace to prevent another corrective ego-threatening cue:

> P01: "I wanted to prevent dropping below the pace, otherwise I'd get another reprimand."

> P02: "I wanted to correct my pace as quickly as possible and make sure I didn't get another harsh comment like that."

Three participants described an affective response, in which the confrontational tone of the cue itself appeared to trigger a motivational impulse:

> P04: “It reminds me of my football coach. That definitely makes me run faster.”

> P05: “I was a bit startled by his tone, but I did immediately start running faster after that.”

> P07: “Nice, motivated me to push harder. Direct but clear. Only sometimes I just really didn’t have it in me anymore."

One participant described a response based on instructional clarity — the cue was perceived as direct and unambiguous, making the required action immediately apparent:

> P06: “Direct and clear. I immediately knew what to do. . . accelerate.”

For three participants, however, the initial response was qualified or broke down over the course of the run. P02 attributed this explicitly to fatigue: “I found it very strict and didn’t want to drop below pace just to avoid a cue. When fatigue made that impossible, I didn’t enjoy it.” P07, despite describing ET as motivating, acknowledged a similar limit: “Nice, it motivated me to push harder. . . but sometimes it just wasn’t possible anymore.” P03 described a disruptive rather than motivating effect: “You really get pulled out of your flow. . . I did accelerate briefly, but it mostly triggered frustration.”

Theme 2: AS Cues Functioned Primarily as an Informational Pacing Tool. Under the AS condition, the most commonly reported function of the cues was reducing the need to monitor pace via a device. Three participants described the cues as a convenient external reference point:

> P01: “Nice that you don’t have to look at anything. I go on autopilot and stay on pace.”

> P02: “I actually quite liked it. It felt like someone was running alongside me.”

> P05: “Good that I don’t have to keep checking my watch.”

One participant found this informational function useful but the tone insufficiently urgent:

> P04: “A bit too cautious for my taste. A bit annoyingly gentle.”

Two participants did not fit this pattern. P06 did not experience the AS cues as supportive or informational, but as a confrontation with her own performance gaps: “I was primarily confronted with my own shortcomings. I found that quite intense and didn’t expect it.” P07 reported progressively disengaging from the cues as fatigue increased:

"Towards the end I switched it off in my head and it didn't really register. I was more focused on motivating myself."

5.10 Summary of Results

Table 6
Summary of hypothesis outcomes.

H	Prediction	Direction	Statistical result	Status
H1	ET faster pace	Counter: AS faster (5/7)	$T = 6.50$, $Z = -1.27$, $p = .203$, $r = .48$	Not supported
H2	ET less cognitive load	No clear direction	$T = 13.50$, $Z = -.09$, $p = .933$, $r = .03$	Not supported
H3	ET more stress	Consistent (4/7 higher, 3 tied)	$T = 0$, $Z = -1.84$, $p = .066$, $r = .70$	Directional support only
H4	GSE moderates ET effect	n insufficient	Not tested	Cannot be evaluated

Note. All statistical tests are Wilcoxon signed-rank. r = rank-biserial correlation as effect size. The order confound (6/7 participants completed AS first) is a primary threat to validity of all between-condition comparisons.

No hypothesis was confirmed. The strongest quantitative signal is H3 (stress), which showed a fully consistent direction and large effect size despite insufficient power.

6. Discussion

No hypothesis was confirmed. The primary performance effect was counter-directional, the cognitive processing measure showed no difference, and only perceived psychological stress produced a consistent directional signal. This chapter discusses the findings in relation to the theoretical framework, examines what the results do and do not reveal about the proposed mechanisms, and reflects on the methodological constraints and directions for future research.

6.1 Interpreting the Performance Null

H1 predicted a pace advantage for ego-threatening feedback. The data showed the opposite: five of seven participants ran faster under the autonomy-supportive condition, with a group mean difference of approximately 3 s/km in favour of AS. This finding cannot, however, be read as evidence that autonomy-supportive feedback outperforms ego-threatening feedback. Because six of seven participants completed the AS condition first, any observed difference between conditions is confounded with run order and cannot be attributed to feedback style alone.

A first-run advantage of 3 s/km is well within the range that could be explained by heightened motivation during initial observation and the absence of physiological carryover from a prior effort. It is equally plausible that a reversed order would have further disadvantaged ET performance, as participants in the second run typically benefit from prior route familiarity and protocol experience. Because counterbalancing effectively failed, the $\approx$3 s/km AS advantage cannot be disentangled from these competing explanations. All performance comparisons must therefore be treated as preliminary observations rather than causally attributable effects.

6.2 The Qualitative–Quantitative Gap: Immediate Reactance Without Sustained Pace Benefit

A notable tension exists between the qualitative and quantitative findings. As reported in Section 5.10, six of seven participants described some form of motivational or behavioural response to ET cues. Yet this did not translate into a group-level pace advantage, and the self-report reactance scale showed no directional difference across conditions.

Two explanations for this gap are plausible and not mutually exclusive. First, the measurement instrument may not have captured what the qualitative data suggest actually occurred. The reactance scale asked participants to retrospectively evaluate their motivational response to the feedback style as a whole. Yet a single post-run score collapses the entire run into one rating, obscuring any within-run variation in response. Moreover, retrospective ratings are vulnerable to recency effects, whereby cues delivered later in the run are likely weighted more heavily in overall recall (Kahneman et al., 1993). In addition, this is most probably when fatigue was highest and any reactive response most suppressed. For example, a participant who felt a brief impulse to accelerate after one cue but disengaged later may have reported the same overall score as a participant who felt nothing at all. The phrasing of the items — "the feedback made me want to prove it wrong" and "the feedback motivated me to push harder immediately" — may also have been interpreted as referring to the general condition experience rather than the specific moment of a corrective cue delivery. This is further compounded by the 500-metre check-in cues, which were delivered at fixed intervals throughout both conditions. Participants' reactance ratings may have been dominated by the check-in cues rather than the threshold-crossing episodes the reactance mechanism specifically targets. Therefore, the corrective cues designed to elicit reactance could have been too infrequent to register in an overall post-run rating. Second, the qualitative evidence suggests that the reactance response was transient. P02 reported that the avoidance motivation remained intact but became physically unachievable as fatigue set in: "I found him very strict and didn't want to drop below the pace so I wouldn't get a cue. When that didn't work out because of fatigue, I didn't like it." P07 acknowledged that ET cues were motivating "but sometimes I just did not have it in me anymore." P03 described disruption rather than motivation. This pattern is consistent with the Challenge–Threat framework (Meijen et al., 2020): reactance may produce a local motivational spike, but its conversion into sustained effort depends on whether the athlete possesses sufficient coping resources at the moment of the cue. Under progressive fatigue, those resources are precisely what diminishes. Third, the order confound may have masked an ET effect that was real but small — as discussed in Section 6.1.

