

Do error predictions of physical exertion inform the level of **running pleasure**?

Stage 1 Registered Report

Damien Brevers¹, Guillaume Martinent², İrem Tuğçe Öz¹, Olivier Desmedt³, Bas de Geus⁴

¹ Louvain Experimental Psychopathology Research Group (LEP), Psychological Sciences Research Institute (IPSY), UCLouvain, Louvain-La-Neuve, Belgium

² Laboratory of Vulnerabilities and Innovation in Sport (EA 7428), University of Claude Bernard Lyon1 – University of Lyon, Lyon, France.

³ Institute of Psychology, University of Lausanne, Lausanne, Switzerland

⁴ Institute for the Analysis of Change in Contemporary and Historical Societies (IACCOS), Faculty of Movement and Rehabilitation Sciences, UCLouvain, Louvain-La-Neuve, Belgium

Corresponding author

Damien Brevers, PhD, Laboratory for Experimental Psychopathology (LEP), Psychological Sciences Research Institute, UCLouvain, Louvain-la-Neuve, Belgium, e-mail:

damien.brevers@uclouvain.be

ABSTRACT

Humans have the ability to mentally project themselves into future events (prospective thinking) to promote the implementation of health-oriented behaviors, such as the planning of daily sessions of physical exercise. Nevertheless, it is currently unclear whether and how prospective thinking can assist individuals in generating future predictions about their own bodily states, such as when anticipating the level of physical exertion to be experienced in a forthcoming session of physical exercise. Based on the literature on the rating of perceived exertion (RPE), we advance that assessing prospective thinking toward physical exertion (prospective RPE) should inform on the remembered level of pleasure that was experienced by an individual during physical exercise (retrospective pleasure). We aim to examine this research question by using ecological momentary assessment of physical exertion to be filled out before (anticipatory RPE) and after (retrospective RPE, retrospective pleasure) each running session of a start-to-run program. By capitalizing on the core dynamic of reward prediction errors, we hypothesize that running sessions that are experienced with a lesser level of physical exertion than anticipated (a positive RPE-based prediction error) should be associated with a higher level of retrospective pleasure during the session of physical exercise, and vice versa (higher score of retrospective RPE than prospective RPE; a negative RPE-based prediction error). The confirmation of this hypothesis will **demonstrate** that the use of prospective and retrospective RPE is beneficial for identifying sessions of physical exercise that lead to an increase (or decrease) in the experience of pleasure. This may ultimately impact future engagement and commitment to physical exercise.

1. INTRODUCTION

Prospective thinking refers to humans' ability to mentally simulate the future (for a review, see Schacter et al., 2017). It allows individuals to effectively prepare for upcoming events and facilitates the enactment of goal-directed actions and the planning of behaviors, including health behaviors (Brevers et al., 2023; D'Argembeau et al., 2010; Schacter et al., 2017). A core feature of prospective thinking is that it enables one to flexibly retrieve and recombine past information into mental simulations related to future events (D'Argembeau et al., 2010; Schacter et al., 2017). These memory-based processes have been extensively studied with experimental tasks that involve the extraction of information about locations, objects, and people, as well as more schematic and conceptual knowledge to envision general goals or events (Schacter et al., 2017). Humans can thus engage in different forms of prospection, including episodic future thinking (for example, by imagining themselves in a particular place at a specific time, bringing specific details to mind) and semantic future thinking (i.e., thinking about the future in a general, abstract manner; Demblon & D'Argembeau, 2014).

Nevertheless, it is currently unclear how prospective thinking unfolds while generating future predictions about one's own bodily states, such as when anticipating the intensity of perceived exertion (i.e., the subjective intensity of effort, strain, discomfort, and/or fatigue that is experienced during physical exercise; Hutchinson, 2020; Robertson and Noble, 1997;) of a forthcoming session of physical exercise. Indeed, the level of physical exertion is usually indexed while exercising (i.e., momentary ratings of perceived exertion, RPE; e.g., "What intensity of exertion do you feel now?") or directly after the exercise session (i.e., retrospective RPE; e.g., "What intensity of exertion did you feel during this session?"; "How was your workout?"; Foster et al., 2001; Haile et al., 2015; Robertson and Noble, 1997). These types of measures have provided a fine-grained understanding of how people manage exercise intensity through pacing strategies (i.e., conscious effort

management throughout an exercise bout) to prevent metabolic and biomechanical failures (e.g., fatigue accumulation, slower rates of neuromuscular recovery, overtraining syndrome; e.g., Meeusen et al., 2013; Thiel et al., 2018; Vieira et al., 2022).

Another key observation from the literature on RPE is that increased perceived levels of exertion are negatively linked with the intensity of pleasure felt during the session of physical exercise (for a theoretical review, see Ekkekakis et al., 2011; for recent studies, see Hartman et al., 2019; Hutchinson et al., 2020; Frazão et al., 2016). It has also been evidenced that decreasing the intensity of a resistance exercise session can elicit higher levels of experienced and retrospective pleasure toward physical exercise (e.g., Hutchinson et al., 2023). Besides, positive changes in hedonic responses during moderate intensity exercise have been linked to future physical activity (Rhodes and Kates, 2015). Taken together, these findings suggest that experienced and retrospective levels of pleasure toward physical exercise substantially affect the individual appraisal of the activity and may ultimately impact future engagement and commitment to physical exercise. In other words, physical exercise will be more likely reinforced by sessions that are experienced as pleasant, whereas if it is perceived as unpleasant it will more likely be avoided (e.g., Teixeira et al., 2022). Here we aim to further increase current knowledge on the impact of perceived exertion on the level of pleasure experienced during physical exercise. Specifically, we aim to better identify sessions of physical exercise that lead to an increase (or decrease) in the remembered level of pleasure that was experienced by an individual during physical exercise, that is, *retrospective pleasure*. We advance that prospective thinking can provide key insight into this research question.

