Current Psychology
https://doi.org/10.1007/s12144-021-01718-3

Preference for and tolerance of the intensity of exercise questionnaire (PRETIE-Q): validity, reliability and gender invariance in Portuguese health club exercisers

D. S. Teixeira[1,2] · P. Ekkekakis[3] · A. Andrade[1] · F. Rodrigues[4,5,6] · A. Evmenenko[1] · J. Faria[1] · P. Marques[1] · L. Cid[4,5] · D. Monteiro[5,7]

Accepted: 5 April 2021

Abstract

The aim of the present study was to translate and adapt the Preference for and Tolerance of the Intensity of Exercise Questionnaire to Portuguese (PRETIE-Q-PT), and to analyze its psychometric properties (factorial validity with gender invariance analyses, reliability, temporal stability, and construct validity). The sample was composed of 445 Portuguese participants (55.9% women), with a mean age of 29.6 years (SD = 8.2). To verify the psychometric properties of the scale, structural equation modeling procedures were used. Results showed that an abbreviated 10-item scale presented good fit, reliability, and convergent validity. Latent mean analysis between samples, sex, and time points showed no differences. Construct validity tested with self-reported exercise frequency, vitality, well-being, and habit supported the use of the PRETIE-Q-PT in exercise settings, highlighting the need for exercise professionals to consider these intensity-related trait variables in the promotion of a pleasant exercise experience. As a general conclusion, the PRETIE-Q-PT seems to be a valid scale that can be used to assess intensity-trait characteristics in health club exercisers, and may allow professionals to better adjust exercise prescription to subjective needs aiming to achieve theoretically suggested promotion of pleasurable exercise experiences.

Keywords Preference · Tolerance · Exercise · Intensity · Health clubs

Physical exercise is probably one of the best health promotion tools available (Pedersen and Saltin, 2015; Warburton and Bredin, 2017). However, exercise professionals often find it difficult to maintain a high level of participant adherence and prevent dropout. According to surveys conducted by several organizations, a high proportion of the world population is not sufficiently physically active, which creates an urgent need to improve professional practice and, in particular, how exercise and physical activity are promoted and supported (EC, 2018; PAG, 2018; Stamatakis et al., 2019).

Besides difficulties that emerge in the initiation of exercise programs among the general population, health clubs face another major challenge, namely high rates of program dropout (or, conversely, the low rate of client maintenance). The average dropout rate in the first six months is generally estimated at over 50% (Buckworth et al., 2013; Edmunds et al., 2007; Radel et al., 2017; Sperandei et al., 2016), which emphasizes the need for further exploration of theories and empirical evidence that would enable professionals to address this problem.

Taking into account the factors that have emerged in research related to exercise adherence and dropout, growing evidence supports the need to revise exercise prescription guidelines. These are presently focused solely on exercise doses (i.e., frequency, intensity, duration) that are considered safe and effective (Ekkekakis et al., 2011; Ladwig et al., 2017). As an alternative, a tripartite rationale (Ekkekakis et al., 2011) was proposed, aiming at extending the previous framework by considering the importance of the 'dose' that

✉ D. S. Teixeira
diogo.teixeira@ulusofona.pt

1 Faculty of Physical Education and Sport (ULHT), Lusófona University, Lisbon, Portugal

2 Research Center in Sport, Physical Education, and Exercise and Health (CIDEFES), Lisbon, Portugal

3 Department of Kinesiology, Iowa State University, Ames, IA, USA

4 Sport Science School of Rio Maior (ESDRM-IPSantarém), Rio Maior, Portugal

5 Research Center in Sport Sciences, Health Sciences and Human Development, Vila Real, Portugal

6 Life Quality Research Center (CIEQV), Santarém, Portugal

7 Polytechnic Institute of Leiria, ESECS-IPLeiria, Leiria, Portugal

Published online: 24 April 2021

participants perceive as pleasant and would be willing to repeat in future sessions.

The pursuit of a pleasurable and enjoyable experience has been proposed as an important contributor to exercise intention, persistence, and continuous adherence to exercise programs (Calder et al., 2020; Rhodes and Kates, 2015; Rodrigues et al., 2020; Williams et al., 2008), and aligns with the proposed tripartite model. One of the factors that has an important relation with pleasure and enjoyment is exercise intensity (Ekkekakis et al., 2011). Several studies have reported that increases in exercise intensity have significant consequences for the participants' affective state, with higher intensities being generally related to reduced pleasure or increased displeasure (Evmenenko and Teixeira, 2020; Rhodes and Kates, 2015; Rose and Parfitt, 2007; Williams et al., 2008).

This intensity-pleasure relation exhibits considerable inter-individual variability, especially at mid-range intensities (not too low, not too high). As stated by Ekkekakis et al. (2008), individuals differ greatly in the intensity of exercise that they prefer and can tolerate, and, consequently, in the degree to which exercise can elicit an optimal affective response. Some authors have reported that affective responses to exercise sessions are related to future physical activity (Rhodes and Kates, 2015), and that intervening to enhance the during-exercise affective state is viable, relevant, and may improve individual affective responses in aerobic, resistance, and mixed activities (Evmenenko and Teixeira, 2020). Moreover, in a qualitative study aimed to identify factors that influence affective forecasts for exercise, four essential aspects emerged, namely exercise intensity, exercise outcomes, exercise environment, and enjoyment (Calder et al., 2020), further reinforcing the importance of tailoring intensities to individuals in order to optimize pleasure and enjoyment experiences.

These aspects are key to understanding how exercise intensity is related to affective states. However, in health clubs, exercise professionals need to plan activities and prescribe exercise according to their clients' needs and goals. When developing exercise plans or giving recommendations, the identification of factors that may contribute to program adherence, increased pleasure and enjoyment, and a decrease in the likelihood of dropout seems particularly important in light of current evidence (Ekkekakis et al., 2019; Rodrigues et al., 2020). Therefore, a tailor-made intensity approach may be an important factor contributing to a more effective exercise practice.

The Preference for and Tolerance of the Intensity of Exercise Questionnaire (PRETIE-Q)

The PRETIE-Q is a 16-item questionnaire that aims to assess the traits of preference (predisposition to select a particular level of exercise intensity) for and tolerance (ability to continue exercising at an imposed level of intensity even when the activity is unpleasant/uncomfortable) of the exercise intensity (Ekkekakis et al., 2005). The scale includes two dimensions (i.e., preference and tolerance) with eight items each, answered using a 5-point bipolar Likert scale ranging from "*I totally disagree*" to "*I totally agree*". In the preliminary validation of the scale, after adjustments allowing some model errors to be correlated, the initial measurement model presented a good fit: χ^2 (99) = 128.0, $p < .03$, χ^2/df = 1.29, CFI = .97, GFI = .92, RMSEA = .04, with the two latent factors moderately correlated ($r = .42$). Subsequent psychometric evaluation of the PRETIE-Q in an independent sample of female college students showed that the two-factor structure provided acceptable fit again: χ^2 (99) = 492.7, CFI = .91, GFI = .89, RMSEA = .08 [.07–.09].

An initial psychometric evaluation of the PRETIE-Q in Brazilian-Portuguese was conducted by Smirmaul et al. (2015). Analyses of internal consistency and test-retest reliability showed that the PRETIE-Q is a comprehensible and practical instrument for the population studied (undergraduate students). However, as noted by the authors, further evaluation of the instrument is needed, both in English and in Portuguese, since different sample characteristics may impact the results.

