Journal of Human Kinetics volume xxx/xxxx, x–x DOI: 10.5114/jhk/219655

Section IV – Psychological and Sociological Aspects of Sport and Exercise

Are Predictive Metrics by Wearables “Unproductive”? The Effect of Deceptive Wearable Monitor Feedback on Performance and Perception in Recreational Runners

by

Andrew J. Carnes [1,], Sara E. Mahoney [1]*

Popular wearables used by endurance athletes feature predictive prompts about recovery and fitness based on heart rate variability (HRV) and/or exercise data. No previous research has examined how receiving these prompts affects the athlete. The aim of this study was to determine the effect of deceptive feedback from a wearable on time trial performance, pacing, perceived effort, perceived readiness and performance quality, and affect in recreational runners. Participants ($N = 19$) completed three 5-km time trials (TT), each preceded by a sham “analysis” by a contrived prototype wearable. Participants received no feedback (FAM) or scripted feedback indicating a positive (POS) or negative (NEG) recovery and readiness status. Perceived readiness and the performance level were assessed before and after each TT. TT performance was similar among conditions ($p = 0.54$). The rating of perceived exertion (RPE) at the 2nd km was lower under the POS (13.8 ± 1.54) than the NEG (14.5 ± 1.31, $p = 0.01$) and FAM (14.6 ± 1.26, $p = 0.02$) conditions. The 4th km was faster under the POS (337 ± 64.7 s) compared to the NEG (342 ± 62.0 s, $p = 0.04$) and FAM (346 ± 65.5 s, $p < 0.01$) conditions. Participants reported higher perceived readiness ($p = 0.02$) under the POS (76.4 ± 14.9 mm) than the NEG condition (66.7 ± 17.7 mm) and higher ratings of performance under the POS (85.6 ± 11.8 mm) than the NEG (73.7 ± 18.2 mm, $p < 0.01$) and FAM (76.1 ± 17.6 mm, $p < 0.01$) conditions. Predictive feedback from a wearable device does not appear to affect running performance, but may affect expectations, perceived effort, pacing, and the perception of the completed exercise. Algorithmically derived scores or ratings by wearables should be interpreted with caution or along with analysis of individual trends in HRV.

Keywords: *running; wearable; heart rate variability; deception; predictive metrics*

Introduction

The use of wearable fitness monitors has proliferated in the last decade (i.e., 2014–2024) (Peake et al., 2018; Shei et al., 2022). Wearable technology was named the top fitness trend in seven out of the nine years between 2016 and 2024 (A'Naja et al., 2024), during which time the wearable technology market expanded dramatically (Halson et al., 2016). Wearables are especially prevalent among distance runners, as >90% report using a GPS tracking device (Moore and Willy, 2019). The measurement of duration, distance, and speed, combined with HR monitoring, enables the modification of training loads and intensity distribution (Casado et al., 2022; Paquette et al., 2020) to target specific event demands and physiologic capacities. Beyond these properties, current running/fitness wearables feature a variety of predictive metrics derived from resting and/or submaximal exercise data. On-screen prompts can inform users of predicted VO_{2max}, the “training effect” or the “strain score” of a session, and recommended recovery duration or a “recovery score”. Several of these metrics rely on monitoring heart rate variability (HRV), which has been extensively researched as an indicator of cardiovascular autonomic balance (Plews et al., 2013). In active populations, increased HRV is generally thought to indicate greater

[1] *Department of Exercise Science, Bellarmine University, Louisville, Kentucky, USA.*
* Correspondence: acarnes@bellarmine.edu

parasympathetic modulation associated with enhanced fitness, positive adaptation to training, and adequate recovery. By contrast, decreased HRV may suggest decreased fitness and performance secondary to training induced fatigue or overreaching (Buchheit et al., 2010; Plews et al., 2013). While the availability of real time data on external work output and internal physiologic responses can help objectively track progress and individualize exercise or sport specific training recommendations, considerable debate (Halson et al., 2016; Shei et al., 2022) surrounds the real-world applicability of predictive data in sport or exercise.

Contrasting a trial-and-error approach to training modification, HRV monitoring could hypothetically detect positive adaptation, excessive training loads, or inadequate recovery, helping to improve fitness while avoiding overtraining or injury (Bosquet et al., 2008). Indeed, numerous studies have shown decreased HRV in runners following increases in training intensity (Pichot et al., 2002) or volume (Manzi et al., 2009) and increased HRV through a training program which enhanced 10-km performance (Buchheit et al., 2010). Furthermore, scheduling high intensity training sessions based on HRV has led to greater performance improvement than following a predefined schedule (Figueiredo et al., 2022; Nuuttila et al., 2017; Vesterinen et al., 2016). However, other studies show conflicting results of increased HRV during periods of overload/overreaching associated with decreased performance (Bellenger et al., 2016, 2017; Le Meur, 2013). Whereas the collective research on HRV monitoring in athletes indicates variation in the HRV response to training, fatigue, and performance changes, the HRV-derived metrics of current fitness wearables have drawn widespread attention throughout the athletic and research communities. However, none have been validated (Peake et al., 2018) and the proprietary algorithms used to generate these metrics can vary considerably between various manufacturers (Shei et al., 2022). As such, caution may be warranted around the application of wearables' predictive metrics to guide real world training and recovery decisions (Halson et al., 2016; Shei et al., 2022).

While the nuance surrounding the application of HRV and derived metrics in endurance athletes has been recognized (Plews et al., 2013), we are not aware of research which has examined the potential physiologic, perceptual, or psychological effects of receiving prompts on "training effect", "readiness" or the "recovery score" from a wearable, independent of their actual validity. However, ample research suggests that placebo or nocebo effects (i.e., the expectation of benefit or harm, respectively) can significantly impact exercise performance (Hurst et al., 2020). Previous studies have used deception about supplement intake, hydration status, speed, and power output, but none have examined the effects of wearable monitor feedback on positive or negative expectations and the resulting performance. It is therefore warranted to examine whether placebo or nocebo effects may arise from predictive prompts delivered by widely used fitness wearables (e.g., "unproductive", "fatigued", "exhausted", "fresh", "optimized").