The Dual Process framework offers a partial explanation for this pattern. The initial corrective acceleration immediately after an ET cue aligns with the theorised System 1-level impulse: a reflexive response to a competence challenge that may bypass deliberate evaluation. However, the framework may have underestimated what follows. Sustaining pace across a 5-kilometre run under fatigue is not a System 1 task. It requires ongoing inhibitory control and deliberate effort regulation, both System 2 processes. A single automatic impulse may interrupt the downward trajectory momentarily, but is unlikely to substitute for the sustained deliberative work that pace maintenance demands. Under progressive fatigue, depleted System 2 resources leave little capacity to convert the initial reactive burst into lasting performance. The current data are consistent with this interpretation, but the sample size and order confound prevent firm conclusions.

6.3 Cognitive Processing and the Scale Validity Problem

H2 predicted that ET feedback would require less deliberate cognitive processing than AS feedback. The scale yielded a near-zero group difference and a statistically null result, but the more informative finding is the internal consistency pattern, which suggests the scale may not have been capable of detecting the theorised effect in the first place.

Two explanations may account for this pattern. First, the negative α under ET ($\alpha = -3.525$) indicates a negative average covariance among the three items. This most likely reflects interpretation variance: the three items asked participants to evaluate abstract, subjective processing qualities that may have been difficult to interpret consistently, regardless of condition.

Second, the scale was likely too broad to distinguish between participants' responses to corrective cues and their responses to the check-in cues delivered throughout both conditions. Participants' overall ratings may have reflected their response to those cues as much as to the corrective cues that H2 targets. Furthermore, asking participants to evaluate abstract processing qualities retrospectively across an entire run obscures any within-run variation — the same limitation identified for the reactance scale in Section 6.2. The implication is not that H2 is empirically resolved, but that the measurement approach was insufficiently sensitive. The cue-response pace profile included in this study offers a behavioural proxy for such effects, but the sample size was insufficient to draw reliable conclusions from it. Future directions for this line of research are discussed in Section 6.7.

6.4 Psychological Stress: The Most Consistent Signal

H3 produced the study's most interpretable finding. Perceived psychological stress was higher under ET for four of seven participants, with the remaining three tied and none reporting higher stress under AS. The effect size was large ($r = .70$), and the direction was consistent. The absence of a corresponding RPE difference confirms that the stress elevation was genuinely psychological rather than an artefact of greater physical effort in the ET condition. Participants experienced the same physical intensity but a meaningfully different psychological experience.

This finding directly validates the core performance–risk trade-off this thesis was designed to examine. Even in the absence of a confirmed performance benefit, the data indicate that ET feedback carries a psychological cost. For a practical coaching application, this cost has direct relevance. An intervention that elevates stress without producing measurable performance gains would be difficult to justify, particularly across repeated use. Steel, Bishop, and Taylor (2021) found that controlled motivational conditions produced greater short-term effort but were accompanied by elevated cortisol. The present results diverge on the performance side — no pace advantage was observed under ET — but converge on the stress dimension. Several participants nonetheless reported preferring the ET condition qualitatively. This suggests that perceived stress and subjective preference do not necessarily align, and that individual differences in coaching style preference may moderate the cost-benefit ratio of ego-threatening feedback. Whether elevated psychological stress accumulates across training sessions in ways that undermine long-term adherence is a question this pilot cannot answer but that any scaled deployment of ego-threatening coaching would need to address.

6.5 Manipulation Integrity and Individual Differences in Feedback Perception

The manipulation check produced a significant difference on confrontational ratings ($p = .040$) but a non-significant difference on supportiveness ratings ($p = .194$). The partial failure of the manipulation on the supportiveness dimension introduces ambiguity into all between-condition comparisons: it is possible that some participants did not experience the conditions as categorically distinct. The anomalous patterns of P06 and P07 are instructive here. P06 perceived the autonomy-supportive condition as more confrontational than the ego-threatening condition, attributing this to confrontation with her own performance shortcomings rather than to the tone of the cues. P07 rated ET as more supportive than AS, explicitly preferring directive coaching.

These cases illustrate that feedback perception is not determined solely by the linguistic framing of cues: prior coaching preferences, self-perception, and the subjective experience of the run itself all mediate how a cue is received. As P06 and P07 demonstrate, an autonomy-supportive cue can be experienced as ego-threatening, and an ego-threatening cue as motivating — meaning the intended psychological distinction between conditions does not always map onto the experienced one. When perception varies this much between individuals, condition labels tell only part of the story. A related issue concerns the manipulation check itself. The confrontational rating used to verify the manipulation captures a different construct than ego-threat: a cue can be experienced as confronting without challenging the athlete's competence or identity in the way ET was theorised to operate. Moreover, even within the same condition, individual cues likely varied in how strongly they were experienced, a variance that a single post-run condition rating cannot capture. A more robust design would supplement the post-run manipulation check with a brief cue-level rating immediately after each corrective episode, allowing researchers to verify perceived tone at the moment the mechanism is theorised to operate.

6.6 Limitations

Several limitations constrain the interpretability of the current findings.