In this study, prospective thinking refers to individuals' ability in anticipating the intensity of a forthcoming session of physical exercise, that is, *before* the physical exercise session has started (e.g., "What intensity of exertion do you expect to feel during this

session?"). We labeled this process as *prospective RPE*. As previously mentioned, few studies have examined prospective or anticipatory types of RPE. Nevertheless, preliminary evidence revealed that mismatches (either overestimation or underestimation) between anticipated and experienced exertion is associated with lower frequency of daily physical activity, negative attitudes about physical exercise, higher body mass index, as well as poor cardiorespiratory fitness (Haile et al., 2008; Hunt et al., 2007; Kane et al., 2010; Poulton et al., 2002). The present study aims to push forward in this direction by examining whether mismatches between anticipated and remembered exertion can inform the level of pleasure that was felt by the individual during a physical exercise session. To do so, we will capitalize on the main dynamic pertaining to *reward prediction errors* (Schultz et al., 2016; Kieslich et al., 2021).

A key tenet from the literature on reward processing is that the reactivity to reward does not depend on the value of rewarding outcomes per se, but is instead driven by the difference between expected and actual outcomes, namely a reward prediction error. This pattern has been evidenced by studies showing that, when a rewarding outcome is better than expected, it induces more pleasure than a reward that matches prior expectations (i.e., a positive reward prediction error; for a review, see Schultz et al., 2016; Kieslich et al., 2021).

Against this background, and given the correspondence between RPE and pleasure, we posit that physical exercise sessions that are experienced with a lesser level of perceived exertion than anticipated (i.e., a positive RPE-based prediction error) should be associated with a higher level of subjective pleasure experienced during a session of physical exercise. In other words, experiencing lesser exertion than expected should induce a higher level of pleasure during physical exercise, and vice versa (i.e., a negative RPE-based prediction error). We will test this hypothesis by using ecological momentary assessment to be filled out

by participants just before (anticipatory RPE) and directly after (retrospective RPE, retrospective running pleasure) running sessions that are part of a start-to-run program.

2. METHODS

2.1. Ethics

The protocol of the study was approved by the Ethics Committee of Saint-Luc University Hospital (UCLouvain; #2022/21JUI/247).

2.2. Participants

We will recruit our participants among UCLouvain students (except from the Faculty of Movement and Rehabilitation Sciences, in order not to interfere with the physical activity programs of the Bachelor/Master of Physical Education and Physiotherapy) who want to participate in our start-to-run study. Participants will be recruited via flyers with a QR code directing them to an online screening tool (LimeSurvey platform). The experimenters will also make announcements in the auditorium (after obtaining the agreement of the Professor in charge of the teaching unit). The online screening tool will first include an informed consent form. An email address and a phone number will be provided to allow participants to ask questions before agreeing or declining to participate in the study. The screening tool will then ask the potential participants (i.e., the ones who have agreed to take part in the study) to complete the International Physical Activity Questionnaire (IPAQ; Craig et al., 2003). Since it is a start-to-run program, we will recruit individuals corresponding to the “low” and “medium” physical activity categories of the IPAQ. To limit the health risks related to running exercise, each participant will be asked to complete the Physical Activity Readiness Questionnaire for Everyone (PAR-Q+; Warburton et al., 2011) in the presence of one of the two team supervisors (BdG) who has 20 years of exercise testing experience. In the first step, only the first questions of the questionnaire will be filled out. Those who will answer NO to

the first 7 questions of the PAR-Q+ receive a “green light” and will be immediately allowed to participate in the study. Those who will answer YES to one or more questions will have to meet the team supervisor who will go through the additional questions (pages 2 and 3 of the PAR-Q+). If participants answer YES to any of the questions, they will not be able to start the study and will be advised to see a (sports) physician. Participants who do not meet the study selection criteria will be informed and will not participate in the start-to-run program.

2.3. Submaximal exercise test

Before the start of the running program, participants will perform a submaximal exercise test (SET) to estimate their Maximal Oxygen Consumption (VO_{2max}). The SET will consist of the “1-mile track jog test” (George et al., 1993). This specific SET is chosen because of its resemblance with George and colleagues’ (1993) study, namely a running exercise test, the comparable study population: both males and females, a similar age category, and college students. Additionally, this SET is chosen over a maximal exercise test because a submaximal exercise test limits the health risks linked to exercise testing in unfit subjects (George et al., 1993). The SET will be repeated before the Louvain-la-Neuve 5 Miles to index the fitness level progression.

2.4. Start-to-run program

The primary goal of the start-to-run program is to provide a context that will allow participants to perform enough running sessions (minimum 5; see also the **Sample size estimation** section) to test our hypothesis on the impact of RPE prediction errors on running pleasure. The start-to-run program is planned to begin in October 2023 and will end on the day of the “Louvain-la-Neuve 5 Miles”, which is a running event that will occur on March 20th, 2024. The end goal of the start-to-run program is that each participant takes part in this running event.

Participants will be asked to undertake weekly “free” running sessions at a self-selected (or preferred) dose. Specifically, participants will be encouraged to self-select their running frequency, intensity, and duration, in which they will be allowed to undertake these sessions alone or in groups, where they want. Participants will also be allowed to listen to music if they want to. In this start-to-run program, “self-selected” running will thus be chosen over “imposed” running. Specifically, when the intensity of physical exercise is self-selected, rather than imposed, it appears to foster greater tolerance to higher intensity levels, a greater sense of autonomy toward physical exercise, and also increased levels of enjoyment and positive affect while exercising (Ekkekakis et al., 2011; Oliveira et al., 2015; Vazou-Ekkekakis and Ekkekakis, 2009). Moreover, because prospective thinking is a crucial factor in maintaining autonomy in daily life (e.g., Blondelle et al., 2022; Kennard and Lewis, 2006), “self-selected” running (i.e., allowing participants to choose the duration, frequency, and intensity of each “free” running session) should also be an optimal approach to facilitate individuals’ ability in anticipating the exertion intensity of a forthcoming session of physical exercise, that is, to generate prospective RPE. This approach fits also well with training procedures derived from the ecological dynamic approach to physical exercise (e.g., David et al., 2016; Rudd et al., 2021). Specifically, this approach advocates for physical exercise behaviors that consider the relationship between individuals’ characteristics and functional aspects of their environment (e.g., running sessions undertaken under multiple contexts).