Given the proliferation of predictive metrics from wearable monitors used expansively by runners and other athletes (Miqueleiz et al., 2025), it is important to determine if the devices' notifications themselves can induce positive or negative expectations which could affect training or competition. Elucidating such an effect may improve the application of wearable data and metrics by endurance athletes and coaches. The aim of the present study was to determine the effect of deceptive feedback from an artificial wearable on time trial performance, pacing, perceived effort, perceptions of readiness and performance quality, and affect in recreational runners. We hypothesized that performance and pacing would be faster, perceived effort would be lower, and perceptions of readiness and performance would be greater following deceptive feedback indicating a positive (versus a negative) recovery and physiologic readiness status.

Methods

Participants

Participants (N = 19, n = 14 female, n = 5 male) were recruited from local running groups, events, and stores via flyers, word-of-mouth, and approved social media posts. Participants were required to be habitual runners (≥ 3 d·wk^{-1} for 6 months) between the ages of 18–65 who competed in ≥2 organized races annually. Exclusion criteria included running-related injury within 3 months or any medical complications that would contraindicate physical activity. Age, body mass, and VO_{2max} of included participants were 40.7 ± 7.67 yr, 62.9 ± 10.5 kg, and 41.6 ± 7.88 ml·kg^{-1}·min^{-1} for women and 50.8 ± 9.78 yr, 79.1 ± 3.21 kg, and 46.9 ± 6.48 ml·kg^{-1}·min^{-1} for men, respectively. All study procedures were approved by the Institutional Review Board of the Bellarmine University, Louisville, Kentucky, United States of America (IRB #1028; approval date: 03 November 2022) and all participants provided written informed consent.

Design and Procedures

The study used a within-subjects repeated measures design. Participants visited the laboratory four times with 4–7 days between visits. The first visit consisted of baseline measurements of anthropometrics and aerobic fitness (i.e., VO_{2max}). The subsequent three visits involved a maximal effort 5-km time trial (TT) on a treadmill (Woodway 4Front, WI). In visit 2, the TT was used to familiarize (FAM) participants and mitigate a learning effect. In visits 3 and 4, participants experienced positive (POS) or negative (NEG) feedback from their wearable device prior to their TT. The order of the POS and NEG conditions was randomized. TT performance, ratings of perceived exertion (RPEs), pacing, subjective ratings of readiness and performance level, perceived stress, and affectual responses were compared between feedback conditions.

Experimental Conditions

Prior to exercise, during visits 2–4 (FAM, POS, NEG), participants sat quietly for two minutes while wearing a nondescript wristwatch resembling a popular wearable, with the screen covered by opaque tape (Xia et al., 2025). Participants were informed that the wristwatch was a prototype monitor with a newly engineered algorithm to monitor recovery status and physiologic "readiness" for optimal performance in a short pre-exercise window (Düking et al., 2021). This purported capability differed from most available wearables, which relied on data collected continuously over days or weeks. The wristwatch (OYV Smart Watch, Denmark), shown in Figure 1, was equipped with an illuminated and functioning optical sensor that was visible to participants, but was used as an inert prop that did not collect any actual data. A researcher applied the wristwatch snugly to the participant's left wrist to prevent any visibility of manufacturer branding on the rear of the device. Following the false analysis, the principal investigator plugged the wristwatch into a computer with the monitor obscured from the participant. After one minute of artificially studying the computer, a scripted feedback statement was delivered verbally. The statement differed between the two conditions, suggesting a positive recovery and readiness status under the POS, but a poor status under the NEG condition. Statements for each condition are presented in Table 1.

Visit 1: Baseline Fitness Testing

Participants received the same instructions to follow prior to all four visits and were instructed to follow their typical training schedule throughout the study period. They were asked to perform only light to moderate intensity training for two days preceding each visit, and on the day before each visit to rest or train <60 min and abstain from resistance exercise and alcohol. Each participant completed all visits at a consistent time of day, with no more than 1 h of difference between any appointment times. Upon their first arrival to the laboratory, participants had their body height and mass measured using a calibrated scale and stadiometer (Seca, Chino, CA), respectively. They then conducted a 10–15-min warm-up at a self-selected pace on a treadmill. The duration and the pace of the warm-up were recorded and replicated for each subsequent visit. Following the warm-up, participants completed an incremental test to exhaustion. Expired gases were analyzed using a calibrated metabolic analyzer (ParvoMedics TrueOne 2400, Sandy, UT). The test protocol began at 8 km·hr^{-1} and a 1% incline, and the speed was increased 0.5 km·hr^{-1} each minute until volitional exhaustion. RPE was assessed every two minutes using the 6–20 Borg Scale (Borg, 1998). VO_{2max} was reported as the highest average VO_2 over 30 s and confirmed by standardized criteria.

After fitness testing, participants completed a short survey on their use of wearable tracking devices and/or online fitness sharing portals. Those indicating wearable use were asked to abstain from measuring HRV for the duration of the study and to temporarily stop using any standalone devices that specifically measured HRV to inform indications of recovery status or training effectiveness (e.g., Whoop©, Oura Ring©).

Participants were free to continue using GPS devices to track speed, distance, and/or the heart rate during exercise. However, the principal investigator assisted and verified that HRV tracking features and indications of training/recovery status, including any associated recommendations, were disabled or hidden. If fitness sharing portals were used, it was further requested that any posts of exercise activities excluded comments regarding the study.