The order confound is the primary threat to internal validity, as discussed in Section 6.1. Its consequences extend beyond the primary outcome: any between-condition difference in perceived stress, cognitive processing, or reactance may also partly reflect order rather than condition. Although counterbalancing was planned, the recruitment losses and scheduling constraints that produced the imbalance were not anticipated. Future studies must treat strict counterbalancing as a non-negotiable design requirement, and should pre-register an analysis that explicitly tests for order effects.

The analytic sample of $n = 7$ renders all statistical tests underpowered. The exclusion of nine of sixteen recruited participants was far beyond what was anticipated at study design. The post-hoc exclusion criteria were methodologically appropriate (participants who never received corrective cues were not exposed to the core manipulation), but the high exclusion rate reflects a design vulnerability. Minimum paces were set collaboratively through a brief pre-experiment discussion, drawing on participants' self-reported habitual pace and researcher guidance to ensure the target felt genuinely challenging, rather than through objective fitness testing. Several participants selected paces that proved too comfortable under actual running conditions, either because the collaborative setting underestimated actual performance capacity or because awareness that a researcher would review their data introduced a performance-boosting observation effect (McCambridge, Witton, & Elbourne, 2014). Future protocols should derive target paces from a baseline time trial — though this adds a third session to an already physically demanding protocol — and consider implementing an adaptive pacing algorithm that adjusts the minimum target in real time when a participant consistently runs far above or below the threshold, safeguarding the integrity of the manipulation throughout the run.

Related to this, several retained participants reported limited prior experience with pace-targeted training. The original inclusion criteria anticipated participants who regularly set and pursued challenging minimum paces. In practice, recruiting sufficient participants within the available timeframe required relaxing this criterion. Variability in pace-goal familiarity may have contributed to the heterogeneity in both behavioural and psychological responses across individuals. Future studies should recruit explicitly from structured training environments where pace-targeted running is standard practice.

The absence of a no-feedback control condition means that the absolute effect of either feedback style cannot be established. The current design answers only the relative question: does ET outperform AS? Whether either style improves on unaided performance, and by how much, remains unknown. A three-condition design including a no-coaching baseline would allow this question to be addressed.

H4 could not be evaluated due to insufficient sample size. The planned moderation analysis required a substantially larger sample than what was retained. Whether GSE determines who responds adaptively to ego-threatening feedback remains an open empirical question and one of the key theoretical questions this pilot leaves unresolved.

The study was conducted using a digital coaching application rather than a human coach. As discussed in Section 2.5, non-human feedback may reduce concerns associated with social evaluation (Dickerson & Kemeny, 2004), potentially allowing participants to engage with ego-threatening cues without the interpersonal defensiveness that human-delivered criticism can produce. Whether the same effects would emerge with a human coach — or whether the human-digital distinction amplifies or attenuates them — cannot be determined from the current data. This limits the generalisability of the findings to digitally mediated coaching contexts specifically.

GPS accuracy, while assessed prior to data collection, could not be verified against an independent speed reference. Within-run standard deviations of 3–4 s/km were observed under near-constant speed conditions, indicating acceptable measurement stability. Crucially, because both runs per participant used the same device and algorithm, any systematic GPS characteristics are held constant across conditions and cannot account for between-condition differences. Absolute accuracy, however, remains unestablished. Moreover, GPS measurement noise may have triggered false corrective cues through threshold crossings that reflected signal fluctuation rather than genuine pace drops. Although the 15-second sustained threshold and EMA smoothing were designed to filter out momentary dips, residual noise cannot be fully excluded. False cues would have introduced irrelevant stimuli into the manipulation, potentially diluting any condition effect.

Weather conditions were not controlled or systematically recorded across sessions. Participants completed the two runs on different days, meaning ambient temperature, humidity, precipitation, and surface conditions could have differed between conditions for any given participant. Elevated heat and humidity are known to increase perceived exertion and reduce running pace independently of psychological factors (Ely et al., 2007), and wet conditions may affect footing and running economy. If conditions differed systematically between runs, this constitutes an uncontrolled within-participant confound.

Finally, the verbal collection of open-ended responses introduces potential interviewer bias, as discussed in Section 4.6. Participants may have moderated their responses in a face-to-face setting compared to written self-report (Podsakoff et al., 2003). This is particularly relevant for the qualitative findings, where social desirability could suppress critical or negative responses about the coaching experience.

6.7 Design Reflection and Future Directions

The present design was more complex than the eventual sample size and data quality could support. A custom mobile application, outdoor GPS-based pace measurement, a within-subjects field protocol, and both quantitative and qualitative outcome measures were combined in a single study. Each of these elements is individually defensible, but their combination introduced a large number of potential confounds that proved difficult to control simultaneously under real-world field conditions. A more controlled first test of the core hypothesis might use a self-pacing treadmill protocol. Treadmills provide a direct, objective speed readout that eliminates GPS measurement error entirely, while auditory cues are delivered through the same headphone setup. This approach also eliminates route familiarity and weather variability as confounds and allows precise experimental control of effort state at the moment cues are delivered. The ecological cost — reduced generalisability to real running conditions — is an acceptable trade-off at the stage of establishing the basic effect before field generalisation is pursued.

The cue-response pace profile included in this study captures the immediate behavioural reaction to each corrective cue and offers a temporally precise test of whether ET cues elicit faster behavioural correction than AS cues, independently of sustained pace differences. The underlying data were logged by the application, but the current sample size was insufficient to draw reliable conclusions. Future studies should treat the cue-response pace profile as a primary outcome rather than a descriptive supplement. For cognitive processing, future research should consider eliminating check-in cues to isolate the effect of corrective cues specifically, and develop scale items that more precisely target the cognitive experience of receiving a single corrective cue rather than asking participants to evaluate an entire run retrospectively.