A key aspect of the start-to-run program is that participants will be asked to record each session on a running app called *Formyfit* (<https://www.formyfit.com/>). Each participant will be asked to download the Formyfit app on their smartphone and will be offered an armband pocket to be able to run with their smartphones. This smartphone app will allow the recording of running session data (while respecting the General Data Protection Regulation, GDPR), including distance, and average speed, as well as the possibility to obtain heart rate

data (in the case the participant runs with a heart rate sensor). Running sessions could be undertaken outdoors or indoors (on a treadmill). For the outdoor session, the GPS of the Smartphone is used to estimate the running distance and average speed. When performing an indoor session, the Formyfit app records the time and participants will be informed that they will have to encode the distance manually in the app at the end of the session. Importantly, since participants will be novice or low-frequent runners, the Formyfit app will recommend running duration based on participants' VO₂max (estimated from the SET). These recommendations will be made available to the participant on the app and could be downloaded in a document format. The participants will be able to choose whether or not they want to follow the proposed running duration. Participants will also have access to the general Formyfit dashboard app featuring summary information on their running sessions (e.g., frequency, average distance, average speed, and heart rate).

In addition to the free run session, participants will be invited on a weekly basis to attend a running session supervised by coaches (i.e., 6 Master, MA, students in Physical Education at UCLouvain will be involved in the start-to-run program). These coaching sessions will occur at different locations on the Louvain-la-Neuve campus of UCLouvain. This type of session will be given in groups, but participants will be asked to run at their preferred pace (e.g., to walk when they feel the need to do so). These coaching sessions will be undertaken without music (i.e., headphones). Four different schedules will be proposed each week with a maximum of 10 participants per group. There will be two coaches in each group session. Each group session will start with a warm-up and ends with a cool-down and stretching routine. These steps will be guided by the coaches. The coaches will run with the participants, with one coach running at the front of the group, and the other at the back. This will allow the coaches to supervise the fastest and slowest runners and give personal advice (e.g., advice on running techniques and running stance) during the running session.

Participants will also receive general information on running techniques, nutrition, and sports injury prevention through the articles that are available on the Formyfit blog

(<http://blog.formyfit.com/category/articlesconseils/nutrition/>). Moreover, because self-selected exercise may also increase the odds of adopting inappropriate exercise intensity (e.g., Johnson & Phipps, 2006), participants will have the possibility to discuss with the coaches (during the weekly “guided” sessions) how to adjust their “free” running session if needed (see also the subsection “*Secondary Measures*” below).

2.5. Primary measures

2.5.1. Prediction error of RPE. RPE will be undertaken directly before (prospective RPE) and after (retrospective RPE) each running session on the Formyfit app (see **Figure 1**). Based on Foster and colleagues’ findings (2001), participants will be asked to provide a prospective or retrospective rating of their perceived RPE of the overall running session. Hence, we will explain to the participants that they will have to provide a global rating of the entire running session (Foster et al., 2001).

RPE will be indexed using the French adaptation of the Borg’s Category Ratio-10 (CR-10) RPE scale (Haddad et al., 2013; see also, Foster et al., 2001; Borg, 1998). Specifically, for prospective RPE, participants will have to estimate the intensity of exertion (“effort” in French) they expect to feel during the forthcoming running session (“What intensity of exertion do you expect to feel during this session?”) on a scale ranging from 0 (“null”, “nulle” in French) to 10 (“maximal”, “maximale”), with other integers on the scale assigned modifiers (1 = “very very light” (“très très légère”), 2 = “light” (“légère”), 3 = “moderate” (Modérée), 4 = “somewhat hard” (“assez dure”), 5 = “hard” (“dure”), 6 = [no verbal anchor], 7 = very hard (“très dure”), 8 = [no verbal anchor], 9 = [no verbal anchor]; see **Figure 1A**). For retrospective RPE, participants will have to report the intensity of exertion they experienced during the running session (“What intensity of exertion did you

feel during this session?”) on a scale ranging from 0 (null) to 10 (maximal), with the same integers on the scale assigned modifiers (see **Figure 1B**). Participants will not have access to their prospective RPE. They will also be asked to formulate their retrospective RPE without trying to remember or reflect on their prospective RPE.

Using this in-app procedure, prediction error will be operationalized using *absolute* and *relative* indexes (e.g., Mattes and Roheger, 2020), with:

$$\text{absolute_prediction_error} = \text{prospective RPE} - \text{retrospective RPE}$$

$$\text{relative_prediction_error} = \frac{\text{prospective RPE} - \text{retrospective RPE}}{\text{prospective RPE}}$$

Absolute and relative indexes of prediction errors complement each other.

Specifically, given the same absolute change, the relative change is larger in magnitude if the prospective RPE value is at a higher level than if it is at a lower level. For instance, (i) with a prospective RPE of 3 and a retrospective RPE of 5, the absolute RPE prediction error = -2 and the relative RPE prediction error = -0.33; (ii) with a prospective RPE of 5 and a retrospective RPE of 7, the absolute RPE prediction error is still = -2, but the relative RPE prediction error is now -0.40. For both the *absolute* and the *relative* indexes, a positive prediction error (i.e., the experienced level of exertion is lower than expected) corresponds to a positive scores difference, and a negative prediction error (i.e., the experienced level of exertion is higher than expected) corresponds to a negative score difference.

2.5.2. Retrospective Running pleasure. Retrospective running pleasure (variable name = running_pleasure) will be indexed using a single item adapted from the “*Single-item measure of enjoyment during exercise*” developed by Stanley and Cumming (2010). The present study adopts a pleasure-oriented approach and aimed to use particular modifiers that are suitably distinct from each other for participants to express differential amounts of pleasure they had experienced during their running session. Specifically, directly after having completed the retrospective RPE, participants will have to estimate the level of pleasure they experienced

during the overall running session (“What intensity of pleasure did you feel during this session?”) on a 7-point Likert scale ranging from 0 (“not at all”) to 6 (“extremely”), with other integers on the scale assigned modifiers (1 = “very little”, 2 = “slightly”, 3 = “moderately”, 4 = “quite a bit”, 5 = “very much”; see **Figure 1B**).

2.6. Secondary measures

2.6.1. Average speed and distance of a running session. For each running session, the total running distance (variable name = *distance*) and average speed (variable name = *average_speed*) will be recorded with the Formyfit app (see **Figure 1C**). These measures will be implemented as covariates in our statistical models (see also the subsection “*data analytic plan*” below).