Since there are linguistic differences between Brazilian Portuguese and European Portuguese, including cultural and contextual differences specific to the fitness and health club domain (e.g., *treinando* vs. *a treinar* [training: gerund vs. infinitive]; *academia* vs. *ginásio* [academy vs. gymnasium]), the meaning of certain items could be misinterpreted or ignore culturally idiosyncratic ways of expression (Widenfelt et al., 2005). Additionally, it is not uncommon when different translations/or adaptations of the same questionnaire exist for one language. Several factors may underlie this decision, and may also be dependent on methodological decisions made by researchers in a given field of work (Widenfelt et al., 2005). Thus, additional psychometric work seems to be necessary for the use of the questionnaire with Portuguese exercisers and within the specific cultural and contextual environment of fitness and health clubs.

Present Research

Considering that no relevant tool is currently used in pre-exercise assessment and evaluation, and no specific recommendations are presently offered in the exercise prescription guidelines of any major scientific or professional organization, an explanation of the role of individual differences in preference for and tolerance of exercise intensity within the context of the tripartite rationale is needed. For example, personal trainers have the option of making continuous adjustments based on

assessments of the pleasure/displeasure response throughout an exercise session. However, in group classes, this may be difficult to apply (Vandoni et al., 2016). Still, in order to promote a positive experience, exercise professionals should aim to enhance pleasure and enjoyment perceived by the exerciser (Brooks, 2004). Thus, an individualized approach that includes an initial assessment of intensity preference and tolerance may help professionals identify activities that serve the participants' goals and needs while also ensuring a pleasurable exercise experience (Ekkekakis and Zenko, 2016).

Considering that past experience has an important role in the prediction of future exercise behavior (Calder et al., 2020; Rodrigues, Teixeira, Cid, & Monteiro 2019), it seems relevant that in the crucial first six months, some additional precautions and adaptations in the behavior of exercise professionals may be vital. It has been shown that people tend to underestimate the psychological benefits that exercise can potentially bring based on the initial discomfort that exercise can induce (Miller and Krizan, 2016; Loehr and Baldwin, 2014; Lutz et al., 2008; Ruby et al., 2011; Teixeira and Palmeira, 2016). Thus, individualized approaches focusing on the promotion of pleasure and avoidance of pain and displeasure (i.e., hedonic principle) may represent an additional tool in improving exercise adherence.

Therefore, considering the importance that an instrument designed to assess individual differences in preference for and tolerance of exercise intensity may have in exercise settings (e.g., health clubs), and considering the need for a more extensive evaluation of the questionnaire in the Portuguese language, the aim of this study was to translate, adapt, and psychometrically evaluate the Preference for and Tolerance of the Intensity of Exercise Questionnaire (PRETIE-Q-PT) in a sample of Portuguese health club exercisers. Given previous considerations, a new translation and adaptation was deemed necessary. Accordingly, following the recommendations of Brislin (1970) for cross-cultural research, a new translation and back-translation in contemporary European Portuguese were conducted.

Ekkekakis et al. (2008) suggested that "the appropriate next step would be to (…) examine the factorial invariance across sexes" (p. 509). Consequently, the measurement invariance of the questionnaire was also examined following this recommendation. Taking into account the methodological suggestions regarding psychometric evaluations (Byrne, 2016), measurement invariance was tested between sexes, two independent samples (i.e., calibration and validation), as well as in a third sample with two data collection time points (3 months apart). These procedures allowed us to extend the analysis of the psychometric properties of the scale in distinct samples, providing evidence of reliability in health club exercisers.

Moreover, considering that the constructs of preference and tolerance were conceptualized as traits, scores on the PRETIE-Q are expected to demonstrate temporal stability (Ekkekakis et al., 2005). Therefore, the data collection time interval used in the third sample resembled the one in the original Ekkekakis et al., (2005) sample (i.e., 3 months). Taking into account that measurement invariance may only reflect the way distinct groups interpret the scale, domains, differences between sexes, samples, and data collection points were tested with latent mean analysis. Ensuring that measurement invariance criteria are met, this analysis may add some information about sample characteristics and exercise settings that should be taken into account when using the PRETIE-Q (Byrne 2016).

Finally, to test construct validity, the associations between the PRETIE-Q-PT and behavioral (weekly exercise frequency), cognitive (habit), and emotional (vitality, positive well-being, psychological distress, fatigue) variables were tested. The two constructs are generally moderately correlated, but previous psychometric testing presented construct discriminant validity (e.g., Ekkekakis et al., 2008). Thus, some shared variance between constructs is expected, and preference and tolerance may present common and distinct associations with several relevant variables. For example, in the study of Ekkekakis et al. (2006), preference, but not tolerance, was a significant predictor of self-selected exercise intensity in a 20-min exercise session. In the work of Ekkekakis et al. (2007), the tolerance scale was hypothesized to be associated with persistence in exercise in incremental tests after reaching the ventilatory threshold, and the two tested samples presented support for this hypothesis. A study conducted by Lochbaum et al. (2009) showed that performance in a shuttle run test had positive associations with preference and tolerance for men, and with tolerance for women. In another example, Hall et al. (2014) showed that both preference and tolerance displayed significant associations with performance in fitness tests (e.g., push-ups, sit-ups), body composition, and physical activity. Given the theoretical framework of the PRETIE-Q, there is support for the hypothesis that preference and tolerance may have distinct associations with several behavioral, cognitive, and emotional variables, and are context/activity-dependent. To the best of our knowledge, the implications of preference and tolerance in the health-club context have not yet been explored, reinforcing the need for further construct validation research in this context.

Additionally, in this particular setting, when exercise prescription takes into account exercisers' individual characteristics and when exercise is adjusted to their personal goals and abilities, several important motivational and behavioral variables are expected (Garber et al., 2011; Ryan and Deci, 2017; ACSM, 2018). Accordingly, a practice in which the individual perceives that exercise intensity is adjusted to his or her own preference and tolerance, should promote positive affective, cognitive, and behavioral responses.

In sum, given previous theoretical and empirical considerations regarding low levels of physical activity participation, health clubs exercise adherence and dropout issues, and the need to add evidence regarding the promotion of better

affective responses in exercise dynamics, the present study sought to translate, adapt, and psychometrically evaluate the PRETIE-Q in health club exercisers. The scale presents itself as a possible tool to help professionals to better adjust exercise prescription, aiming to obtain a better affective response related to individual exercise intensity, thus adding to the possible factors associated to program adherence, as is the case of a pleasurable exercise experience.

Method

Participants

Data were collected from three independent samples. The first sample comprised a total of 174 gym exercisers (75 men; 99 women) aged between 18 and 61 years ($M = 31.25$; $SD = 6.5$). Data from this sample were collected using an online survey. These participants were involved in several fitness activities (individual and group) in gyms and health clubs. This sample was considered as the calibration sample and was used to test the initial factor structure consisting of 16 items and two correlated factors.

Data from the second sample were also collected using an online survey. The sample comprised 146 (50 men; 96 women) gym exercisers aged between 19 and 66 years ($M = 29.87$; $SD = 9.5$). This sample also included exercisers enrolled in health club activities (individual and group). These participants were considered a cross-validation sample that was used to test whether the measurement model would demonstrate invariance compared to the calibration sample.

Finally, the third sample was composed of 125 gym exercisers (71 men; 54 women) aged 18 to 55 years ($M = 26.38$; $SD = 9.95$). In-person data collection was carried out individually in three health clubs in the central region of Portugal. Exercisers were enrolled in several fitness activities provided by health clubs (individual, group, and water activities). In order to test for longitudinal invariance and test-retest reliability, data were collected twice from this sample with a three-month interval. There was 12% attrition between measurements.