Visits 2–4: Familiarization and Experimental Time Trials

Upon arrival, researchers verbally verified the participant's abstinence from HRV-derived metrics. Participants then underwent the 2-min sham "analysis" while wearing the contrived prototype monitor. Prior to the FAM condition, the analysis was described as a baseline measurement and no follow-up feedback was given. In visits 3 and 4, the statement corresponding to POS or NEG was delivered after the 2-min sham analysis. Following delivery of feedback, participants were asked to indicate their perceived readiness for the TT using a visual analog scale (VAS), and characterize their current perception of cumulative stress using the Daily Analysis of Life Demands for Athletes (DALDA). Exercise then began with the same warm-up completed prior to baseline fitness testing, followed by a maximal 5-km TT at a 1% incline. Participants were instructed to complete the TT as quickly as possible and were free to adjust the speed at any time. The distance covered was visible, but speed and elapsed time were obscured. The researchers did not communicate with the participant during the TT, except for assessing RPE. The total elapsed time (s) to complete each TT was recorded as the primary outcome measure. Intermediate times (s) and RPE were assessed at each kilometer. After a 5-min walking cool down, participants were asked to indicate their perceived performance level via VAS, before completing the Exercise Induced Feeling Inventory (EFI).

At the conclusion of the fourth visit, each participant was debriefed of the true nature of the study and granted access to the performance times and intermediate splits for each TT.

Subjective Measures

Rating of Perceived Exertion (RPE). Undifferentiated, whole body RPE was assessed using the 15-point, 6–20 Borg scale (Borg, 1998). The scale was displayed in front of participants on a large poster, and participants were free to speak or point to the number corresponding to their effort at that time.

Perceived Readiness (VAS_{pre}). Participants marked a 100 mm VAS anchored by the statements "I am not ready at all to perform my best" at 0 mm and "I am as ready as I can possibly be to perform my best" at 100 mm (McCormack et al., 1988).

Perceived Stress. Participants responded to the Daily Analysis of Life Demands for Athletes (DALDA). The DALDA is a validated self-report instrument for stress tolerance in athletic populations (Rushall, 1990). Thirty-four questions are separated into two sections. Section A consists of 9 questions pertaining to sources of life stress, while Section B includes 25 questions on symptoms of stressors. Each question is answered by selecting a (worse than normal), b (normal), or c (better than normal). The number of a, b, and c responses was tallied from each section.

Perceived Performance Level (VAS_{post}). Participants marked a 100 mm VAS anchored by the statements "that was my worst possible performance" at 0 mm and "that was my best possible performance" at 100 mm (Bath et al., 2012; McCormack et al., 1988).

Post Exercise Affect. Participants responded to the Exercise Induced Feeling Inventory (EFI), a validated 12-item questionnaire to assess four specific feeling states following physical activity: revitalization, tranquility, positive engagement, and physical exhaustion (Gauvin and Rejeski, 1993). Participants rated 12 specific feelings pertaining to these feeling states (3 questions for each feeling state) on a five-point Likert scale ranging from 0 (do not feel) to 4 (feel very strongly).

Statistical Analysis

All data are presented as mean ± SD. Statistical analyses were performed using SPSS (version 28, IBM SPSS Inc., Illinois, USA) with an α level = 0.05. Parametric data were analyzed using univariate analyses of variance (ANOVAs) to detect differences in each outcome variable between the three feedback conditions. Paired samples *t*-tests were then used to elucidate any main effect of condition detected by ANOVA. Nonparametric data were analyzed using the Friedman test with post hoc Wilcoxon signed rank tests. Effect sizes for any significant differences between feedback conditions were quantified using Cohen's *d*, with values of 0.2, 0.5, and 0.8 indicating small, medium, and large effect sizes, respectively.

Results

Performance, Perceived Exertion, and Pacing

Univariate ANOVA showed that 5-km TT performance did not significantly differ between conditions (F = 0.364, p = 0.54). A main effect of condition occurred for RPE at the 2-km split (F = 4.23, p = 0.02) and the time to complete the 4th km (F = 3.95, p = 0.03). RPE at the 2-km split was significantly lower under the POS (13.8 ± 1.54) than NEG (14.5 ± 1.31, p = 0.01, d = 0.53) and FAM (14.6 ± 1.26, p = 0.02, d = 0.50) conditions. The 4th km was faster under the POS (337 ± 64.7 s) than NEG (342 ± 62.0 s, p = 0.04, d = 0.42) and FAM (346 ± 65.5 s, p < 0.01, d = 0.61) conditions. TT completion times, average and intermediate RPE values, and elapsed time for each kilometer are shown in Table 2. RPE and pacing profiles are shown in Figures 2 and 3, respectively.

Expectations and Subjective Ratings of Performance

A main effect of condition occurred for perceived readiness (F = 3.14, p = 0.05) and the perceived performance level (F = 4.22, p = 0.02). Prior to the TT, participants reported higher (p = 0.02, d = 0.46) VAS_{pre} under the POS (76.4 ± 14.9 mm) than the NEG (66.7 ± 17.7 mm) condition. Following the TT, VAS_{post} was greater under the POS (85.6 ± 11.8 mm) than NEG (73.7 ± 18.2 mm, p < 0.01, d = 0.70) and FAM (76.1 ± 17.6 mm, p < 0.01, d = 0.67) conditions. Pre and post VAS ratings are shown for each condition in Figure 4.

Perceived Stress Tolerance and Affect

The number of a, b, and c responses for each section of the DALDA did not differ among conditions ($p \geq 0.12$). A significant difference in the positive engagement dimension of the EFI occurred between conditions (χ^2 = 8.82, p = 0.01). Participants reported higher positive engagement (p = 0.02) under the POS (10.2 ± 1.35) than the NEG (9.05 ± 1.98) condition, but no difference was observed in the other three dimensions of the EFI ($p \geq 0.06$).

Discussion

The present study adds a novel context to a rapid proliferation of HRV monitoring in athletes and HRV-derived metrics by wearable monitors. No significant effect on 5-km TT performance occurred following deceptive feedback on recovery and physiologic "readiness" for maximal performance. However, runners reported higher perceptions of readiness after receiving positive feedback, and likewise perceived better performance afterwards than when provided negative or no feedback. Additionally, RPE was lower at 2-km and the 4th km was faster under the POS relative to NEG and FAM conditions. Positive feedback also led to higher ratings of "positive engagement" following the TT. Prior to this study, the question of whether the wearables' messages themselves could impact exercise performance or related variables had not been addressed. However, receiving unfavorable prompts about fitness or recovery is a common source of frustration and confusion among runners. These findings provide preliminary answers to this issue and highlight the need for additional research on the balance between the utility of HRV-derived metrics and the athlete's interpretation and interaction with those metrics.