Future studies might also consider varying race distance. The present study used a 5-kilometre protocol, a distance short enough that progressive fatigue sets in relatively late. At longer distances, the interplay between fatigue and the reactance mechanism may differ. If the reactance response is genuinely transient, longer efforts provide more opportunities for repeated cue delivery, but also greater depletion of the coping resources on which the response depends. Whether distance moderates the effectiveness of ego-threatening feedback is an open question the present design cannot address.

The qualitative data also point to a secondary finding. Several participants reported valuing the check-in cues primarily as a source of real-time pace and distance information, reducing the need to monitor their device. This is theoretically relevant: visual display monitoring during running disrupts cadence and produces micro-accelerations (Van Hooren et al., 2020), which deplete energy and reduce running economy (Bood et al., 2013). Whether audio check-in cues can fully substitute for device-monitoring — and if so, in what tone — has not, to our knowledge, been directly examined. This is a question future work could pursue, but is distinct from the question this thesis aims to answer.

The central empirical question this study raises — whether ET feedback triggers an immediate corrective acceleration that fades before it can accumulate into a faster overall pace — remains open. Answering it will require a design that separates cue-level behavioural responses from run-level averages, ensures adequate counterbalancing, and recruits a sample with both sufficient size and homogeneous experience with pace-targeted training.

7. Conclusion

This thesis set out to examine whether ego-threatening auditory feedback, delivered by a digital coach, produces faster average running pace than autonomy-supportive feedback during a high-intensity 5-kilometre run, and at what cognitive and psychological cost. The research question cannot be answered affirmatively on the basis of the present data. No hypothesis was confirmed, and the methodological constraints preclude causal attribution of any observed between-condition difference to feedback style alone.

The absence of confirmed hypotheses does not mean the study found nothing. First, ego-threatening feedback carried a consistent psychological cost: perceived stress was higher under ET for every participant who showed any difference at all, with a large effect size that held despite insufficient statistical power. This cost is relevant for any practical deployment of ego-threatening coaching within digital coaching interventions and aligns with the cortisol elevation reported by Steel, Bishop, and Taylor (2021) under controlled motivational conditions.

Second, the qualitative data reveal a pattern that the quantitative instruments were unable to capture. Six of seven participants described some form of motivational or behavioural response to ET cues. Yet average pace was not faster under ET and the reactance scale showed no group-level difference ($M_{\text{AS}} = 3.36$ vs. $M_{\text{ET}} = 3.57$, $p = .588$). This suggests that ET cues may trigger a brief local acceleration that fades before it can accumulate into a lower overall pace. Whether the response fades because progressive fatigue depletes the coping resources needed to sustain it, or because the current design was simply too underpowered and confounded to detect a real effect, cannot be determined from the present data. This distinction is the central open question the study leaves behind.

Third, the cognitive processing scale failed as a measurement instrument under ET conditions. The negative Cronbach's alpha ($\alpha = -3.525$) indicates that the reliability model's assumptions were violated entirely. This means H2 remains unresolved. The theorised System 1 advantage of ego-threatening feedback remains a legitimate hypothesis that requires an appropriate instrument, likely a cue-level behavioural measure rather than a retrospective self-report scale, to test.

From a theoretical standpoint, the findings are most consistent with a modified reading of the Dual Process framework: ET cues may trigger an initial automatic fight impulse, but sustaining pace across a 5-kilometre run under progressive fatigue is a System 2 task, and a single reactive burst is insufficient to substitute for the deliberate effort regulation that pace maintenance demands. The reactance mechanism proposed by Brehm (1966) may be real and reachable through a digital medium, but its conversion into sustained performance depends on coping resources that fatigue progressively depletes.

For the design of digital coaching applications, these findings suggest caution. An intervention that raises psychological stress without improving performance is difficult to justify for practical use, particularly across repeated sessions. The autonomy-supportive condition, by contrast, functioned effectively as an informational pacing tool and was received positively by the majority of participants. This is a finding that points to real value in audio-based bandwidth feedback even without ego-threatening content.

The most important contribution of this thesis may be methodological. It demonstrates that studying the cue-level dynamics of ego-threatening feedback requires a design that can separate immediate behavioural responses from run-level averages, isolate corrective cues from check-in cues, ensure strict counterbalancing, recruit participants with verified pace-goal experience, and control for environmental conditions that independently affect performance. None of these requirements proved easy to satisfy in a real-world field protocol. In retrospect, testing a novel and psychologically complex intervention in a field setting before its basic mechanism has been established under controlled conditions may have been overambitious. The next test of the provocative coach hypothesis should begin not with the field, but with the controlled environment where its basic mechanism can first be established.

References

Alkasasbeh, W. J., & S. H. Akroush. 2025. Sports motivation: A narrative review of psychological approaches to enhance athletic performance. *Frontiers in Psychology*, 16:1645274. https://doi.org/10.3389/fpsyg.2025.1645274.

Ames, Carole. 1992. Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, 84(3):261–271. https://doi.org/10.1037/0022-0663.84.3.261.

Bandura, Albert. 1997. *Self-Efficacy: The Exercise of Control*. W. H. Freeman and Company.

Barger, A. S. 2025. Artificial intelligence vs. human coaches: Examining the development of working alliance in a single session. *Frontiers in Psychology*, 15:1364054. https://doi.org/10.3389/fpsyg.2024.1364054.

Bishop, P. A., E. Jones, & A. K. Woods. 2008. Recovery from training: A brief review. *Journal of Strength and Conditioning Research*, 22(3):1015–1024. https://doi.org/10.1519/JSC.0b013e31816eb518.

Bood, Robert Jan, Marijn Nijssen, John van der Kamp, & Melvyn Roerdink. 2013. The power of auditory-motor synchronization in sports: Enhancing running performance by coupling cadence with the right beats. *PLoS ONE*, 8(8):e70758. https://doi.org/10.1371/journal.pone.0070758.

Borg, Gunnar A. 1982. Psychophysical bases of perceived exertion. *Medicine and Science in Sports and Exercise*, 14(5):377–381. https://doi.org/10.1249/00005768-198205000-00012.

Brehm, Jack W. 1966. *A Theory of Psychological Reactance*. Academic Press.