Procedures

Data Collection

For the two online samples, an invitation to complete an online battery of psychometric questionnaires was distributed through social media and via mailing lists from January to March of 2020. Data were collected anonymously through Google Forms from anyone willing to participate. Face-to-face data collection with the third sample was conducted over a period of two months. Authorizations from club managers were obtained beforehand. After reading a letter of explanation, signed informed consent was provided prior to completing the questionnaires.

To be included, exercisers had to be ≥18 years old, speak and understand the Portuguese language, have a minimum experience of three months in the health clubs, and be enrolled in health club activities. Responses were screened to verify whether the inclusion criteria were met, and two individuals had to be excluded because they were underage. The time of completion was approximately 10 min for all data collection procedures.

Translation of the PRETIE-Q

The translation of the PRETIE-Q from English to Portuguese was done through the committee approach methodology (Brislin, 1980). The process includes five steps, namely:

1) Preliminary Translation: This first stage was carried out by the researchers, and a translation of the original questionnaire was completed, resulting in the first translated version;
2) First Evaluation Panel: An analysis of the initial version of the PRETIE-Q was carried out individually by three specialists from different areas, including a Ph.D. in sport sciences and two Ph.Ds. in exercise psychology, who were experienced researchers with knowledge of psychometrics;
3) Second Evaluation Panel: A revised version of the questionnaire was sent for a new evaluation to another panel formed by another three specialists in the same areas as the previous ones. After that, the panel examined all the items of the questionnaire and made suggestions for small changes, which were promptly accepted, resulting in a new version;
4) The Pilot Study: The third version of the questionnaire was administered to 36 bilingual college students with exercise experience so as to test English/Portuguese syntax and comprehension, resulting in the fourth version;
5) Final revision, in which two Portuguese language teachers reviewed the fourth version for syntax, spelling, and grammar, and produced the fifth and the final version.

Instruments

The Preference for and Tolerance of the Intensity of Exercise Questionnaire (PRETIE-Q)

Following the structure of the original English version of the PRETIE-Q (Ekkekakis et al., 2005), the Portuguese version

(PRETIE-Q-PT) comprised 16 items representing the two scales (eight items per scale) that corresponded to the intensity-preference (e.g., "The faster and harder the workout, the more pleasant I feel") and intensity-tolerance (e.g., "Feeling tired during exercise is my signal to slow down or stop") traits. Half of the items in each scale are reversed-scored. The stem asks respondents what best describes what they believe and how they feel when exercising, and the answers are given on a 5-point bipolar Likert scale, ranging from 1 ("Totally disagree") to 5 ("Totally agree"). For example, "*Feeling tired during exercise is my signal to slow down or stop*" is an example of a tolerance item, and "*Low-intensity exercise is boring*" is an example of a preference item.

For descriptive purposes, two additional questions were developed to assess the perception of the exercisers regarding their current training intensity and their preference and tolerance. The questions ("The intensity of my training is in accordance with my preference"; "The intensity of my training is in accordance with my tolerance") were answered and coded with 0 (not in agreement) or 1 (in agreement).

Subjective Vitality Scale (SVS)

This 6-item scale, developed by Ryan and Frederick (1997; Portuguese version by Moutão, Alves & Cid, 2013), assesses subjective vitality (the state of feeling alive/alert, and having energy available to the self) in the general population. Responses are given using a 7-point bipolar Likert scale, ranging from 1 ("Totally disagree") to 7 ("Totally agree") and reflect the participants' global perception of subjective vitality. The score is calculated by averaging the values from each item. The scale has been used extensively in physical activity-related studies, demonstrating adequate reliability (e.g., Ryan and Deci, 2001). Internal consistency in the present study was acceptable (Cronbach's alpha = .88).

Subjective Exercise Experiences Scale (SEES)

The SEES (McAuley & Courneya, 1994; Portuguese version by Palmeira, 2006) is a 12-item questionnaire that measures three domains of affective experience in exercise settings, namely: Positive Well-Being (4 items; e.g., great, strong); Psychological Distress (4 items; e.g., awful, crummy), and Fatigue (4 items; e.g., tired, fatigued). Responses were provided using a 7-point scale, ranging from 1 ("Not at all") to 7 ("Very much so"). The stem asks participants to indicate how they feel "now, at this point in time". The original SEES instrument seems to be sensitive to exercise intensity and is supported in young and middle-age adults. In the present sample, Cronbach's alpha was .87 (well-being), .75 (distress), and .64 (fatigue).

Self-Report Behavioral Automaticity Index (SRBAI)

The SRBAI (Gardner et al., 2012) is a 4-item scale developed to assess behavioral habit development (acquired behavioral patterns dependent on learned context-behavior associations). Since the SRBAI has not been translated nor adapted to the exercise context, a Portuguese version was developed for this study following the translation procedures recommended by Brislin (1970, 1980), with further psychometric testing being underway. The statement "*Exercise is something*" preceded the four items ("*I do automatically*", "*I do without having to consciously remember", "I do without thinking", "I start doing before I realize I am doing it*"), and participants rated how true each statement was for them on a 7-point bipolar Likert scale, ranging from 1 ("Totally disagree") to 7 ("Totally agree"). In the present sample, the internal consistency coefficient was acceptable (Cronbach's alpha .89).

Data Analyses

Descriptive statistics, including mean, standard deviation, as well as bivariate correlations, were calculated for the variables studied. Cronbach's coefficients were calculated to determine internal consistency reliability of participants' responses on the PRETIE-Q-PT. For test-retest analysis, the third sample of 125 Portuguese exercisers was used. The interval among survey administration was three months, consistent with the method used in the original psychometric evaluation.

A confirmatory factor analysis (CFA) using maximum likelihood estimation was performed with AMOS v. 23 to examine the model fit of the PRETIE-Q-PT. The CFA was performed following standard recommendations (e.g., Byrne, 2016; Hair et al., 2019; Marsh et al., 2004). To determine model fit, the following indices were considered: Comparative Fit Index (CFI); Tucker-Lewis Index (TLI); Standardized Root Mean Square Residual (SRMR), and Root Mean Square Error of Approximation (RMSEA), with a 90% Confidence Interval (CI 90%). For these indices, the following cutoffs were considered: CFI and TLI $\geq .90$; SRMR and RMSEA $\leq .08$ (Byrne, 2016; Hair et al., 2019; Kline, 2016; Marsh et al., 2004).

To evaluate convergent validity, Average Variance Extracted (AVE) was calculated considering scores $\geq .50$ as acceptable (Hair et al., 2019). Discriminant validity was established when the AVE score for each factor exceeded the squared correlations between the two factors (Hair et al., 2019). Internal consistency was examined considering composite reliability coefficients calculated by using the Raykov (1997) formula and accepting values $\geq .70$ (Raykov et al., 2015).

Multigroup Analysis

The analyses of measurement invariance between groups (i.e., calibration sample and validation sample; between gender) and longitudinally (i.e., time 1 vs. time 2 in the third sample) were based on standard recommendations (Byrne, 2016; Chen, 2007; Cheung and Rensvold, 2002). Firstly, we examined whether the measurement model presented a good fit to the data for each group. Secondly, for invariance testing, model re-specification between groups examined whether (i) the same number of factors was present in both groups and remained associated with the same items (configural invariance); (ii) the model factors had similar factor loadings in both groups (metric invariance); (iii) latent and observable means were valid in both groups when means were compared (scalar invariance); and (iv) comparison between observable items were supported (residual invariance). The following criteria were used to accept measurement invariance between groups: configural invariance with ΔCFI less than .01 (Cheung and Rensvold, 2002); metric invariance with ΔSRMR less than .03 and ΔRMSEA less than .015; scalar invariance with ΔSRMR less than .010, and ΔRMSEA less than .015 (Chen, 2007).