2.6.2. Additional covariates for the effect of the RPE prediction error on running pleasure.

Previous research has shown that running in a group impacts the level of pleasantness of physical exercise sessions (e.g., Xie et al., 2020). Hence, we will examine whether running with or without another person during the “free” sessions (**running alone vs. running with another person vs. running with more than one person**) or **running during the coaching session per se** modulates the impact of RPE prediction error on running pleasure (**variable name = running_group**). We will also examine whether the impact of RPE prediction error on running pleasure **is modulated by the degree of familiarity linked to the running program (variable name = familiarity)**. Indeed, individuals might get better at predicting their level of perceived exertion for habitual running trails, which can decrease the impact of RPE prediction error on running pleasure. **These data will be recorded directly before (Figure 1A) each running session on the Formyfit app by the participant. In addition, because listening to music might modulate the level of perceived exertion during physical exercise (for a review, see Ballmann et al., 2021), we will also examine whether running with music modulates the effect of PE prediction error on running (variable name = music)**. To do so, participants will

have to report, directly after the running session, whether or not they ran with music (see **Figure 1B**). Lastly, participants will have the option to write a free commentary on the Formyfit app (see **Figure 1B**).

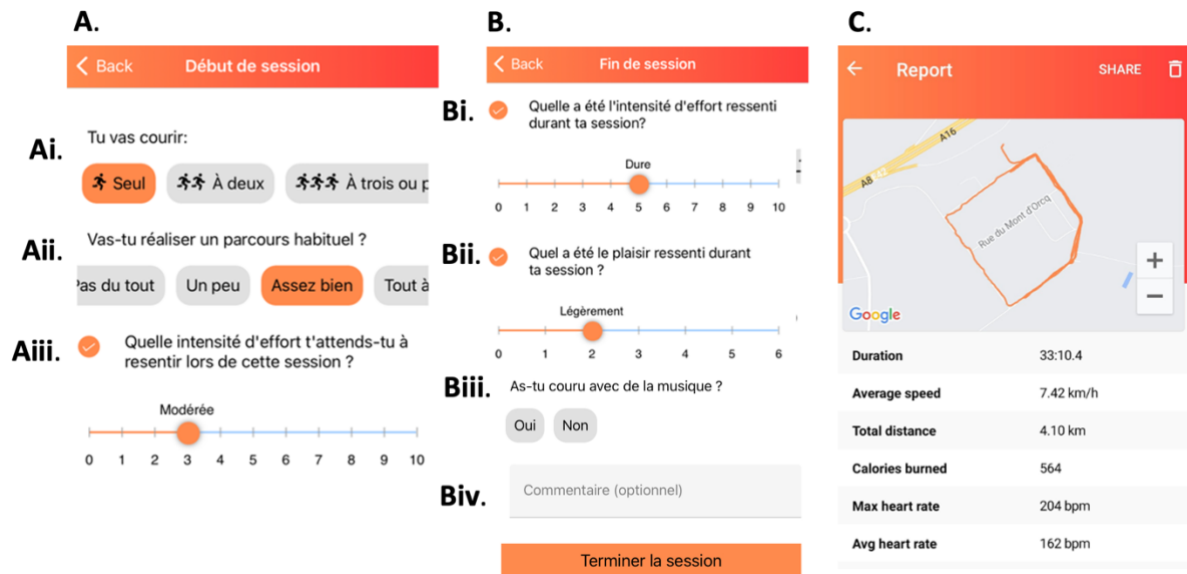


Figure 1. **A.** Pre-session measurements. **Ai:** reporting on the inter-individual nature of the running session (running a free session alone, running a free session with another person, running a free session with more than one person, or running a coaching session); **Aii:** habit level of the running session (not at all, a little bit, quite well, very much); **Aiii:** prospective RPE. **B.** Post-session measurements: **Bi:** retrospective RPE; **Bii:** retrospective running pleasure; **Biii:** use of music while running (yes or no); **Biv:** free comment option. **C:** Running session data (total distance, duration, average speed, heart rate).

2.7. Data analytical plan

To test our hypothesis, we will run linear mixed models (LMM). LMM is becoming a popular alternative to repeated-measures ANOVA analyses in experimental psychology (Magezi, 2015). In looking at the effects of (absolute and relative) *prediction error RPE* on *running pleasure* in our study, there are three main advantages of adopting an LMM approach over typical repeated-measures ANOVA. First, when using LMM, it is possible to specify random effects (i.e., here participants are treated as nested random factors). Instead of bundling this variance into an error term, LMM partitions the variance that is associated with these differences explicitly. Second, LMM allow to account for individual differences in the

effect of a predictor by adding random slopes. In the present study, the size and direction of the *prediction-error_RPE* effect on *running_pleasure* could differ across individuals. Third, by contrast to repeated-measures ANOVA, LMM can handle missing measurements and different numbers of measurements per subject. In the case of this dataset, the number of running sessions undertaken across the start-to-run program will differ between each participant. For these reasons, the LMM approach was more appropriate.

To run the LMM, we will use the lme4 package (Bates et al., 2015) and run it on Jamovi (Version 2.3.21.0). Significance will be calculated using the lmerTest package (Kuznetsova et al., 2017), which applies Satterthwaite's method to estimate degrees of freedom and generate *p*-values for mixed models. The model will be run twice: once with *absolute_prediction_error* (model 1) and another with *relative_prediction_error* (model 2) as a fixed effect. All predictor variables will be grand-mean centered. We will build our multilevel model by adopting the following three-steps sequence, separately for absolute prediction error and relative prediction_error:

Step 1 (null model). We will first run the null model by including participants as a cluster variable with random effect, and *running_pleasure* as the dependent variable with the following model specification:

$$running_pleasure \sim + (1|participants)$$

Step 2. As a second step in the model-building process, we will add the fixed effect of *absolute_prediction_error* (step 2a) or *relative_prediction_error* (step 2b) prediction_error, average_speed, distance, running_group, familiarity, and music with fixed slope:

$$Step\ 2a: \quad running_pleasure \sim 1 + absolute_prediction_error + distance + average_speed + running_group + familiarity + music + (1|participants)$$

$$Step\ 2b: \quad running_pleasure \sim 1 + relative_prediction_error + distance + average_speed + running_group + familiarity + music + (1|participants)$$

Step 3. As a third and final step, we will run the model with `absolute_prediction_error` (*step 3a*) or `relative_prediction_error` (*step 3b*) as a fixed effect with random slope, and `average_speed`, `distance`, `running_group`, `familiarity`, and `music` with fixed slope:

Step 3a: $running_pleasure \sim 1 + absolute_prediction_error + distance + average_speed + running_group + familiarity + music + (1 + -absolute_prediction_error|participants)$

Step 3b: $running_pleasure \sim 1 + relative_prediction_error + distance + average_speed + running_group + familiarity + music + (1 + relative_prediction_error|participants)$

2.8. Pilot data

Between October 2022 and December 2022, we ran a pilot study to obtain estimates for fixed and random effects and effect sizes. This pilot study also allowed us to pretest the procedure pertaining to the start-to-run program using a beta version of the Formyfit app. These pilot data were obtained on a sample of 19 participants (4 males, 15 females; age: mean = 20.8, median = 21, SD = 2.51, range = 18-25; height (centimeters): mean = 170, median = 167, SD = 8.57, range = 160-192; weight (kilograms): mean = 69.5, median = 64.1, SD = 13.4, range = 52-92; VO_{2max}: mean = 40.0, median = 41.3, SD = 6.22, range = 29-51). Participants were UCLouvain students and corresponded to the “low” and “medium” physical activity categories of the IPAQ and received a “green light” on the PAR-Q+.

In the first two weeks of October 2022, all participants undertook the SET (i.e., the 1-mile track jog test) under standard conditions on an indoor 250-meter track. The start-to-run program was similar to the procedure described in *subsection 2.4* (i.e., self-selected mode of running, weekly guided running session), except that: (i) there was no end goal of participating at a running event (i.e., the Louvain-la-Neuve 5 Miles), (ii) the program lasted less than four months (it ended in December, not in March), and (iii) only primary, not secondary, measures were recorded on the beta version of the Formyfit app (i.e., prospective and retrospective RPE, total running distance, and average running speed).

This start-to-run program allowed us to obtain pilot data on 19 participants across 228 running sessions (mean of 12.39 running sessions per participant, median = 11.50, SD = 6.51; minimum = 5, maximum = 32). Initially, the total number of recorded running sessions was 261, but 10 sessions were deleted because the running distance was very low relative to the other running sessions (< 1 kilometer), and 23 sessions were deleted due to at least one missing event (i.e., when a participant did not report prospective RPE, retrospective RPE, and/or running pleasure rating). The SET sessions ($n = 19$) were not used for this primary data analysis.

Using this pilot data set, we ran linear mixed model analysis using the lme4 package (Bates et al., 2015) on Jamovi (Version 2.3.21.0) The results from these analyses are detailed in **Table 1** (model 1) and **Table 2** (model 2) and illustrated in **Figure 2**. We built our multilevel model by adopting the following three-steps sequence described in subsection 2.7, with the difference that only covariate measures on “distance” and “average_speed” were recorded for the pilot study:

Step 1 (null model). $running_pleasure \sim 1 + (1|participants)$

Step 2a: $running_pleasure \sim 1 + absolute_prediction_error + distance + average_speed + (1|participants)$

Step 2b: $running_pleasure \sim 1 + relative_prediction_error + distance + average_speed + (1|participants)$

Step 3a: $running_pleasure \sim 1 + absolute_prediction_error + distance + average_speed + (1 + -absolute_prediction_error|participants)$

Step 3b: $running_pleasure \sim 1 + relative_prediction_error + distance + average_speed + (1 + relative_prediction_error|participants)$

Step 1 (null model). The first step in the model indicated that ICC = .21, which means that differences across participants account for about 21% of the variability in individuals’

level of running pleasure. As shown in **Table 1**, the intercept variance is .37 and the within-participant variance is 1.38. In short, results provide evidence for a nested data structure that requires multilevel modeling rather than a single-level data analytic approach. Specifically, an ICC, even as small as .10 (Kahn, 2011), suggests that participants (Level 2 variable) explain the heterogeneity of running pleasure scores. ICC value near zero suggests that a model including Level 1 variables only is appropriate, and, hence, there may be no need to use multilevel modeling (a simpler OLS regression approach may be more parsimonious).

Step 2 (model 2 and model 3). This second step involves testing a random intercept and fixed slope model. In other words, the relationship between running pleasure and absolute prediction error RPE is assumed to be identical across all participants, while also taking into account the effect of running distance and average speed on running pleasure. As described earlier (*subsection 2.7*), we used grand-mean centered scores for our analyses. As shown in **Table 1**, model 2 results indicate that a 1-unit increase in absolute prediction error RPE is associated with a significant ($p < .001$) .15 increase in running pleasure (see also **Figure 2A**). Model 3 results indicate that a 1-unit increase in absolute prediction error RPE is associated with a significant ($p < .01$) .57 increase in running pleasure (see **Table 2** and **Figure 2B**). Importantly, -2 Log likelihood and AIC values indicate that there is an increased model fit between Step 1 and Step 2 (see Table 2). For models 2 and 3, the conditional R^2 (which considers the variance of both the fixed and random effects) are .42 and .41 respectively, which is indicative of moderate effect sizes.

Step 3 (model 4 and model 5). This third step involves testing a random intercept and random slope for the variable absolute prediction error. In other words, it answers the question of whether the relationship between prediction error RPE and running pleasure varies across participants. For both model 4 and model 5, we observed a similar effect of prediction error RPE on running pleasure. -2 Log likelihood and AIC values also indicate that

there is no increase in model fit between Step 2 and Step 3 (see Table 2). Moreover, the random effect variances were close to zero, which indicates that there is little variance to be accounted for in the random slope in the data (Rights and Jason, 2019). Hence, these findings suggest that the relationship between prediction error RPE and running pleasure does not vary across participants.

Taken together, the pilot findings provide a preliminary step in the validation of our hypothesis, by showing that absolute and relative prediction error RPE significantly impact the level of pleasure experienced during a running session. These findings are important as they not only offer preliminary support for the hypothesis of the study but also suggest that the model with random intercept and fixed slope (step 2) is the best model. Indeed, the model with random intercept and random slope does not result in a better fit.

2.8. Sample size estimation

A priori power analysis was performed using Power IN Two-level designs software which is designed to estimate standard errors of regression coefficients in hierarchical linear models for power calculations (Snijders and Bosker, 1993). In line with recent guidelines that suggest running power analysis based on the lowest meaningful estimate of the effect size (Dienes, 2021), we ran sample size estimation analyses with a conditional R^2 of .40 based on the results obtained in the step 2 model of relative prediction error. Accordingly, if α is chosen at .05, an effect size of .40 is what we expect, and a power of .80 is desired, then a sample of 28 participants along 5 measurement points (i.e., a running session) is required for third-step models.