Although deceptive feedback did not consistently affect TT performance, the observed effects on intermediate RPE, pacing, performance expectations, perceived performance quality, and post-exercise affect provide the first empirical evidence that algorithmic prompts delivered by wearables have the capacity to exert an effect on high intensity exercise, irrespective of their accuracy or validity. Put simply, the present data suggest that generalized feedback from a wearable can affect how an athlete expects to feel and actually feels during and after training or competition. RPE is inextricably linked to the development of fatigue (Marcora and Staiano, 2010) and the decision to maintain, increase, or decrease power output or running speed (Foster et al., 2023). Ample research across a range of modalities and distances shows that RPE increases linearly relative to the percentage of a maximal exercise task completed, and reaches its maximum at the point of exhaustion or task completion (Foster et al., 2023). Endurance athletes may rely on RPE to gauge the tolerable level of homeostatic disturbance and thereby maintain the highest speed possible over a set distance while avoiding premature exhaustion (Azevedo et al., 2021). In the present study, RPE at the 2-km split was lower under the POS compared to NEG and FAM conditions. Such a delay in the rise of RPE could potentially prolong reaching a maximal RPE, allowing a longer maintenance of the desired pace. Conversely, a higher intermediate RPE resulting from unfavorable wearable feedback could augment the RPE slope and lead to premature fatigue. Although the altered 2-km RPE did not consistently alter the time to completion in the present 5-km TTs, it is undetermined if this could occur in a real-world competitive scenario or a longer event, in which central and/or psychological factors likely have a heavier contribution to fatigue (Thomas et al., 2016).

Table 1. Experimental Condition Statements.

Experimental Condition 1 (POS):
The cardiovascular data indicate that in the past week, your training and recovery were balanced in a very productive way. Based on these measurements, your aerobic fitness is improving and your recovery is sufficient to deliver an excellent performance.
Experimental Condition 2 (NEG):
The cardiovascular data indicate that in the past week, there may have been a sub optimal balance between your training and recovery. It is possible that your training load was excessive, or recovery was not adequate. However, it may still be possible to deliver a good performance.

Table 2. Outcome measure data.

	POS	NEG	FAM
TT (s)	1722 ± 333.5	1732 ± 324.9	1741 ± 338.7
1-km split (s)	374 ± 74.1	371 ± 68.6	371 ± 74.0
2-km split (s)	348 ± 68.0	349 ± 67.4	349 ± 69.0
3-km split (s)	342 ± 65.1	343 ± 65.5	346 ± 66.7
4-km split (s)	337 ± 64.6*	341 ± 62.0	346 ± 65.5
5-km split (s)	323 ± 60.0	328 ± 57.9	329 ± 61.0
1-km RPE	12.3 ± 1.82	12.6 ± 1.61	13.0 ± 1.27
2-km RPE	13.8 ± 1.54*	14.5 ± 1.31	14.6 ± 1.26
3-km RPE	15.7 ± 1.11	15.9 ± 1.33	15.8 ± 1.07
4-km RPE	17.2 ± 1.01	16.8 ± 1.32	16.9 ± 1.39
5-km RPE	19.1 ± 0.76	19.2 ± 1.20	18.9 ± 0.99
VAS Pre (mm)	76.4 ± 14.9#	76.2 ± 14.3	66.7 ± 17.7
VAS Post (mm)	85.6 ± 11.8*	73.7 ± 18.2	76.0 ± 17.6

** significantly different from NEG and FAM ($p \leq 0.04$); # significantly different from NEG ($p \leq 0.02$)*

Figure 1. Deceptive wearable device.

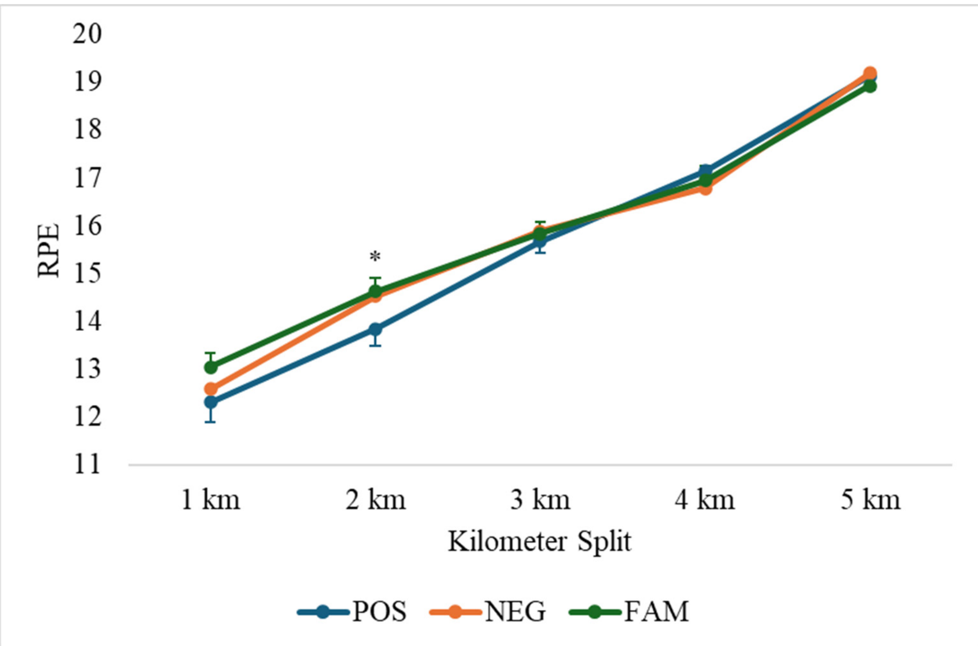

Figure 2. RPE by kilometer.