Brehm, Jack W., & Elizabeth A. Self. 1989. The intensity of motivation. *Annual Review of Psychology*, 40:109–131. https://doi.org/10.1146/annurev.ps.40.020189.000545.

Brick, Noel, Tadhg MacIntyre, & Mark Campbell. 2014. Attentional focus in endurance activity: New paradigms and future directions. *International Review of Sport and Exercise Psychology*, 7(1):106–134. https://doi.org/10.1080/1750984X.2014.885554.

Buman, Matthew P., Jens W. Omli, Peter R. Giacobbi, & Britton W. Brewer. 2008. Experiences and coping responses of “hitting the wall” for recreational marathon runners. *Journal of Applied Sport Psychology*, 20(3):282–300. https://doi.org/10.1080/10413200802078267.

Chen, Gilad, Stanley M. Gully, & Dov Eden. 2001. Validation of a new general self-efficacy scale. *Organizational Research Methods*, 4(1):62–83. https://doi.org/10.1177/109442810141004.

Cona, Giorgia, Anna Cavazzana, Antonio Paoli, Giuseppe Marcolin, Alberto Grainer, & Patrizia S. Bisiacchi. 2015. It’s a matter of mind! cognitive functioning predicts the athletic performance in ultra-marathon runners. *PLoS ONE*, 10(7):e0132943. https://doi.org/10.1371/journal.pone.0132943.

Conroy, David E., Jason P. Willow, & Jonathan N. Metzler. 2002. Multidimensional fear of failure measurement: The performance failure appraisal inventory. *Journal of Applied Sport Psychology*, 14(2):76–90. https://doi.org/10.1080/10413200252907752.

Deci, Edward L., & Richard M. Ryan. 1985. *Intrinsic Motivation and Self-Determination in Human Behavior*. Springer. https://doi.org/10.1007/978-1-4899-2271-7.

Deci, Edward L., & Richard M. Ryan. 2000. The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4):227–268. https://doi.org/10.1207/S15327965PLI1104_01.

Diamond, Adele. 2013. Executive functions. *Annual Review of Psychology*, 64:135–168. https://doi.org/10.1146/annurev-psych-113011-143750.

Dickerson, Sally S., & Margaret E. Kemeny. 2004. Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, 130(3):355–391. https://doi.org/10.1037/0033-2909.130.3.355.

Dickerson, Sally S., Peggy J. Mycek, & Frank Zaldivar. 2008. Negative social evaluation, but not mere social presence, elicits cortisol responses to a laboratory stressor task. *Health Psychology*, 27(1):116–121. https://doi.org/10.1037/0278-6133.27.1.116.

Dietrich, Arne. 2006. Transient hypofrontality as a mechanism for the psychological effects of exercise. *Psychiatry Research*, 145(1):79–83. https://doi.org/10.1016/j.psychres.2005.07.033.

Elo, Anna-Liisa, Anneli Leppänen, & Antti Jahkola. 2003. Validity of a single-item measure of stress symptoms. *Scandinavian Journal of Work, Environment & Health*, 29(6):444–451. https://doi.org/10.5271/sjweh.752.

Ely, Matthew R., Samuel N. Cheuvront, William O. Roberts, & Scott J. Montain. 2007. Impact of weather on marathon-running performance. *Medicine & Science in Sports & Exercise*,

39(3):487–493. https://doi.org/10.1249/mss.0b013e31802d3aba.
Eriksson, M., K. A. Halvorsen, & L. Gullstrand. 2011. Immediate effect of visual and auditory feedback to control the running mechanics of well-trained athletes. *Journal of Sports Sciences*, 29(3):253–262. https://doi.org/10.1080/02640414.2010.523088.
Eysenck, Michael W., Nazanin Derakshan, Rita Santos, & Manuel G. Calvo. 2007. Anxiety and cognitive performance: Attentional control theory. *Emotion*, 7(2):336–353. https://doi.org/10.1037/1528-3542.7.2.336.
Giraldo-Pedroza, A., W. C. Lee, W. K. Lam, R. Coman, & G. Alici. 2020. Effects of wearable devices with biofeedback on biomechanical performance of running – a systematic review. *Sensors*, 20(22):6637. https://doi.org/10.3390/s20226637.
Inzlicht, Michael, & Samuele M. Marcora. 2016. The central governor model of exercise regulation teaches us precious little about the nature of mental fatigue and self-control failure. *Frontiers in Psychology*, 7:656. https://doi.org/10.3389/fpsyg.2016.00656.
Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. Farrar, Straus and Giroux.
Kahneman, Daniel, Barbara L. Fredrickson, Charles A. Schreiber, & Donald A. Redelmeier. 1993. When more pain is preferred to less: Adding a better end. *Psychological Science*, 4(6):401–405. https://doi.org/10.1111/j.1467-9280.1993.tb00589.x.
Kannampallil, T., C. R. Ronneberg, N. E. Wittels, V. Kumar, N. Lv, J. M. Smyth, B. S. Gerber, E. A. Kringle, J. A. Johnson, P. Yu, L. E. Steinman, O. A. Ajilore, & J. Ma. 2022. Design and formative evaluation of a virtual voice-based coach for problem-solving treatment: Observational study. *JMIR Formative Research*, 6(8):e38092. https://doi.org/10.2196/38092.
Lazarus, Richard S., & Susan Folkman. 1984. *Stress, Appraisal, and Coping*. Springer.
Liu, Yunshan, Yuqi Yang, Xuejun Bai, Yue Chen, & Lei Mo. 2022. Do immediate external rewards really enhance intrinsic motivation? *Frontiers in Psychology*, 13:853879. https://doi.org/10.3389/fpsyg.2022.853879.
Lyle, John, & Christopher Cushion. 2017. *Sport Coaching Concepts: A Framework for Coaching Practice*, 2 edition. Routledge. https://doi.org/10.4324/9780203126424.
Mageau, Géneviève A., & Robert J. Vallerand. 2003. The coach–athlete relationship: A motivational model. *Journal of Sports Sciences*, 21(11):883–904. https://doi.org/10.1080/0264041031000140374.
Marcora, Samuele M., & Walter Staiano. 2010. The limit to exercise tolerance in humans: Mind over muscle? *European Journal of Applied Physiology*, 109:763–770. https://doi.org/10.1007/s00421-010-1418-6.
Marcora, Samuele M., Walter Staiano, & Victoria Manning. 2009. Mental fatigue impairs physical performance in humans. *Journal of Applied Physiology*, 106(3):857–864. https://doi.org/10.1152/japplphysiol.91324.2008.
McCambridge, Jim, John Witton, & Diana R. Elbourne. 2014. Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects. *Journal of Clinical Epidemiology*, 67(3):267–277. https://doi.org/10.1016/j.jclinepi.2013.08.015.
McCormick, Alister, Carla Meijen, & Samuele Marcora. 2015. Psychological determinants of whole-body endurance performance. *Sports Medicine*, 45(7):997–1015. https://doi.org/10.1007/s40279-015-0319-6.
Meijen, Carla, Martin Turner, Marc V. Jones, David Sheffield, & Paul McCarthy. 2020. A theory of challenge and threat states in athletes: A revised conceptualization. *Frontiers in Psychology*, 11:126. https://doi.org/10.3389/fpsyg.2020.00126.
Micklewright, Dominic, S. Kegerreis, John Raglin, & Florentina Hettinga. 2017. Will the conscious-subconscious pacing quagmire help elucidate the mechanisms of self-paced exercise? *Sports Medicine*, 47(7):1231–1239. https://doi.org/10.1007/s40279-016-0642-6.
Morris, Laurel S., Morgann M. Grehl, Sarah B. Rutter, Mitul Mehta, & Margaret L. Westwater. 2022. On what motivates us: A detailed review of intrinsic v. extrinsic motivation. *Psychological Medicine*, 52(10):1801–1816. https://doi.org/10.1017/S0033291722001611.
Mossman, Lara H., Gavin R. Slemp, Kelsey J. Lewis, Rachel H. Colla, & Paul O'Halloran. 2022. Autonomy support in sport and exercise settings: A systematic review and meta-analysis. *International Review of Sport and Exercise Psychology*, 17(1):1–24. https://doi.org/10.1080/1750984X.2022.2031252.
Nass, Clifford, & Youngme Moon. 2000. Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1):81–103.