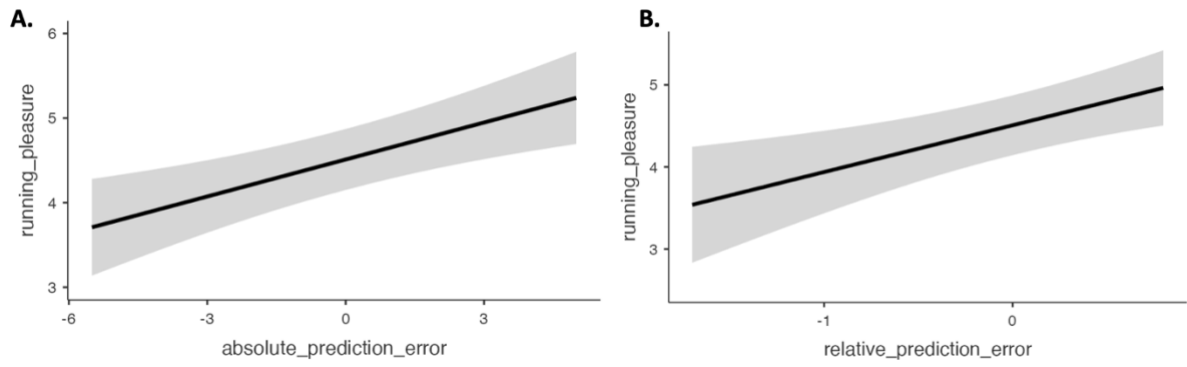


Figure 2. A. Fixed effect of (A) absolute and (B) relative prediction error RPE and on running pleasure. Residual and standard error. Semi-transparent grey areas indicate the 95% CI of the fixed effect.

Table 1. Linear Mixed Model with absolute_prediction_error.

	Null (Step 1)	Random Intercept and Fixed Slope (Step 2)	Random Intercept and Random Slope (Step 3)
<i>Variable</i>			
Intercept	4.47*** (0.16)	4.51*** (0.16)	4.51*** (0.17)
Absolute_prediction_error		0.15*** (0.05)	0.14*** (0.05)
Average_speed		0.31*** (0.07)	0.31*** (0.07)
Distance		0.15** (0.05)	0.15** (0.05)
<i>Variance components</i>			
Within-participant variance	1.38	1.17	1.17
Intercept variance	0.37	0.43	0.43
Absolute_prediction_error			0.001
<i>Additional information</i>			
ICC	0.21		
-2 Log likelihood (FILM)	739.78	713.18***	713..16
Number of estimated parameters	3	6	7
Conditional R^2	0.21	0.42	0.42
Pseudo R^2		0.20	0.20
AIC	747.78	725.76	729.16

*Note: FIML = full information maximum likelihood estimation; Total number of running sessions = 228, number of participants = 19. Values in parentheses are standard errors; t-statistics were computed as the ratio of each regression coefficient divided by its standard error. * $p < .05$. ** $p < .01$. *** $p < .001$.*

Table 2. Linear Mixed Model with relative_prediction_error.

	Null (Step 1)	Random Intercept and Fixed Slope (Step 2)	Random Intercept and Random Slope (Step 3)
<i>Variable</i>			
Intercept	4.47*** (0.16)	4.51*** (0.17)	4.51*** (0.17)
Relative_prediction_error		0.57** (0.19)	0.58** (0.18)
Average_speed		0.31*** (0.07)	0.31*** (0.07)
Distance		0.15** (0.05)	0.15** (0.05)
<i>Variance components</i>			
Within-participant variance	1.38	1.18	1.18
Intercept variance	0.37	0.44	0.44
Relative_prediction_error			0.001
<i>Additional information</i>			
ICC	0.21		
-2 Log likelihood (FILM)	739.78	713.10***	713.30
Number of estimated parameters	3	6	7
Conditional R^2	0.21	0.41	0.41
Pseudo R^2		0.20	0.20
AIC	747.78	727.32	731.30

*Note: FIML = full information maximum likelihood estimation; Total number of running sessions = 228, number of participants = 19. Values in parentheses are standard errors; t-statistics were computed as the ratio of each regression coefficient divided by its standard error. * $p < .05$. ** $p < .01$. *** $p < .001$.*

4. RESULTS

[the Results section is only to be completed in Stage 2 following data collection]

5. DISCUSSION

[the Discussion section is only to be completed in Stage 2 following data collection]

Data, code, and materials availability

The Jamovi file containing the data, and LMM analyses (including model specification for reproducing the LMM analyses using other statistical software) of the pilot study (i.e., the analyses presented in the pre-registration) are openly available on the Open Science Framework (OSF) website <https://osf.io/2sb86/>. All data, analysis code, and any other materials of the main study will be made openly available at the same OSF link.

REFERENCES

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48.

<https://doi.org/10.18637/jss.v067.i01>

Blondelle, G., Sugden, N., & Hainselin, M. (2022). Prospective memory assessment: Scientific advances and future directions. *Frontiers in psychology*, 13, 958458.

<https://doi.org/10.3389/fpsyg.2022.958458>

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

Brevers, D., Billieux, J., de Timary, P., Desmedt, O., Maurage, P., Perales, J. C., Suárez-Suárez, S., & Bechara, A. (2023). Physical Exercise to Redynamize Interoception in Substance use Disorders. *Current neuropharmacology*, Advance online publication.

<https://doi.org/10.2174/1570159X21666230314143803>

Craig, C. L., Marshall, A. L., Sjöström, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., Pratt, M., Ekelund, U., Yngve, A., Sallis, J. F., & Oja, P. (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine and science in sports and exercise*, 35(8), 1381–1395. <https://doi.org/10.1249/01.MSS.0000078924.61453.FB>

D'Argembeau, A., Ortoleva, C., Jumentier, S., & Van der Linden, M. (2010). Component processes underlying future thinking. *Memory & cognition*, 38(6), 809–819. <https://doi.org/10.3758/MC.38.6.809>

Dauids, K., Araújo, D. & Brymer, E. Designing Affordances for Health-Enhancing Physical Activity and Exercise in Sedentary Individuals. *Sports Med* 46, 933–938 (2016).