*RPE at the 2-km split was significantly lower under the POS (13.8 ± 1.54) than NEG (14.5 ± 1.31, p = 0.01) and FAM (14.6 ± 1.26, p = 0.02) conditions; * significantly different from NEG and FAM conditions (p ≤ 0.02)*

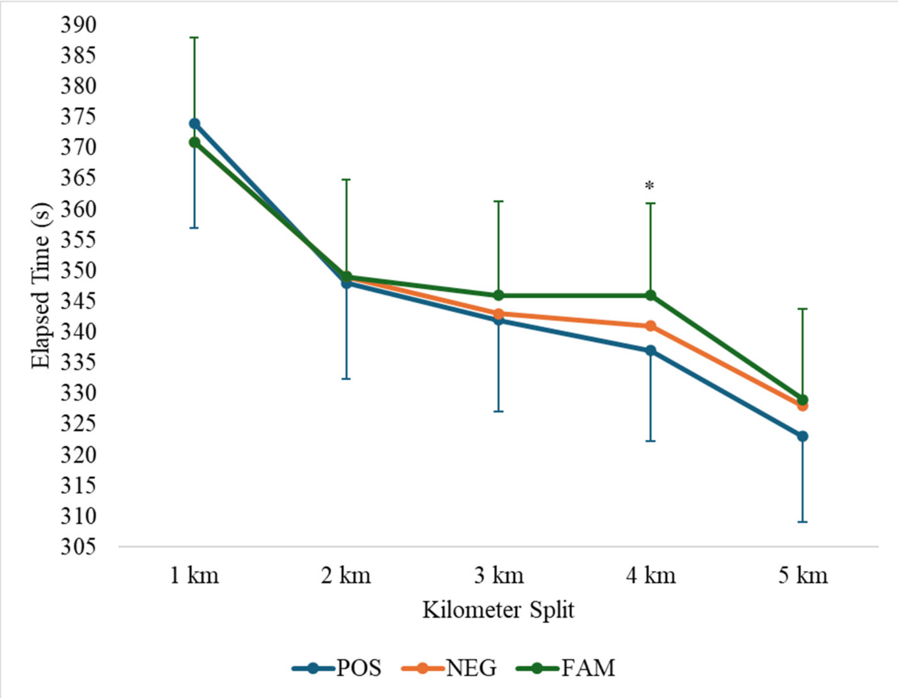

Figure 3. Kilometer pacing.
*The 4th km was faster under the POS (337 ± 64.7 s) than the NEG (342 ± 62.0 s, p = 0.04) and FAM (346 ± 65.5 s, p < 0.01) conditions; * significantly different from NEG and FAM conditions (p ≤ 0.04)*

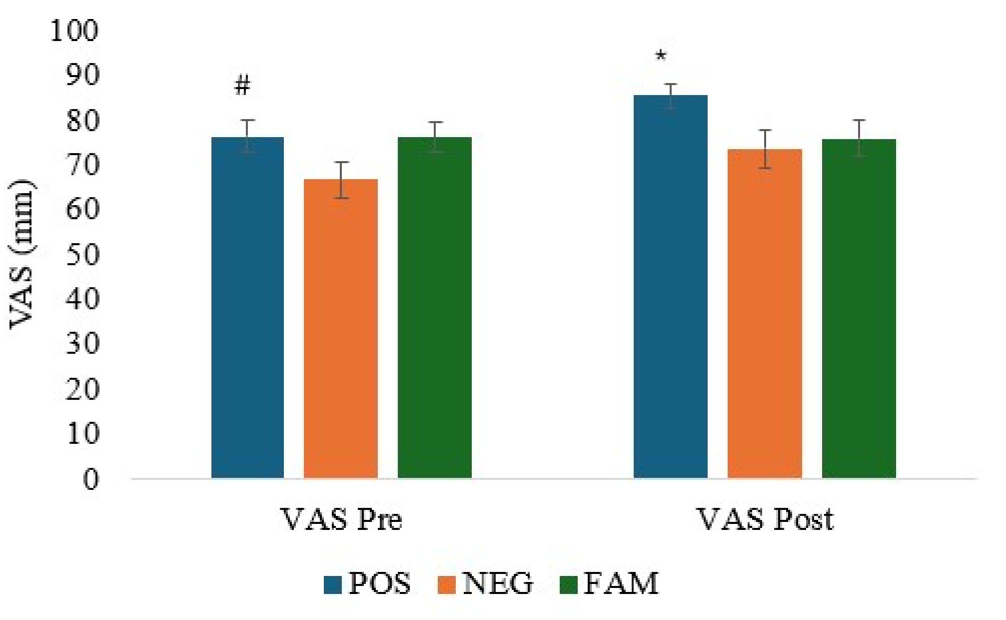

Figure 4. Perceived readiness and performance level.
*VAS_{pre} was greater ($p = 0.02$) under the POS (76.4 ± 14.9 mm) than the NEG (66.7 ± 17.7 mm) condition; VAS_{post} was greater under the POS (85.6 ± 11.8 mm) than NEG (73.7 ± 18.2 mm, $p < 0.01$) and FAM (76.1 ± 17.6 mm, $p < 0.01$, $d = 0.67$) conditions; * significantly different from NEG and FAM conditions ($p < 0.01$); # significantly different from the NEG condition ($p = 0.02$)*

Furthermore, whereas overall TT performance was not affected, the lower RPE at 2 km was followed by a faster 4th km under the POS condition, aligning with research pointing to RPE as the key determinant in pacing decisions (Foster et al., 2023). In essence, the ability of wearable feedback to alter RPE could potentially affect the athlete's perception of what he/she "has left" and consequently impact his/her pacing profile, either to benefit or harm. After the TT, altered perceptions of performance quality and "positive engagement" following the deceptive feedback aligned with the widespread anecdotal dismay of endurance athletes when receiving discouraging ratings of an exercise session from their wearable (e.g., "unproductive"). As suggested by research on "psychological momentum" (Briki et al., 2013), these perceptions and affectual responses may impact future exercise sessions. While the current data are not conclusive, they constitute the first available evidence that predictive prompts from wearable devices can affect the athlete.