https://doi.org/10.1111/0022-4537.00153.
Nicholls, John G. 1989. *The Competitive Ethos and Democratic Education*. Harvard University Press.
Noakes, Timothy D. 2012. Fatigue is a brain-derived emotion that regulates the exercise behavior to ensure the protection of whole body homeostasis. *Frontiers in Physiology*, 3:82. https://doi.org/10.3389/fphys.2012.00082.
Podsakoff, Philip M., Scott B. MacKenzie, Jeong-Yeon Lee, & Nathan P. Podsakoff. 2003. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(6):879–903. https://doi.org/10.1037/0021-9010.88.5.879.
Puleo, Nicholas A., & Kirk A. Abraham. 2018. External feedback does not affect running pace in recreational runners. *International Journal of Exercise Science*, 11(5):384–390.
Sakalidis, K. E., S. G. P. Menting, M. T. Elferink-Gemser, & F. J. Hettinga. 2022. The role of the social environment in pacing and sports performance. *International Journal of Environmental Research and Public Health*, 19(23):16131. https://doi.org/10.3390/ijerph192316131.
Schmidt, Benedikt, Sebastian Benchea, Ruediger Eichin, & Christian Meurisch. 2015. Fitness tracker or digital personal coach: How to personalize training. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 1063–1067. https://doi.org/10.1145/2800835.2800961.
Schwarzer, Ralf, & Matthias Jerusalem. 1995. Generalized self-efficacy scale. In John Weinman, Stephen Wright, & Marie Johnston, editors, *Measures in Health Psychology: A User's Portfolio. Causal and Control Beliefs*. NFER-NELSON, pages 35–37.
Steel, Richard P., Nicolette C. Bishop, & Ian M. Taylor. 2021. The effect of autonomous and controlled motivation on self-control performance and the acute cortisol response. *Psychophysiology*, 58(11):e13915. https://doi.org/10.1111/psyp.13915.
Steindl, Christina, Eva Jonas, Sandra Sittenthaler, Eva Traut-Mattausch, & Jeff Greenberg. 2015. Understanding psychological reactance: New developments and findings. *Zeitschrift für Psychologie*, 223(4):205–214. https://doi.org/10.1027/2151-2604/a000222.
Stragier, Jeroen, Mariek Vanden Abeele, Peter Mechant, & Lieven De Marez. 2016. Understanding persistence in the use of online fitness communities: Comparing novice and experienced users. *Computers in Human Behavior*, 64:34–42. https://doi.org/10.1016/j.chb.2016.06.013.
Tóth, R., V. Resperger, & L. Tóth. 2024. Optimizing athletes' mindsets: Application of rational emotive behavior therapy (rebt) in sport. *SPRINT – Sports Research International*, 1:25–29. https://doi.org/10.1556/1020.2024.00008.
Van Hooren, Bas, Joel Goudsmit, Juan Restrepo, & Steven Vos. 2020. Real-time feedback by wearables in running: Current approaches, challenges and suggestions for improvements. *Journal of Sports Sciences*, 38(2):214–230. https://doi.org/10.1080/02640414.2019.1690960.
Van Hooren, Bas, Guy Plasqui, & Kenneth Meijer. 2024. The effect of wearable-based real-time feedback on running injuries and running performance: A randomized controlled trial. *The American Journal of Sports Medicine*, 52(3):750–765. https://doi.org/10.1177/03635465231222464.
Weir, Joseph P., Travis W. Beck, Joel T. Cramer, & Terry J. Housh. 2006. Is fatigue all in your head? a critical review of the central governor model. *British Journal of Sports Medicine*, 40(7):573–586. https://doi.org/10.1136/bjsm.2005.023028.
Wickens, Christopher D. 2008. Multiple resources and mental workload. *Human Factors*, 50(3):449–455. https://doi.org/10.1518/001872008X288394.
Wright, Rex A., Stephanie Agtarap, & Christopher Mlynski. 2015. Conversion of reactance motives into effortful goal pursuit. *Zeitschrift für Psychologie*, 223(4):267–276. https://doi.org/10.1027/2151-2604/a000228.