Additionally, if invariance criteria were met, latent mean analysis procedures were followed. Mean and covariance structure analyses for the preference and tolerance factors allow for the test of groups differences. The male, calibration, and time 1 samples were constrained to zero, while the female, validation, and time 2 samples, respectively, were freely estimated using Z scores and p values for preference and tolerance. According to Kline's (2016) recommendations and Cohen's (1988) d criteria, effect sizes were calculated and were characterized as (a) trivial (0–.19), (b) small (.20), (c) medium (.50), or (d) large (.80).

Finally, the concurrent validity of the PRETIE-Q-PT was tested by performing a correlational analysis between the number of exercise sessions per week in the health club, SRBAI (habit development), SVS (vitality) and SEES (positive well-being, psychological distress, and fatigue).

Results

In the preliminary analysis, and after initial data screening, no missing values were detected. Seven univariate ($z > 3.00$) and multivariate (Mahalanobis distance = $p1 < .001$; $p2 < .001$) outliers were detected and removed from the data before further analysis, according to Byrne's (2016) recommendations. As seen in Table 1, participants used the full range of the scale (1–5) (except for item 12 in the calibration sample), presenting similar values for preference and tolerance items. The skewness and kurtosis values ranged between −2 to +2 and − 7 to 7, respectively, and can be considered as approximating a normal distribution (Gravetter and Wallnau, 2014). For detailed information, see Table 1. Regarding multivariate kurtosis, Mardia's coefficient was >5.0, exceeding a common cut-off value for assuming multivariate normality, and, to account for this, a Bollen-Stine bootstrap with 2000 samples was used for subsequent analysis (Byrne, 2016; Nevitt and Hancock, 2001).

The test-retest reliability results obtained with the third sample are depicted in Table 2. According to the results, all correlation coefficients were above .70, suggesting adequate test-retest reliability. Additionally, alpha coefficients of internal consistency were above .70 for both factors at both time points (3 months apart).

The psychometric properties of the measurement model are reported in Table 3. The initial calibration model did not provide acceptable fit to the data. Thus, factor loadings and cross-loadings were examined in accordance with recommendations (Hair et al., 2019). As a result, items 6, 10, and 16 (Preference), and 5, 7 and 11 (Tolerance) were removed due to low factor loadings (<.40) or high cross-loadings (>.15). A final calibration model with 10 items (5 items per factor) was found to have acceptable fit to the data, and subsequent analyses were based on this model. In addition, all group samples (i.e., calibration, validation, female, male, and two time points) also presented acceptable model fit, in line with the predefined cutoff values.

Convergent and discriminant validity, as well as reliability coefficients, are reported in Table 4. Regarding convergent validity, the AVE scores did not meet the recommended cut-off values (all observed values < .50) in the calibration and validation samples. However, in both models, all factor loadings (λ) were above .50, $p < .05$ (except item 14 in the calibration model, $\lambda = .44$, $p < .05$). In the discriminant validity analysis, a few issues appeared only in the calibration model, where the squared value of the correlations (r^2) were higher than the AVE. The composite reliability coefficients met the criteria for internal consistency.

Since the measurement model provided acceptable fit to the data from the calibration and validation groups, sex and the two time-points (see Table 3), multigroup invariance procedures were followed. Multigroup model comparison was based on the assumption that differences in CFI, SRMR, and RMSEA should conform to the proposed cutoffs. This was verified for all criteria except residual invariance in the time 1 vs. time 2 analysis (ΔCFI = .01). Additionally, ΔSRMR and ΔRMSEA between models were met in all the tested samples, as seen in Table 5.

Given that the model displayed evidence of invariance across groups, latent mean analysis was conducted to further understand possible group differences among latent means. As seen in Table 6, no significant differences in latent means between groups were found.

Finally, to investigate construct validity, correlations of the PRETIE-Q-PT with behavioral (number of exercise sessions per week), cognitive (habit), and affective (vitality, positive

Table 1 Descriptive analysis of the answered items on the PRETIE-Q (calibration, validation, and longitudinal samples)

	Calibration							Validation						Longitudinal (1st assesment)					
	Min-Max	M	SD	Skewness	z-value	Kurtosis	z-value	M	SD	Skewness	z-value	Kurtosis	z-value	M	SD	Skewness	z-value	Kurtosis	z-value
Item 1 (Tolerance)	1–5	3,28	1,20	–,16	–,86	–,99	–2,66	3,31	1,22	–,27	–1,30	–,84	–1,99	5,28	1,30	–,30	–1,31	–,93	–2,05
Item 2 (Preference)	1–5	3,56	1,20	–,55	–2,96	–,56	–1,50	3,64	1,23	–,59	–2,80	–,68	–1,59	4,57	1,42	–,46	–2,01	–,83	–1,82
Item 3 (Tolerance)	1–5	2,76	1,17	,18	,99	–,82	–2,19	2,62	1,24	,30	1,41	–,98	–2,31	4,31	1,44	–,40	–1,75	–,56	–1,23
Item 4 (Preference)	1–5	3,32	1,30	–,28	–1,51	–1,02	–2,74	3,40	1,32	–,26	–1,25	–1,14	–2,69	3,53	1,53	–1,19	–5,22	,78	1,71
Item 5 (Tolerance)	1–5	3,23	1,19	–,43	–2,32	–,66	–1,77	3,35	1,25	–,55	–2,58	–,81	–1,91	3,32	1,25	,04	,17	–,99	–2,18
Item 6 (Preference)	1–5	3,80	1,15	–,74	–3,97	–,16	–,42	3,83	1,20	–,88	–4,14	–,09	–,22	3,71	1,20	–,55	–2,41	–,29	–,64
Item 7 (Tolerance)	1–5	3,06	1,07	–,28	–1,52	–,51	–1,36	2,90	1,10	,06	,26	–,67	–1,58	2,75	1,21	–,52	–2,29	–,32	–,71
Item 8 ((Preference)	1–5	3,74	1,12	–,64	–3,44	–,47	–1,25	3,63	1,17	–,52	–2,47	–,60	–1,41	3,37	1,35	–,30	–1,34	–1,07	–2,35
Item 9 (Tolerance)	1–5	3,84	1,09	–,81	–4,34	–,07	–,20	3,80	,98	–,51	–2,38	–,50	–1,19	3,29	1,16	–,72	–3,17	–,35	–0,77
Item 10 (Preference)	1–5	3,39	1,25	–,25	–1,35	–,99	–2,68	3,41	1,32	–,31	–1,47	–1,06	–2,49	3,95	1,21	,17	,73	–,91	–2,00
Item 11 ((Tolerance)	1–5	3,27	1,04	–,34	–1,84	–,36	–,97	3,24	1,21	–,16	–,75	–,95	–2,24	2,97	1,04	–,23	–1,01	–1,18	–2,59
Item 12 (Preference)	2–5	4,32	,84	–1,43	–7,68	2,31	6,21	4,28	,88	–,98	–4,64	–,03	–,07	3,87	1,10	–,59	–2,58	–,40	–,89
Item 13 (Tolerance)	1–5	3,35	1,06	,13	,71	–1,10	–2,96	3,22	1,21	–,06	–,30	–1,08	–2,56	3,97	1,02	–1,00	–4,40	,13	,29
Item 14 (Preference)	1–5	3,48	1,05	–,55	–2,97	–,09	–,24	3,40	1,16	–,32	–1,52	–,71	–1,67	3,53	1,25	–,18	–,79	–,40	–,87
Item 15 (Tolerance)	1–5	3,47	1,09	–,54	–2,93	–,47	–1,26	3,49	1,02	–,45	–2,10	–,48	–1,14	3,20	1,11	–,85	–3,72	–,05	–,10
Item 16 (Preference)	1–5	3,32	1,30	–,31	–1,69	–1,01	–2,71	3,43	1,33	–,33	–1,57	–1,10	–2,60	4,45	,74	–1,01	–4,46	,54	1,18

Note. Min – Minimum; Max = Maximum; M = Mean; SD = Standard Deviation

well-being, psychological distress, fatigue) variables were examined. As reported in Table 7, in the total sample, preference and tolerance were positively associated with the number of exercise sessions per week, habit, vitality, and positive well-being. Additionally, preference demonstrated a weak but statistically significant positive association with fatigue. In the sample that reported that their exercise regimen was not in accordance with their intensity preference or tolerance, the majority of the significant results were not replicated. The exceptions were the associations of preference with the number of sessions/week and habit.