Because of the novelty of the research question, opportunities to compare the current findings to those of previous research are limited. However, parallels do exist with previous research on other types of deceptive information inducing placebo or nocebo effects in sport. Prior studies have shown that deceptive "placebo" information conferred during high intensity exercise can enhance performance and/or the perception of effort (Hurst et al., 2020). Cyclists (Stone et al., 2012) and triathletes (Taylor and Smith 2014) enhanced performance when riding against a deceptive avatar or receiving incorrect pacing information, and running economy has increased following deceptive positive feedback on movement efficiency (Stoate et al., 2012). Fewer studies have considered the "nocebo" (i.e., deleterious) effect on exercise performance in trained individuals (Hurst et al., 2020), but a compelling study (Funnell et al., 2024) revealed a 6% decrement in cycling performance when participants were told that they were dehydrated by 2% body mass, despite actual hydration status being controlled by an intra-gastric tube. While the informational deception deployed here did not influence overall performance as in those studies, the significant effects on several perceptual variables contribute additional evidence that information, such as that displayed in a wearable's messaging, can exert meaningful effects whether true or not.

The current findings warrant caution in the use and interpretation of algorithmically derived scores or ratings by a wearable device. However, they do not refute or discredit the utility of individualized HRV analysis to guide endurance training periodization. In fact, growing research suggests that endurance training prescription guided by HRV is likely to yield greater performance enhancement than predefined scheduling of key sessions (Figueiredo et al., 2022; Nuuttila et al., 2017; Vesterinen et al., 2016). Contrasting a predetermined schedule, HRV-guided periodization is suggested to promote a greater training response and avoid nonfunctional overreaching (Düking et al., 2021; Medellin Ruiz et al., 2020). However, the predictive prompts generated by many wearables' default "out-of-the-box" settings do not equate to the systematic analysis of individual HRV trends utilized in these studies. In those studies demonstrating greater performance improvements over 3-km to 5-km, training decisions were based on whether the 7-day rolling average of HRV fell within a "smallest worthwhile change" calculated for each participant based on daily HRV measurements during several weeks preceding the intervention (Figueiredo et al., 2022; Nuuttila et al., 2017; Vesterinen et al., 2016). Although software applications have adopted individualized HRV analysis (Carrasco-Poyatos et al., 2022), no proprietary "score" or on-board predictive metric offered by a wearable device has been validated to gauge internal recovery status or as a basis for the arrangement of training sessions. Runners, coaches, and practitioners are thus cautioned that algorithmic scores or subjective indications (e.g., "unproductive", "detraining") do not equate to monitoring HRV under consistent daily conditions and scheduling high intensity sessions according to deviations from an individually defined range.

While the findings of this study are novel, specific limitations should be considered in their interpretation and application. Our sample was predominantly female, which did not allow for between-sex comparisons (Seo et al., 2024) of the impact of POS versus NEG feedback. The scripted feedback on recovery and readiness status following the false wearable analysis was delivered verbally to participants, which differs from the digital prompts commonly received by users of these devices. However, the development of a device capable of providing externally manipulated feedback, as opposed to algorithmic analysis of user data, was outside the scope of this study. Furthermore, past research implementing other forms of a placebo or a nocebo in sport, as in the case of hydration status or supplement intake (Hurst et al., 2020), used verbal delivery of deceptive information to influence physical performance or related variables. Another potential limitation was that the artificial wearable monitor used was unfamiliar to participants and may have differed from their own device. Anecdotally, however, when we asked participants about their belief in the feedback at the conclusion of their participation, only a single participant expressed suspicion about the validity of the false wearable's analysis. We also recognize that while all participants in the study reported using a wearable monitor while running, none reported using predictive metrics from the devices, based on HRV or otherwise, to plan or adjust their training or competition. As such, these results should not be assumed to apply to athletes who regularly monitor and/or use HRV or its derived predictive metrics to guide training decisions. Future research should seek to further elucidate the effects of wearables' predictive messaging by longitudinally investigating larger samples consisting of athletes who habitually monitor HRV, including those at higher levels of performance and training load. While deception is not a necessity of ongoing research, further investigation of how the predictive information that is standard on many current wearables can affect performance and perception is warranted.

Conclusions

In conclusion, these findings indicate that predictive feedback from a wearable monitor does not consistently affect high intensity endurance exercise performance, but may affect performance expectations, perceived effort, pacing, and the perception of the completed exercise. Based on these results, athletes and coaches are not discouraged from using wearable technology or monitoring HRV, but are advised to do so in conjunction with subjective measures of athlete recovery and well-being status, to collect subjective measures prior to evaluating wearable feedback, and to observe individual trends in HRV as opposed to accepting algorithmically derived ratings based on HRV.

Author Contributions: Conceptualization: A.J.C.; methodology: A.J.C. and S.E.M.; formal analysis: A.J.C.; investigation: A.J.C. and S.E.M.; data curation: A.J.C.; writing—original draft preparation: A.J.C.; writing—review & editing: S.E.M.; visualization: A.J.C.; supervision: A.J.C.; project administration: A.J.C. All authors have read and agreed to the published version of the manuscript.

ORCID iD:

Andrew J. Carnes: https://orcid.org/0000-0003-3715-4276

Sara E. Mahoney: https://orcid.org/0009-0001-8481-2887

Funding Information: This research received no external funding.

Institutional Review Board Statement: This study was conducted following the principles of the Declaration of Helsinki, and approved by the Institutional Review Board of the Bellarmine University, Louisville, Kentucky, United States of America (IRB #1028; approval date: 03 November 2022).

Informed Consent: Informed consent was obtained from all participants included in the study.

Conflicts of Interest: The authors declare no conflict of interest.

Acknowledgements: The authors wish to thank Lauren Marshall for assistance with data collection, Dr. Jacob Barkley for assistance with statistical analysis, and Lindsay Johnson for assistance with manuscript formatting. Most importantly, we wish to express our gratitude to the participants who enthusiastically took part in this study.