A. Audio Cue Scripts

This appendix provides the complete set of audio cue scripts used in the experiment. Table 1 (in Section 4.3.2) presents the corrective threshold cues. The remaining cue types are presented below.

Table 7
Start and finish cues per condition.

Trigger	**Condition A — Autonomy-Supportive**	**Condition B — Ego-Threatening**
Run started	"Your run has started! Find your rhythm and listen to your body."	"Your run has started! Let's see if you can actually handle the pace you set."
Run ended	"Five kilometres covered! Your run has ended. Great effort out there today."	"Five kilometres covered! Your run has ended. Let's push even harder next time."

Table 8
500-metre check-in cues (above pace). Triggered at 500-metre markers when pace is above the minimum target pace.

km	Condition A — Autonomy-Supportive	Condition B — Ego-Threatening
0.5	"Half a kilometre in, great start! You're right on pace, keep it going."	"Half a kilometre in! This should still feel easy, if it's already hard, it's going to be a long run for you."
1	"First kilometre done! Your pace is right on. Take it one kilometre at a time and keep listening to your body."	"One kilometre in! Plenty of runners look good at 1k. Most fall apart after. Let's see which type you are."
1.5	"One and a half kilometres in! You're finding your rhythm and you're doing really well."	"One and a half kilometres in! Losing pace at this point would be a weak performance."
2	"Two kilometres in and you are holding steady! Great rhythm, stay comfortable and try to keep it going."	"Two kilometres done! Two kilometres is where excuses start. Prove this wasn't just a good opening."
2.5	"Two and a half kilometres in! You're doing great. Try to keep this rhythm if it feels okay."	"Two and a half kilometres in! Giving up now would be disappointing."
3	"Three kilometres done! Your pace is solid and you are past halfway. Keep trusting yourself."	"Three kilometres, on pace! The last two kilometres will expose exactly what you're worth. Don't embarrass yourself now!"
3.5	"Three and a half kilometres. Almost there and you are still doing great!"	"Three and a half kilometres. You can not give up now!"
4	"Four kilometres in! Pace is right where it needs to be."	"Four kilometres in. Pace is holding. Prove you can finish what you started."
4.5	"Five hundred metres to go! If you have any energy left, try to push one final time!"	"Five hundred metres left! If you slow down now, your effort was worth nothing! Keep going!"

Table 9
500-metre check-in cues (below pace). Triggered at 500-metre markers when pace is below the minimum target pace.

km	Condition A — Autonomy-Supportive	Condition B — Ego-Threatening
0.5	"Half a kilometre in, but you have already fallen below your target pace. Try to pick it up a little if you feel able to."	"Half a kilometre in and already off pace. This is not good enough."
1	"First kilometre done, but your pace is below your target. Try to bring it back. You still have four kilometres to turn this around."	"One kilometre in and already behind pace. At this rate, you will never reach your goal."
1.5	"One and a half kilometres in. You are below your target pace. If you feel up to it, try to push a little harder. There is still plenty of time."	"One and a half kilometres in and still off pace. Dropping this early would be embarrassing."
2	"Two kilometres in, but your pace is below target. Try to bring it back if you can. Three kilometres is still enough time to turn it around."	"Two kilometres done and behind pace. You are not performing at the level this goal requires."
2.5	"Two and a half kilometres in! Halfway there, but below your target. Try to push through if that feels okay."	"Two and a half kilometres in and still off pace. Giving in at the halfway point would be deeply disappointing."
3	"Three kilometres done. Past halfway but your pace has dropped. Try to pick it back up if it feels okay."	"Three kilometres in and off pace. The second half will decide whether this run meant anything."
3.5	"Three and a half kilometres. You are doing great, but are still below target. If it feels okay, try to pick it back up."	"Three and a half kilometres, off pace. You are about to fail this in front of yourself. Is that really how you want this to end?"
4	"One kilometre to go. Your pace has dropped, but you can still finish strong. Push if you have any energy left."	"Four kilometres in and still off pace. Do not make this run a complete waste."
4.5	"Five hundred metres to go! You are below pace, but try to push through to the end. This finish is yours!"	"Five hundred metres left! You are off pace. This is your last chance to make this run worth something. Keep going!"

Table 10
Recovery confirmation cues. Triggered when pace returns to minimum after a corrective cue. Single cue per condition, not cycled.

Condition A — Autonomy-Supportive	Condition B — Ego-Threatening
"Pace is back! Well done for pushing through, keep that rhythm going!"	"Pace recovered! Now don't let it drop again so easily!"

Table 11
Total audio file count.

Cue type	Count
Start and finish cues (Table 7)	4 (2 per condition)
Corrective threshold (Table 1)	16 (8 per condition)
Recovery confirmation (Table 10)	2 (1 per condition)
500-metre check-in, above pace (Table 8)	18 (9 per condition)
500-metre check-in, below pace (Table 9)	18 (9 per condition)
Total	**58 audio files**

B. Baseline Questionnaire

This appendix presents the complete baseline questionnaire administered to participants prior to the first run. All items were presented in Dutch via Qualtrics. English translations are provided in parentheses for reference.