Discussion

The present study aimed to examine the validity and reliability of the PRETIE-Q-PT in a sample of Portuguese exercisers in three distinct samples. Additionally, we explored whether the factor structure of the scale would exhibit invariance across samples with different characteristics. Lastly, we examined the construct validity of preference and tolerance measures, analyzing the associations of these factors with behavioral, cognitive, and affective variables. The results of the study show that the abbreviated 10-item scale presented a good fit, adequate reliability and construct validity. Additionally, latent mean analysis did not detect any differences between samples, sex, and time points. The results suggest that the PRETIE-Q-PT is a valid scale that can be used in exercise settings to assess intensity-trait characteristics, adding possibilities for professionals to adjust and refine their exercise prescriptions.

According to classical test theory (Crocker and Algina, 1986), test scores consist of two hypothetical components, a "true score" (presumed to be error-free) and a component that reflects random measurement error (such as the kinds of errors respondents may make due to carelessness, fatigue, or inconsistent instructions, explanations, or testing conditions). In the context of cross-cultural adaptations of self-report measures, the likelihood of random measurement error may be increased due to such factors as lack of clarity in the instructions or the use of ambiguous or polysemous terms in translated items. Reliability can be conceptualized as the extent to which scores on a psychometric scale reflect the hypothetical "true score" and are, therefore, unaffected by random measurement error (or, said differently, reliability is the ratio of true-score variance to observed-score variance). Considering that neither true-score variance nor error variance can be measured directly, they must be estimated indirectly via such methods as test-retest or calculating the internal consistency of test scores.

Reliability is a fundamentally important psychometric property because it has far-reaching implications for research. The susceptibility of test scores to sources of random measurement error lowers the precision of estimates of populations values (broadening confidence intervals) and biases

Table 2 Test-retest reliability analysis

Items	M±SD	*r*	*p*	α
Item 1 Pre-Post	3.31±1.36 – 3.19±1.26	.82	<.001	–
Item 2 Pre-Post	3.44±1.36 – 3.07±1.27	.77	<.001	–
Item 3 Pre-Post	2.59±1.06 – 2.54±1.09	.89	<.001	–
Item 4 Pre-Post	2.96±1.33 – 2.90±1.26	.83	<.001	–
Item 5 Pre-Post	3.22±1.22 – 2.91±1.25	.79	<.001	–
Item 6 Pre-Post	3.56±1.51 – 3.37±1.46	.88	<.001	–
Item 7 Pre-Post	3.30±1.02 – 3.22±1.04	.83	<.001	–
Item 8 Pre-Post	3.56±1.18 – 3.30±1.16	.80	<.001	–
Item 9 Pre-Post	3.72±.980 – 3.44±1.02	.77	<.001	–
Item 10 Pre-Post	3.39±1.38 – 3.29±1.18	.81	<.001	–
Item 11 Pre-Post	3.11±1.04 – 2.78±.710	.72	<.001	–
Item 12 Pre-Post	4.37±.850 – 3.94±.940	.78	<.001	–
Item 13 Pre-Post	3.30±.960 – 3.11±.960	.88	<.001	–
Item 14 Pre-Post	3.24±1.34 – 2.94±1.31	.75	<.001	–
Item 15 Pre-Post	3.24±1.03 – 2.98±1.04	.78	<.001	–
Item 16 Pre-Post	3.39±1.31 – 3.11±1.37	.81	<.001	–
Preference	17.57±4.64 – 16.17±4.33	.87	<.001	.814 – .772
Tolerance	16.17±3.76 – 15.26±3.66	.91	<.001	.727 – .706

Note. M = Mean; SD = Standard Deviation; r = Correlation; *p* = Sig. value; α = Cronbach's Alpha

statistical tests toward the null (underestimating effect sizes), thus reducing statistical power. In other words, unreliable test scores result in attenuated estimates of both statistical associations (i.e., correlations) and experimental effects (Baugh, 2002; Muchinsky, 1996; Schmidt and Hunter, 1996). The consequences of this problem could be an increased rate of Type 2 errors (i.e., failing to reject then null hypothesis when it is false at the level of the population) or having to increase the required sample sizes in order to compensate for the loss of statistical power. For example, in the case of the PRETIE-Q, low reliability of test scores could result in reduced estimates of the extent to which preference and tolerance can predict affective responses or physical activity participation and adherence. Moreover, reliability sets an upper boundary for validity; because only the true-score part of the variance can be valid, test scores cannot be any more valid than they are reliable.

Validity is another fundamental psychometric property that describes an evaluative judgment, derived from a combination of empirical evidence and theoretical reasoning, about the adequacy and appropriateness of inferences and interpretations based on test scores (Messick, 1988). Factorial or structural validity refers to the extent to which test scores faithfully reflect the theorized (unidimensional or multidimensional) structure of a construct. To illustrate the importance of factorial validity for the interpretation of test scores, consider that the item "I feel tense" is scored as part of a unidimensional questionnaire that measures state anxiety. When respondents score this item in the context of a session of psychotherapy, it is reasonable to infer that increases in reported tension would likely signify increases in state anxiety. However, when the same item is scored during a session of strenuous exercise, heightened feelings of tension would likely reflect increased level of somatic stress and autonomic nervous system activity rather than state anxiety per se. Therefore, interpreting test scores as reflecting a unidimensional construct (a unitary "state anxiety"), would result in serious inferential errors, such as concluding that strenuous exercise increases state anxiety (Rejeski et al., 1991). Similar inferential errors can occur when measurement fails to properly reflect the distinction between intensity-preference and intensity-tolerance. For example, researchers may interpret the lack of a pronounced decline in pleasure in response to high-intensity exercise as an indication that participants "like" or "prefer" this type of exercise, when this pattern of results might reflect a high level of tolerance. To avoid such interpretational pitfalls, it is crucial to demonstrate that a measure such as the PRETIE-Q can distinguish between items intended to assess intensity-preference from those intended to assess intensity-tolerance - and do so invariably across different population segments (e.g., men and women, young and old).

Correlations between responses to each item at the first and second administration of the questionnaire were acceptable. Additionally, alpha coefficients were > .70, suggesting acceptable internal consistency as proposed by Raykov (1997). Thus, the items within the PRETIE-Q-PT had a high degree of temporal stability and reliability, and are in line with previous psychometric testing (Ekkekakis et al., 2005), thus suggesting that this questionnaire may be applied, for example, in an usual evaluation moment (e.g., fitness physical assessment) aiming to determine the individual profile for exercise intensity, without the need for regular and new assessment applications (at least in the three months period). These temporal stability indicators allowed us to proceed with the testing of the measurement model.