<https://doi.org/10.1007/s40279-016-0511-3>

Demblon, J., & D'Argembeau, A. (2014). The organization of prospective thinking: evidence of event clusters in freely generated future thoughts. *Consciousness and cognition*, 24, 75–83. <https://doi.org/10.1016/j.concog.2014.01.002>

Ekkekakis, P., Parfitt, G., & Petruzzello, S. J. (2011). The pleasure and displeasure people feel when they exercise at different intensities: decennial update and progress towards a tripartite rationale for exercise intensity prescription. *Sports medicine (Auckland, N.Z.)*, 41(8), 641–671. <https://doi.org/10.2165/11590680-000000000-00000>

Foster, C., Florhaug, J. A., Franklin, J., Gottschall, L., Hrovatin, L. A., Parker, S., Doleshal, P., & Dodge, C. (2001). A new approach to monitoring exercise training. *Journal of strength and conditioning research*, 15(1), 109–115.

Frazão, D. T., de Farias Junior, L. F., Dantas, T. C. B., Krinski, K., Elsangedy, H. M., Prestes, J., Hardcastle, S. J., & Costa, E. C. (2016). Feeling of pleasure to high-intensity interval exercise is dependent of the number of work bouts and physical activity status. *PLoS ONE*, 11(3), Article e0152752. <https://doi.org/10.1371/journal.pone.0152752>

George, J. D., Vehrs, P. R., Allsen, P. E., Fellingham, G. W., & Fisher, A. G. (1993).

VO₂max estimation from a submaximal 1-mile track jog for fit college-age individuals.

Medicine and Science in Sports and Exercise, 25(3), 401-406.

Haile, L., Gallagher, M., Jr., & Robertson, R. J. (2015). *Perceived exertion laboratory manual: From standard practice to contemporary application*. Springer Science + Business Media. <https://doi.org/10.1007/978-1-4939-1917-8>

Haddad, M., Chaouachi, A., Castagna, C., Hue, O., Tabben, M. et al.. (2013). Validity and psychometric evaluation of the French version of RPE scale in young fit males when monitoring training loads. *Science & Sports*, 28(2), 29-35.

Haile, L., Ledezma, C. M., Koch, K. A., Shouey, L. B., Aaron, D. J., Goss, F. L., Robertson, R. J (2008). Predicted, Actual and Session Muscle Pain and Perceived Exertion During Cycle Exercise in Young Men: 1793. *Medicine & Science in Sports & Exercise* 40(5):p S301, DOI: [10.1249/01.mss.0000323631.85365.e1](https://doi.org/10.1249/01.mss.0000323631.85365.e1)

Hartman, M. E., Ekkekakis, P., Dicks, N. D., & Pettitt, R. W. (2019). Dynamics of pleasure-displeasure at the limit of exercise tolerance: conceptualizing the sense of exertional physical fatigue as an affective response. *The Journal of experimental biology*, 222(Pt 3), jeb186585. <https://doi.org/10.1242/jeb.186585>

Hunt, S. E., DiAlesandro, A., Lambright, G., Williams, D., Aaron, D., Goss, F., Robertson, R. (2007). Predicted and actual leg pain and perceived exertion during cycle exercise in young women. *Medicine & Science in Sports & Exercise*, 39, S485..

Hutchinson, J. C., Jones, L., Ekkekakis, P., Cheval, B., Brand, R., Salvatore, G. M., Adler, S., & Luo, Y. (2023). Affective Responses to Increasing- and Decreasing-Intensity Resistance Training Protocols, *Journal of Sport & Exercise Psychology*, 45(3), 121-137.

<https://doi.org/10.1123/jsep.2022-0243>

Hutchinson, J. C., Zenko, Z., Santich, S., & Dalton, P. C. (2020). Increasing the Pleasure and Enjoyment of Exercise: A Novel Resistance-Training Protocol, *Journal of Sport and Exercise Psychology*, 42(2), 143-152. <https://doi.org/10.1123/jsep.2019-0089>

Johnson, J. H., & Phipps, L. K. (2006). Preferred method of selecting exercise intensity in adult women. *Journal of strength and conditioning research*, 20(2), 446–449. <https://doi.org/10.1519/R-17935.1>

Kahn, J. H. 2011. Multilevel modeling: Overview and applications to research in counseling psychology. *Journal of Counseling Psychology*, 58, 257-271.

Kane, I., Robertson, R. J., Fertman, C. I., McConnaha, W. R., Nagle, E. F., Rabin, B.S., Rubinstein, E.N. (2010). Predicted and actual exercise discomfort in middle school children. *Medicine & Science in Sports & Exercise*. 42, 1013–21.

Kennard, J., Lewis, K. (2006). Maintaining autonomy: The role of prospective memory in rehabilitation. *International Journal of Therapy and Rehabilitation*, 13(4), 150-150.

Kieslich, K., Valton, V., & Roiser, J. P. (2022). Pleasure, Reward Value, Prediction Error and Anhedonia. *Current topics in behavioral neurosciences*, 58, 281–304. https://doi.org/10.1007/7854_2021_295

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>

Magezi D. A. (2015). Linear mixed-effects models for within-participant psychology experiments: an introductory tutorial and free, graphical user interface (LMMgui). *Frontiers in psychology*, 6, 2. <https://doi.org/10.3389/fpsyg.2015.00002>

Mattes, A., & Roheger, M. (2020). Nothing wrong about change: the adequate choice of the dependent variable and design in prediction of cognitive training success. *BMC medical research methodology*, 20(1), 296. <https://doi.org/10.1186/s12874-020-01176-8>

Meeusen, R., Duclos, M., Foster, C., Fry, A., Gleeson, M., Nieman, D., Raglin, J., Rietjens, G., Steinacker, J., Urhausen, A., European College of Sport Science, & American College of Sports Medicine (2013). Prevention, diagnosis, and treatment of the overtraining syndrome: joint consensus statement of the European College of Sport Science and the American College of Sports Medicine. *Medicine and science in sports and exercise*, 45(1), 186–205. <https://doi.org/10.1249/MSS.0b013e318279a10a>