Informed Consent

Q1. Ik ga akkoord met deelname.
(*I agree to participate.*)
Response options: *Ja* (Yes) / *Nee* (No)

Demographics

Q2. Leeftijd
(*Age*) — open text field

Q3. Geslacht
(*Gender*)
Response options: *Man* (Male) / *Vrouw* (Female) / *Anders* (Other) / *Zeg ik liever niet* (Prefer not to say)

Q4. Hoe vaak loop je per week?
(*How often do you run per week?*)
Response options: *1 keer* (Once) / *2–3 keer* (2–3 times) / *4+ keer* (4+ times)

Q5. Heb je ervaring met lopen op een doeltempo?
(*Do you have experience running at a target pace?*)
Response options: *Ja* (Yes) / *Nee* (No)

General Self-Efficacy Scale

De volgende vragen gaan niet specifiek over hardlopen, maar over hoe je in het algemeen in het leven staat en met uitdagingen omgaat. Kies het antwoord dat het beste bij jou als persoon past.

(*The following questions are not specifically about running, but about how you approach life in general and deal with challenges. Choose the answer that best fits you as a person.*)

Rated on a 4-point scale: *Helemaal niet waar* (1 = Not at all true) — *Nauwelijks waar* (2 = Hardly true) — *Enigszins waar* (3 = Somewhat true) — *Helemaal waar* (4 = Exactly true). Source: Schwarzer & Jerusalem (1995).

1. Ik slaag er altijd in om moeilijke problemen op te lossen als ik maar genoeg mijn best doe.
(*I can always manage to solve difficult problems if I try hard enough.*)

2. Als iemand mij tegenwerkt, vind ik manieren en wegen om te krijgen wat ik wil.
(*If someone opposes me, I can find the means and ways to get what I want.*)

3. Het valt me makkelijk om aan mijn voornemens vast te houden en mijn doelen te bereiken.
   (*It is easy for me to stick to my aims and accomplish my goals.*)
4. Ik heb er vertrouwen in dat ik efficiënt kan omgaan met onverwachte gebeurtenissen.
   (*I am confident that I could deal efficiently with unexpected events.*)
5. Dankzij mijn vindingrijkheid weet ik hoe ik onvoorziene situaties moet aanpakken.
   (*Thanks to my resourcefulness, I know how to handle unforeseen situations.*)
6. Ik kan de meeste problemen oplossen als ik er de benodigde moeite in steek.
   (*I can solve most problems if I invest the necessary effort.*)
7. Ik kan kalm blijven als ik voor moeilijkheden sta, omdat ik kan vertrouwen op mijn vermogen om hiermee om te gaan.
   (*I can remain calm when facing difficulties because I can rely on my coping abilities.*)
8. Wanneer ik met een probleem word geconfronteerd, kan ik meestal meerdere oplossingen vinden.
   (*When I am confronted with a problem, I can usually find several solutions.*)
9. Als ik in de problemen zit, kan ik meestal wel een oplossing bedenken.
   (*If I am in trouble, I can usually think of a solution.*)
10. Ik kan meestal alles aan wat er op mijn pad komt.
    (*I can usually handle whatever comes my way.*)

C. Post-Run Questionnaire

This appendix presents the complete post-run questionnaire administered to participants after each 5-kilometre run. All items were presented in Dutch via Qualtrics. English translations are provided in parentheses for reference. Items were administered in the order shown below.

Rating of Perceived Exertion

Hoe zwaar was de lichamelijke inspanning tijdens deze run?
(*How physically demanding was the exertion during this run?*)

Measured using the Borg CR10 scale (Borg, 1982):

Score	Label (Dutch / English)
0	Rust / Rest (no exertion at all)
1	Zeer licht / Very light
2	Licht / Light
3	Matig / Moderate
4	Vrij zwaar / Somewhat hard
5	Zwaar / Hard
6	—
7	Zeer zwaar / Very hard
8	—
9	—
10	Maximaal / Maximal (the hardest exertion you have ever experienced)

Manipulation Check

Ik ervaarde de feedback als ...
(*I experienced the feedback as ...*)

Rated on a 5-point Likert scale (1 = *Helemaal mee oneens* / Strongly disagree; 5 = *Helemaal mee eens* / Strongly agree):

1. Ondersteunend (*Supportive*)
2. Confronterend (*Confrontational*)

Perceived Cognitive Processing (H2)

All items rated on a 5-point Likert scale (1 = *Helemaal mee oneens* / Strongly disagree; 5 = *Helemaal mee eens* / Strongly agree):

1. Ik moest bewust nadenken over de betekenis van de feedback.
   (*I had to consciously think about the meaning of the feedback.*)

2. Mijn reactie op de feedback voelde automatisch aan.[R]
   (*My reaction to the feedback felt automatic.*)
3. De feedback onderbrak mijn concentratie tijdens het lopen.
   (*The feedback interrupted my concentration while running.*)

[R] Reverse-coded item.

Motivational Reactance

All items rated on a 5-point Likert scale (1 = *Helemaal mee oneens* / Strongly disagree; 5 = *Helemaal mee eens* / Strongly agree):

1. De feedback motiveerde mij om te bewijzen dat ik het aankon.
   (*The feedback motivated me to prove I could handle it.*)
2. De feedback motiveerde mij om onmiddellijk harder te pushen.
   (*The feedback motivated me to push harder immediately.*)

Perceived Psychological Stress (H3)

Hoe stressvol vond je de feedback die je tijdens het hardlopen kreeg?
(*How stressful did you find the feedback you received during the run?*)

Rated on a 5-point verbal scale (Elo, Leppänen, & Jahkola, 2003):

Score	Label (Dutch / English)
1	Helemaal niet / Not at all
2	Een beetje / A little
3	Enigszins / Somewhat
4	Erg / Very
5	Heel erg / Extremely

Open-Ended Questions

The following two questions were administered verbally by the researcher immediately after questionnaire completion, and responses were recorded in written notes:

1. Beschrijf wat er door je heen ging toen het corrigerende geluidsfragment afspeelde, en wat je daarna deed.
   (*Describe what went through your mind when the corrective cue played, and what you did next.*)
2. Was er iets aan de coaching dat je stoorde of juist hielp? Wat precies?
   (*Was there anything about the coaching that bothered or helped you? What exactly?*)