The original 16-item measurement model did not present an acceptable fit in the calibration sample. Modification indices and factor loadings were examined in order to improve fit. This led to the removal of items 6, 10, and 16 (Preference), and 5, 7 and 11 (Tolerance) due to low factor loadings (<.40) or high cross-loadings (>.15), following the recommendations of Hair et al. (2019). A CFA of the resultant 10-item model (5 items per factor) showed good fit for the two-factor solution assessing preference and tolerance. A second CFA was subsequently performed on data obtained from an independent (validation) sample. The revised 10-item measurement model again was found to provide acceptable fit to the data.

The AVE scores indicated that the convergent validity of tolerance and preference was somewhat below the suggested cut-off values (< .50) in the calibration and validation samples.

Table 3 Goodness-of-fit indices of PRETIE-Q models (including original version)

Models	N	χ^2	df	B-S *p*	SRMR	TLI	CFI	RMSEA [90% CI]
Initial calibration model	174	300.53	103	<.001	.100	.694	.737	.105 [.092–.119]
Final calibration model	174	77.43	34	<.001	.066	.902	.916	.082 [.061–.111]
Final validation model	146	60.07	34	.004	.068	.917	.937	.076 [.043–.107]
Male model	196	63.77	34	.001	.074	.901	.924	.084 [.052–.118]
Female model	249	79.05	34	<.001	.068	.900	.917	.080 [.060–.109]
Time 1	125	56.73	34	.009	.070	.902	.919	.076 [.039–.110]
Time 2	125	60.73	34	.003	.069	.901	.918	.080 [.046–.113]
Ekkekakis et al. model[1]	184	216.3	103	<.001	–	–	.872	.078 [–]
Ekkekakis et al. model[1]	184	128.0	99	.030	–	–	.967	.040 [–]

Note. Initial model = 16 items; Final models = 10 items; χ^2 = chi-squared; df = degrees of freedom; B-Sp = Bollen-Stine *p* value; SRMR = Standardized Root Mean Square Residual; TLI = Tucker-Lewis Index; CFI = Comparative Fit Index; RMSEA = Root Mean Squared Error of Approximation; 90% CI = Confidence Interval of RMSEA; [1] Ekkekakis et al. (2005) initial model with uncorrelated errors; [2] Ekkekakis et al. (2005) initial model allowing correlated errors

However, in both samples, factor loadings (λ) were above .50 (except item 14) and all items loaded significantly on their respective factor (Hair et al., 2019). Moreover, no significant cross-loadings were detected, suggesting acceptable validity (Byrne, 2016).

Regarding discriminant validity, the correlation between preference and tolerance in the calibration sample was somewhat above the AVE scores. In contrast, in the validation sample, the correlation between the scales was below the AVE score, indicating satisfactory discriminant validity as suggested by Hair et al. (2019). Concerning the lack of discriminant validity in the calibration sample, similar results were found in the study conducted by Ekkekakis et al. (2008). Thus, these results suggest that the preference and tolerance dimensions of the PRETIE-Q partially overlap with each other, in accordance with the theoretical postulate proposed by Ekkekakis et al. (2005). However, this was only seen in the calibration sample and the difference among AVE and the squared correlation between constructs was somewhat low (Δ0.03). In line with recommendations by Hair et al. (2019), factor loadings and cross-loadings should be revisited and items displaying several issues should be considered for elimination. The item 14 ("The faster and harder the workout, the more pleasant I feel", part of the preference scale) did present a factor loading below the .50 guideline and should, therefore, be considered for elimination. However, this suggestion by Hair et al. (2019) should be viewed as a rule of thumb when examining the psychometric properties of a measure. Other considerations are that item 14 enhances content validity, removing it would not improve model fit, and the overall 10-item model that includes item 14 exhibits good fit. Therefore, we retained item 14. General recommendations used for addressing items removal in CFA can be obtained in other works (e.g., Byrne, 2016; Kline, 2016; Hair et al., 2019). Although our results with Portuguese exercisers did not support the original 16-item model proposed by Ekkekakis et al., (2005) based on American respondents, the results show adequate fit for the adapted 10-item model (Byrne, 2016; Hair et al., 2019; Marsh et al., 2004) in both the calibration and validation samples.

Table 4 Internal reliability, convergent and discriminant validity, and average variance extracted for the 10-item, two-factor model

Factors (calibration)	CR	AVE	Preference	Tolerance
Preference	.78	.43	1	–
Tolerance	.77	.41	.46*	1
Factors (validation)	**CR**	**AVE**	**Preference**	**Tolerance**
Preference	.81	.47	1	–
Tolerance	.76	.41	.40*	1

Note. CR = Composite Reliability; AVE = Average Variance Extracted; *r^2

With respect to the internal consistency analysis, the results of the present study showed that both factors yielded acceptable coefficients for both samples (Hair et al., 2019; Raykov, 1997). These results are similar to those reported in the original development sample (Ekkekakis et al., 2005) and the subsequent validation sample with female college students (Ekkekakis et al., 2008). In general, our results demonstrated acceptable fit (Byrne, 2016; Hair et al., 2019), as well as adequate convergent and discriminant validity (Hair et al., 2019).

The results also revealed that the measurement model provided adequate fit to the data in all samples (i.e., calibration sample, validation sample, male, female, and sample 3 at two time points: Time 1 and Time 2, three months apart). Looking at each criterion of invariance, the multigroup analysis across groups showed that the factor structure of the PRETIE-Q-PT demonstrated configural invariance. Additionally, metric

Table 5 Criteria for measurement invariance in comparisons across samples, gender, and the two time-points*

	χ^2	df	$\Delta\chi^2$	Δdf	*p*	CFI	ΔCFI	SRMR	ΔSRMR	RMSEA	ΔRMSEA
Calibration-Validation											
Configural Invariance	137.500	42	–	–	–	.925	–	.066	–	.058	–
Metric Invariance	141.929	34	4.429	8	<.001	.929	.004	.067	.001	.053	.005
Scalar Invariance	143.036	31	5.536	11	<.001	.931	.006	.068	.002	.051	.007
Residual Invariance	158.365	21	20.865	21	<.001	.925	<.001	.067	.001	.050	.008
Male – Female											
Configural Invariance	142.842	68	–	–	–	.920	–	.074	–	.060	–
Metric Invariance	150.808	76	7.966	8	<.001	.920	<.001	.077	.003	.057	.003
Scalar Invariance	151.575	79	8.733	11	<.001	.922	.002	.079	.005	.055	.005
Residual Invariance	163.682	89	20.840	21	<.001	.920	<.001	.078	.004	.052	.008
Time 1 vs. Time 2											
Configural Invariance	117.460	68	–	–	–	.919	–	.074	–	.056	–
Metric Invariance	131.152	76	13.692	8	<.001	.917	.002	.079	.005	.055	.001
Scalar Invariance	131.628	79	14.168	11	<.001	.913	.006	.080	.006	.053	.003
Residual Invariance	139.429	89	21.969	21	.001	.909	.010	.081	.008	.049	.007

Note. M = Male; F = Female; χ^2 = chi-squared; df = degrees of freedom; $\Delta\chi^2$ = differences in the value of chi-squared; Δdf = differences in degrees of freedom; CFI = Comparative Fit Index; ΔCFI = differences in the value of the Comparative Fit Index; * testing between the validation and time-point 1 samples presented model invariance

invariance was also achieved, since the factor loadings were found to be invariant between groups. The criteria for scalar and residual invariance were also satisfied, since item thresholds and item residuals, respectively, were found to be invariant between groups, as suggested by several authors (Byrne, 2016; Chen, 2007; Hair et al., 2019; Sass, 2011). Hence, all invariance criteria were met, further supporting the validity of the structure of the measure across groups with different characteristics.