Oliveira, B. R., Deslandes, A. C., & Santos, T. M. (2015). Differences in exercise intensity seems to influence the affective responses in self-selected and imposed exercise: a meta-analysis. *Frontiers in psychology*, 6, 1105. <https://doi.org/10.3389/fpsyg.2015.01105>

Poulton, R., Trevena, J., Reeder, A. I., & Richard, R. (2002). Physical health correlates of overprediction of physical discomfort during exercise. *Behaviour research and therapy*, 40(4), 401–414. [https://doi.org/10.1016/s0005-7967\(01\)00019-5](https://doi.org/10.1016/s0005-7967(01)00019-5)

Rhodes, R. E., Kates, A. (2015). Can the Affective Response to Exercise Predict Future Motives and Physical Activity Behavior? A Systematic Review of Published Evidence, *Annals of Behavioral Medicine*, 49(5), 715–731, <https://doi.org/10.1007/s12160-015-9704-5>

Robertson, R. J., & Noble, B. J. (1997). Perception of physical exertion: methods, mediators, and applications. *Exercise and sport sciences reviews*, 25, 407–452.

Rudd, J. R., Woods, C., Correia, V., Seifert L. & Davids, K. (2021) An ecological dynamics conceptualisation of physical ‘education’: Where we have been and where we could go next, *Physical Education and Sport Pedagogy*, 26(3), 293-306. <https://doi.org/10.1080/17408989.2021.1886271>

Schacter, D. L., Benoit, R. G., & Szpunar, K. K. (2017). Episodic Future Thinking: Mechanisms and Functions. *Current opinion in behavioral sciences*, 17, 41–50. <https://doi.org/10.1016/j.cobeha.2017.06.002>

- Schultz W. (2016). Dopamine reward prediction error coding. *Dialogues in clinical neuroscience*, 18(1), 23–32. <https://doi.org/10.31887/DCNS.2016.18.1/wschultz>
- Snijders, T. A. B., and Bosker, R. J. (1993). Standard errors and sample sizes for two-level research. *Journal of Educational and Behavioral Statistics*, 18, 237–259. <https://doi.org/10.3102/10769986018003237>
- Stanley, D. M., & Cumming, J. (2010). Are we having fun yet? Testing the effects of imagery use on the affective and enjoyment responses to acute moderate exercise. *Psychology of Sport and Exercise*, 11(6), 582–590. <https://doi.org/10.1016/j.psychsport.2010.06.010>
- Teixeira, D. S., Rodrigues, F., Cid, L., & Monteiro, D. (2022). Enjoyment as a Predictor of Exercise Habit, Intention to Continue Exercising, and Exercise Frequency: The Intensity Traits Discrepancy Moderation Role. *Frontiers in psychology*, 13, 780059. <https://doi.org/10.3389/fpsyg.2022.780059>
- Thiel, C., Pfeifer, K. & Sudeck, G. Pacing and perceived exertion in endurance performance in exercise therapy and health sports. *Ger J Exerc Sport Res* 48, 136–144 (2018). <https://doi.org/10.1007/s12662-017-0489-5>
- Vazou-Ekkekakis, S., & Ekkekakis, P. (2009). Affective consequences of imposing the intensity of physical activity: Does the loss of perceived autonomy matter? *Hellenic Journal of Psychology*, 6(2), 125–144.
- Vieira-Cavalcante, V., Venancio-Dallan, L. P., Pereira-Santana, O., Bertuzzi, R., Tomazini, F., Bishop, D. J., Cristina-Souza, G., & Lima-Silva, A. E. (2023). Effect of different pacing strategies on 4-km cycling time trial performance. *Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas*, 55, e12351. <https://doi.org/10.1590/1414-431X2022e12351>

Warburton, D.E.R., Jamnik, V.K., Bredin, S.S.D., & Gledhill, N. (2011). The Physical Activity Readiness Questionnaire (PAR-Q+) and electronic Physical Activity Readiness Medical Examination (ePARmed-X+). *Health & Fitness Journal of Canada*, 4(2): 3–23.

Xie, H., Chen, Y, & Yin, R. (2020) Running together is better than running alone: a qualitative study of a self-organised distance running group in China, *Leisure Studies*, 39(2), 195-208. <https://doi.org/10.1080/02614367.2019.1698647>

Design Table.

Question	Hypothesis	Sampling plan (e.g., power analysis)	Analysis plan	Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis	Interpretation given to different outcomes	Theory that could be shown wrong by the outcomes
Do mismatches between prospective and retrospective RPE inform on retrospective running pleasure?	Running sessions with a higher retrospective than prospective RPE (i.e., a positive RPE-based prediction error) are associated with a higher level of retrospective pleasure, and vice versa (i.e., negative RPE-	Pilot data suggest an effect size of $R^2 > 0.40$ for the effect of RPE-based prediction error on running pleasure (while taking into account the effect of running distance and average speed of each running session). Our power simulation suggests	LMM analyses on the pilot data revealed that the model with random intercept and fixed slope (step 2 model) is the best model. Indeed, the model with random intercept and random slope (step 3 model) does not result in a better fit.	We determined the relevant effect size for statistical power analysis based on effect sizes obtained in our pilot study (see, <i>Sample size estimation</i> section, for details).	If the prediction error RPE is significantly and positively associated with retrospective running pleasure, we will conclude finding evidence for our hypothesis. We also expect to observe an effect of medium to large size of our step 2	Our study tests whether one main tenet of reward prediction error (i.e., when rewarding outcome is better than expected, it induces more pleasure than a reward that matches a prior expectation) translates to (a self-selected mode of) physical exercise.

	<p>based prediction error).</p>	<p>that 16 participants (each running at least 5 sessions) is required for third step models.</p>			<p>model (conditional $R^2 > .40$). This will lead us to the interpretation that using prospective and retrospective RPE may be beneficial for better identifying sessions of physical exercise that lead to increased (or decreased) experience of pleasure. In the case of nonsignificant effect of RPE prediction errors on running pleasure, this will lead us to discuss how the</p>	<p>A failure to confirm this hypothesis would either question the occurrence of RPE-based prediction errors or indicate that our procedure is not appropriate to test this hypothesis.</p>
--	---------------------------------	---	--	--	---	--

					current design and procedure of the physical exercise program (i.e., “self-selected” running sessions) could be adapted (e.g., standardized running sessions) for observing a significant effect of RPE-based prediction error on running pleasure.	
--	--	--	--	--	---	--