Received: 13 June 2025

Accepted: 18 March 2026

References

A'Naja, M. N., Reed, R., Sansone, J., Batrakoulis, A., McAvoy, C., & Parrott, M. W. (2024). 2024 ACSM worldwide fitness trends: future directions of the health and fitness industry. *ACSM's Health & Fitness Journal*, *28*(1), 14–26. doi:10.1249/FIT.0000000000000933

Azevedo, R. d. A., Silva-Cavalcante, M. D., Lima-Silva, A. E., & Bertuzzi, R. (2021). Fatigue development and perceived response during self-paced endurance exercise: state-of-the-art review. *European Journal of Applied Physiology*, *121*, 687–696. doi:10.1007/s00421-020-04549-5

Bath, D., Turner, L. A., Bosch, A. N., Tucker, R., Lambert, E. V., Thompson, K. G., & Gibson, A. S. C. (2012). The effect of a second runner on pacing strategy and RPE during a running time trial. *International Journal of Sports Physiology and Performance*, *7*(1), 26–32. doi: 10.1123/ijspp.7.1.26

Bellenger, C. R., Karavirta, L., Thomson, R. L., Robertson, E. Y., Davison, K., & Buckley, J. D. (2016). Contextualizing Parasympathetic Hyperactivity in Functionally Overreached Athletes With Perceptions of Training Tolerance. *International Journal of Sports Physiology and Performance*, *11*(5), 685–692. doi:10.1123/ijspp.2015-0495

Bellenger, C. R., Thomson, R. L., Robertson, E. Y., Davison, K., Nelson, M. J., Karavirta, L., & Buckley, J. D. (2017). The effect of functional overreaching on parameters of autonomic heart rate regulation. *European Journal of Applied Physiology*, *117*, 541–550. doi:10.1007/s00421-017-3549-5

Borg, G. (1998). *Borg's perceived exertion and pain scales*. Human Kinetics.

Bosquet, L., Merkari, S., Arvisais, D., & Aubert, A. E. (2008). Is heart rate a convenient tool to monitor over-reaching? A systematic review of the literature. *British Journal of Sports Medicine*, *42*(9), 709–714. doi:10.1136/bjsm.2007.042200

Briki, W., Den Hartigh, R. J., Markman, K. D., Micallef, J.-P., & Gernigon, C. (2013). How psychological momentum changes in athletes during a sport competition. *Psychology of Sport and Exercise*, *14*(3), 389–396. doi:10.1016/j.psychsport.2012.11.009

Buchheit, M., Chivot, A., Parouty, J., Mercier, D., Al Haddad, H., Laursen, P., & Ahmaidi, S. (2010). Monitoring endurance running performance using cardiac parasympathetic function. *European Journal of Applied Physiology*, *108*, 1153–1167. doi:10.1007/s00421-009-1317-x

Carrasco-Poyatos, M., González-Quílez, A., Altini, M., & Granero-Gallegos, A. (2022). Heart rate variability-guided training in professional runners: Effects on performance and vagal modulation. *Physiology & Behavior*, *244*, 113654. doi:10.1016/j.physbeh.2021.113654

Casado, A., González-Mohíno, F., González-Ravé, J. M., & Foster, C. (2022). Training Periodization, Methods, Intensity Distribution, and Volume in Highly Trained and Elite Distance Runners: A Systematic Review. *International Journal of Sports Physiology and Performance*, *17*(6), 820–833. doi:10.1123/ijspp.2021-0435

Düking, P., Zinner, C., Trabelsi, K., Reed, J. L., Holmberg, H.-C., Kunz, P., & Sperlich, B. (2021). Monitoring and adapting endurance training on the basis of heart rate variability monitored by wearable technologies: A systematic review with meta-analysis. *Journal of Science and Medicine in Sport*, *24*(11), 1180–1192. doi:10.1016/j.jsams.2021.04.012

Figueiredo, D. H., Figueiredo, D. H., Bellenger, C., & Machado, F. A. (2022). Individually guided training prescription by heart rate variability and self-reported measure of stress tolerance in recreational runners: Effects on endurance performance. *Journal of Sports Sciences*, *40*(24), 2732–2740. doi:10.1080/02640414.2023.2191082

Foster, C., De Koning, J. J., Hettinga, F. J., Barroso, R., Boullosa, D., Casado, A., Cortis, C., Fusco, A., Gregorich, H., & Jaime, S. (2023). Competition between desired competitive result, tolerable homeostatic disturbance, and psychophysiological interpretation determines pacing strategy. *International Journal of Sports Physiology and Performance*, *18*(4), 335–346. doi:10.1123/ijspp.2022-0171

Funnell, M. P., Moss, J., Brown, D. R., Mears, S. A., & James, L. J. (2024). Perceived dehydration impairs endurance cycling performance in the heat in active males. *Physiology & Behavior*, *276*, 114462. doi:10.1016/j.physbeh.2024.114462

Gauvin, L., & Rejeski, W. J. (1993). The exercise-induced feeling inventory: Development and initial validation. *Journal of Sport and Exercise Psychology*, *15*(4), 403–423. doi:10.1123/jsep.15.4.403

Halson, S., Peake, J., & Sullivan, J. (2016). Wearable technology for athletes: information overload and pseudoscience? *International Journal of Sports Physiology and Performance*, *11*(6), 705–706. doi:10.1123/IJSPP.2016-0486

Hurst, P., Schipof-Godart, L., Szabo, A., Raglin, J., Hettinga, F., Roelands, B., Lane, A., Foad, A., Coleman, D., & Beedie, C. (2020). The placebo and nocebo effect on sports performance: a systematic review. *European Journal of Sport Science*, *20*(3), 279–292. doi:10.1080/17461391.2019.1655098

Le Meur, Y., Pichon, A., Schaal, K., Schmitt, L., Louis, J., Gueneron, J., Vidal, P. P., & Hausswirth, C. (2013). Evidence of parasympathetic hyperactivity in functionally overreached athletes. *Medicine and Science in Sports and Exercise*, *45*(11), 2061–2071. doi:10.1249/MSS.0b013e3182980125