The present study found no significant differences in terms of latent means between samples in the PRETIE-Q-PT. Specifically, no differences were found in latent means between the calibration and validation samples, between sexes, and between two time points in the third independent sample. This means that average scores of preference and tolerance are comparable across groups, independently of group characteristics, reinforcing once again that the PRETIE-Q-PT can be used in both sexes, aligning with the health clubs common users' socio-demographic characteristics.

Table 6 Latent mean differences between samples, gender (validation sample) and time-points on constructs

	Difference	Z	*p*	*d*
Samples				
Preference	−.004	−.030	.976	−.01
Tolerance	−.065	−.604	.546	−.09
Gender				
Preference	−.037	−2.98	.766	−.05
Tolerance	−.103	−.944	.345	−.15
Time-points				
Preference	−.194	−1.56	.120	−.30
Tolerance	−.111	−.93	.352	−.17

Note. Z = test score; *p* = sig. Value; *d* = Cohens' effect size

The final analysis intended to test the construct validity of the PRETIE-Q-PT against several variables related to exercise. The correlations of the preference and tolerance scales with several behavioral variables have been examined in other contexts (e.g., Hall et al., 2014), particularly with exercise frequency in Ekkekakis et al. (2005, 2007, 2008), showing that both preference and tolerance were positively correlated with the number of exercise days per week. Additionally, several studies have reported that exercise is positively related with psychological well-being (McAuley et al., 2000; Sjögren et al., 2006) in intensity-dependent fashion (Teixeira and Palmeira, 2016), and may induce improvements in vitality (Couto et al., 2017; Ju, 2017; Park et al., 2019). Moreover, studies have highlighted the importance of habit in exercise settings (Garber et al., 2011; Kaushal and Rhodes, 2015), aligning with the hedonic principle on which the PRETIE-Q is grounded (Ekkekakis et al., 2019).

Results in the total sample indicate that the preference and tolerance scales exhibited significant, albeit generally weak, positive associations with exercise frequency, habit, vitality, psychological well-being, psychological distress, and fatigue (only for preference). Ekkekakis et al. (2005) reported similar associations with exercise frequency (e.g., .18 to .33 for Preference, .22 to .39 for Tolerance). Similarly, Hall et al. (2014) found that both preference and tolerance correlated significantly but modestly (r = .29 for both) with physical activity. Ekkekakis et al. (2008) found that the weekly frequency

Table 7 Correlational analysis of PRETIE-Q-PT

	Sessions/week (frequency)	Habit	Vitality	Well-Being	Distress	Fatigue
Total sample (N=445)						
Preference	.134*	.176***	.119*	.146**	−.053	.095*
Tolerance	.229***	.185***	.166**	.167**	−.034	.085
Preference-discrepant subsample (N=62)						
Preference	.191	.113	.171	.121	−.027	.011
Tolerance	.504**	.385*	.140	.210	−.038	.002
Tolerance-discrepant subsample (N=58)						
Preference	.091	.187	.130	.112	−.036	.123
Tolerance	.267	−.051	.127	.145	−.021	.021

Note. * $p < .05$; ** $p < .01$; *** $p < .001$

of strenuous exercise was significantly correlated with preference ($r = .30$) and tolerance ($r = .36$), whereas the frequency of moderate or mild exercise did not. Along similar lines, in the total sample in the present study, we found significant associations of exercise frequency with both preference ($r = .13$) and tolerance ($r = .23$). A unique and interesting finding of the present study was that among the subsample of participants who reported a discrepancy between their preferred and actual exercise intensity, tolerance exhibited a higher correlation ($r = .50$) with exercise frequency, whereas no associations were found in the tolerance-discrepant subsample. These results underscore the role that tolerance may have in exercise frequency, especially for individuals who perceive a discrepancy between their preferred and actual training intensity.

It is important to emphasize that, given their content, the scales of intensity-preference and intensity-tolerance should mainly be expected to exhibit associations with the habitual intensity, rather than the frequency, of exercise and physical activity. Moreover, as stated by Ekkekakis et al. (2005, p. 354), these two intensity-related traits should not be "considered to be the sole determinants of intensity selection or tolerance", since numerous other factors are also likely to play an important role.

As an example, an exerciser whose preference and tolerance scores indicate an inclination toward moderate-intensity activities, would be more likely to experience a state of psychological well-being in response to exercise performed in accordance with these intensity-trait variables rather than intensity that is high or low. In apparent support of this assumption, in subsamples of participants in the present studies who reported a discrepancy between their preference and their tolerance with their actual exercise, preference and tolerance were not associated with the frequency of exercise or well-being-related variables. This finding seems in line with theoretical postulates that emphasize the need to develop individualized approaches to exercise prescriptions and physical activity recommendations (Ekkekakis et al., 2011; Evmenenko and Teixeira, 2020; Rhodes and Kates, 2015).

The fact that PRETIE-Q-PT subscales scores are related (albeit weakly) with exercise habit in the total sample may suggest a possible role of these parameters in habit formation. As stated by Box and Petruzzello (2020) and Brand and Ekkekakis (1), the hedonic theory suggests that experiencing a more pleasant response to exercise increases the possibility of continuing to engage in that behavior. This view is grounded in the Affective-Reflective Theory of physical inactivity and exercise (ART; for a review, see Brand and Ekkekakis, 2018), claiming that an automatic and a reflective systems jointly influence exercise behavior. Thus, it has been suggested that habit formation, an evaluation that accounts for unconscious and automatic processes, may be influenced by both positive affective experiences and positive judgments about the behavior (Kaushal and Rhodes, 2015; Rhodes et al., 2009). As seen in the results section, in the total sample, both preference and tolerance were positively associated with habit. Additionally, while this result should be considered preliminary, it is intriguing that in the subsample of respondents who perceived that their exercise intensity was discrepant from their individual preference, preference was not associated with habit but tolerance was. This may suggest that, in the absence of agreement between preference and the actual training regimen, tolerance may assume a relevant role in habit formation, in line with the relation between tolerance and exercise frequency in this sample subsample.

Strengths, Limitations, and Future Implications

The present study presents the first translation, adaptation, and psychometric evaluation of the PRETIE-Q from American-English to the European-Portuguese language. An important strength of the study were the procedures for testing the invariance of the scale across groups and

over time. Additionally, the concurrent validity data provided information about the relevance of the preference and tolerance traits to exercise behavior and to psychological outcomes associated with exercise in non-student samples.

Readers should also be aware of certain limitations that are important in interpreting the results reported herein. Specifically, the analyses of discriminant and convergent validity revealed certain, relatively minor, deviations from guidelines. In these cases, the retention of items was decided on the basis of conceptual arguments. However, in accordance with the conceptualization of validation as an ongoing process, the continued psychometric evaluation of the Portuguese version of the PRETIE-Q is recommended, particularly with regard to the item with low factorial weight (item 14) and the items that were removed from the original 16-item version. Another limitation pertains to the cross-sectional nature of the study, with the exception of the data on temporal stability. Future validation studies should adopt an experimental design to test theory-derived predictions. Finally, further validation efforts should be enhanced by the addition of device-based methods of assessing exercise and physical activity behavior, and should test a broader range of hypotheses related to discrepancies between individual levels of intensity preference and tolerance and the intensity of current exercise participation. In the present study, low sub-sample size of preference-discrepancy, tolerance-discrepancy, and preference and tolerance-discrepancy do not allow for a more detailed and in-depth analysis, and we recommend further hypothesis testing based on these assumptions with more heterogeneous samples of exercisers and other statistical approaches.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12144-021-01718-3.