Manzi, V., Castagna, C., Padua, E., Lombardo, M., D'Ottavio, S., Massaro, M., Volterrani, M., & Iellamo, F. (2009). Dose-response relationship of autonomic nervous system responses to individualized training impulse in marathon runners. *American Journal of Physiology-Heart and Circulatory Physiology*, *296*(6), H1733–H1740. doi:10.1152/ajpheart.00054.2009

Marcora, S. M., & Staiano, W. (2010). The limit to exercise tolerance in humans: mind over muscle? *European Journal of Applied Physiology*, *109*(4), 763–770. doi:10.1007/s00421-010-1418-6

McCormack, H. M., de L. Horne, D. J., & Sheather, S. (1988). Clinical applications of visual analogue scales: a critical review. *Psychological Medicine*, *18*(4), 1007–1019. doi:10.1017/S0033291700009934

Medellin Ruiz, J. P., Rubio-Arias, J. Á., Clemente-Suarez, V. J., & Ramos-Campo, D. J. (2020). Effectiveness of training prescription guided by heart rate variability versus predefined training for physiological and aerobic performance improvements: a systematic review and meta-analysis. *Applied Sciences*, *10*(23), 8532. doi:10.3390/app10238532

Miqueleiz, U., Aguado-Jimenez, R., Lecumberri, P., Garcia-Tabar, I. & M. Gorostiaga, E. (2025). Reliability of inertial measurement unit-based spatiotemporal and kinetic variables in endurance runners during treadmill running. *Journal of Human Kinetics*, *97*, 65–76. https://doi.org/10.5114/jhk/195895

Moore, I. S., & Willy, R. W. (2019). Use of wearables: tracking and retraining in endurance runners. *Current Sports Medicine Reports*, *18*(12), 437–444. doi:10.1249/JSR.0000000000000667

Nuuttila, O.-P., Nikander, A., Polomoshnov, D., Laukkanen, J. A., & Häkkinen, K. (2017). Effects of HRV-guided vs. predetermined block training on performance, HRV and serum hormones. *International Journal of Sports Medicine*, *38*(12), 909–920. doi:10.1055/s-0043-115122

Paquette, M. R., Napier, C., Willy, R. W., & Stellingwerff, T. (2020). Moving beyond weekly "distance": optimizing quantification of training load in runners. *Journal of Orthopaedic & Sports Physical Therapy*, *50*(10), 564–569. doi:10.2519/jospt.2020.9533

Peake, J. M., Kerr, G., & Sullivan, J. P. (2018). A critical review of consumer wearables, mobile applications, and equipment for providing biofeedback, monitoring stress, and sleep in physically active populations. *Frontiers in Physiology*, *9*, 743. doi:10.3389/fphys.2018.00743

Pichot, V., Busso, T., Roche, F., Garet, M., Costes, F., Duverney, D., Lacour, J.-R., & Barthélémy, J.-C. (2002). Autonomic adaptations to intensive and overload training periods: a laboratory study. *Medicine & Science in Sports & Exercise*, *34*(10), 1660–1666. doi:10.1097/00005768-200210000-00019

Plews, D. J., Laursen, P. B., Kilding, A. E., & Buchheit, M. (2013). Evaluating training adaptation with heart-rate measures: a methodological comparison. *International Journal of Sports Physiology and Performance*, *8*(6), 688–691. doi:10.1007/s40279-013-0071-8

Rushall, B. S. (1990). A tool for measuring stress tolerance in elite athletes. *Journal of Applied Sport Psychology*, *2*(1), 51–66. doi:10.1080/10413209008406420

Seo, M., Park, T. & Jung, H. C. (2024). Sex Differences in Heart Rate Variability and Vascular Function Following High-Intensity Interval Training in Young Adults. *Journal of Human Kinetics*, *90*, 89–100. https://doi.org/10.5114/jhk/170964

Shei, R.-J., Holder, I. G., Oumsang, A. S., Paris, B. A., & Paris, H. L. (2022). Wearable activity trackers–advanced technology or advanced marketing? *European Journal of Applied Physiology, 122*(9), 1975–1990. doi:10.1007/s00421-022-04951-1

Stoate, I., Wulf, G., & Lewthwaite, R. (2012). Enhanced expectancies improve movement efficiency in runners. *Journal of Sports Sciences*, *30*(8), 815–823. doi:10.1080/02640414.2012.671533

Stone, M., Thomas, K., Wilkinson, M., Jones, A., St Clair Gibson, A., & Thompson, K. (2012). Effects of deception on exercise performance: implications for determinants of fatigue in humans. *Medicine & Science in Sports & Exercise*, *44*(3), 534–541. doi:10.1249/MSS.0b013e318232cf77

Taylor, D., & Smith, M. F. (2014). Effects of deceptive running speed on physiology, perceptual responses, and performance during sprint-distance triathlon. *Physiology & Behavior*, *133*, 45–52. doi:10.1016/j.physbeh.2014.05.002

Thomas, K., Elmeua, M., Howatson, G., & Goodall, S. (2016). Intensity-dependent contribution of neuromuscular fatigue after constant-load cycling. *Medicine and Science in Sports and Exercise*, *48*(9), 1751–1760. doi:10.1249/MSS.0000000000000950

Vesterinen, V., Nummela, A., Heikura, I., Laine, T., Hynynen, E., Botella, J., & Häkkinen, K. (2016). Individual endurance training prescription with heart rate variability. *Medicine and Science in Sports and Exercise*, *48*(7), 1347–1354. doi:10.1249/MSS.0000000000000910

Xia, Z., Sun, D., Qian, Y., Xu, Y., Zhu, C., Cen, X., Song, Y., Xiang, L., Jemni, M. & Gu, Y. (2025). Biomechanical Strategies for Minimizing Force Plate Targeting Effects during Running: Efficacy of Masked Force Plate Integration with Augmented Visual Feedback. *Journal of Human Kinetics*, Advance online publication. https://doi.org/10.5114/jhk/212382