Funding This research did not receive any grant or funding from agencies in the public, commercial, or not-for-profit sectors. The authors have no conflict of interest to disclose.

Data Availability The datasets generated during and/or analysed during the current study are not publicly available due to study privacy and ethics committee definitions, but are available from the corresponding author on reasonable request.

Declarations

Nothing to declare.

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee (Beira Interior University Ethics Committee, reference number CE-UBI-pJ-2018-044:ID683) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

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

References

American College of Sports Medicine. (2018). *ACSM's guidelines for exercise testing and prescription* (10th ed.). Philadelphia: Wolters Kluwer.

Baugh, F. (2002). Correcting effect sizes for score reliability: A reminder that measurement and substantive issues are linked inextricably. *Educational and Psychological Measurement, 62*(2), 254–263. https://doi.org/10.1177/0013164402062002004.

Box, A., & Petruzzello, S. (2020). Why do they do it? Differences in high-intensity exercise-affect between those with higher and lower intensity preference and tolerance. *Psychology of Sport and Exercise*, 47. https://doi.org/10.1016/j.psychsport.2019.04.011.

Brand, R., & Ekkekakis, P. (2018). Affective–reflective theory of physical inactivity and exercise: Foundations and preliminary evidence. *German Journal of Exercise and Sport Research, 48*, 48–58. https://doi.org/10.1007/s12662-017-0477-9.

Brislin, R. W. (1980). Translation and content analysis of oral and written material. In H. C. Triandis & R. W. Brislin (Eds.), *Handbook of cross-cultural psychology* (pp. 389–444). Boston: Allyn and Bacon.

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1*(3), 185–216. https://doi.org/10.1177/135910457000100301.

Brooks, D. (2004). *The complete book of personal training*. Champaign: Human Kinetics.

Buckworth, J., Dishman, R., O'Connor, P., & Tomporowski, P. (2013). *Exercise psychology* (2nd ed.). Champaign: Human Kinetics.

Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications and programming* (3rd ed.). Abingdon: Routledge.

Calder, A. J., Hargreaves, E. A., & Hodge, K. (2020). Great expectations: A qualitative analysis of the factors that influence affective forecasts for exercise. *International Journal of Environmental Research and Public Health*, *17*(2). doi.org/10.3390/ijerph17020551.

Chen, F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464–504. https://doi.org/10.1080/10705510701301834.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Couto, N., Antunes, R., Monteiro, D., Moutão, J., Marinho, D., & Cid, L. (2017). Validation of the Subjective Vitality Scale and study of the vitality of elderly people according to their physical activity. *Revista Brasileira de Cineantropometria & Desempenho Humano, 19*(3), 261–269. https://doi.org/10.5007/1980-0037.2017v19n3p261.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart, and Winston.

EC (2018) Special Eurobarometer 472 – Sport and physical activity. Available at: http://data.europa.eu/euodp/data/dataset/S2164_88_4_472_ENG

Edmunds, J., Ntoumanis, N., & Duda, J. L. (2007). Adherence and well-being in overweight and obese patients referred to an exercise on prescription scheme: A self-determination theory perspective. *Psychology of Sport and Exercise, 8*(5), 722–740. https://doi.org/10.1016/j.psychsport.2006.07.006.

Ekkekakis, P., Hall, E. E., & Petruzzello, S. J. (2005). Some like it vigorous: Measuring individual differences in the preference for and

tolerance of exercise intensity. *Journal of Sport and Exercise Psychology, 27*(3), 350–374. https://doi.org/10.1123/jsep.27.3.350.

Ekkekakis, P., Hartman, M. E., & Ladwig, M. A. (2019). Conceptual foundations of exercise psychology: Facilitators, inhibitors, and a road map toward establishing societal relevance. In *APA handbook of sport and exercise psychology, volume 2: Exercise psychology (Vol. 2)* (pp. 27–56). https://doi.org/10.1037/0000124-002.

Ekkekakis, P., Lind, E., Hall, E. E., & Petruzzello, S. J. (2007). Can self-reported tolerance of exercise intensity play a role in exercise testing? *Medicine and Science in Sports and Exercise, 39*(7), 1193–1199. https://doi.org/10.1249/mss.0b013e318058a5ea.

Ekkekakis, P., Lind, E., & Joens-Matre, R. R. (2006). Can self-reported preference for exercise intensity predict physiologically defined self-selected exercise intensity? *Research Quarterly for Exercise and Sport, 77*(1), 81–90. https://doi.org/10.1080/02701367.2006.10599334.

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, 41*(8), 641–671. https://doi.org/10.2165/11590680-000000000-00000.

Ekkekakis, P., Thome, J., Petruzzello, S. J., & Hall, E. E. (2008). The preference for and tolerance of the intensity of exercise questionnaire: A psychometric evaluation among college women. *Journal of Sports Sciences, 26*(5), 499–510. https://doi.org/10.1080/02640410701624523.

Ekkekakis, P., & Zenko, Z. (2016). Escape from cognitivism: Exercise as hedonic experience. In M. Raab, P. Wyllemann, R. Seiler, A. M. Elbe, & A. Hatzigeorgiadis (Eds.), *Sport and exercise psychology research. From theory to practice* (pp. 389–414). Amsterdam: Elsevier. https://doi.org/10.1016/B978-0-12-803634.

Evmenenko, A., & Teixeira, D. S. (2020). The circumplex model of affect in physical activity contexts: A systematic review. *International Journal of Sport and Exercise Psychology.*, 1–34. https://doi.org/10.1080/1612197X.2020.1854818.

Garber, C. E., Blissmer, B., Deschenes, M. R., Franklin, B. A., Lamonte, M. J., Lee, I. M., et al. (2011). Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: Guidance for prescribing exercise. *Medicine and Science in Sports and Exercise, 43*(7), 1334–1359. https://doi.org/10.1249/MSS.0b013e318213fefb.

Gardner, B., Abraham, C., Lally, P., & de Bruijn, G. J. (2012). Towards parsimony in habit measurement: Testing the convergent and predictive validity of an automaticity subscale of the self-report habit index. *International Journal of Behavioral Nutrition and Physical Activity, 9*. https://doi.org/10.1186/1479-5868-9-102.

Gravetter, F., & Wallnau, L. (2014). Essentials of statistics for the behavioral sciences (8th ed.) Cengage learning: Wadsworth, Belmont, CA.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). Multivariate data analysis. *(8th ed.)*. Hoboken, NJ: Pearson Educational.

Hall, E. E., Petruzzello, S. J., Ekkekakis, P., Miller, P. C., & Bixby, W. R. (2014). Role of self-reported individual differences in preference for and tolerance of exercise intensity in fitness testing performance. *Journal of Strength and Conditioning Research, 28*(9), 2443–2451. https://doi.org/10.1519/JSC.0000000000000420.

Ju, H. (2017). The relationship between physical activity, meaning in life, and subjective vitality in community-dwelling older adults. *Archives of Gerontology and Geriatrics, 73*, 120–124. https://doi.org/10.1016/j.archger.2017.08.001.

Kaushal, N., & Rhodes, R. E. (2015). Exercise habit formation in new gym members: A longitudinal study. *Journal of Behavioral Medicine, 38*(4), 652–663. https://doi.org/10.1007/s10865-015-9640-7.

Kline, R. (2016). *Principles and practice of structural equation modelling* (3rd ed.). New York: The Guildford Press.

Ladwig, M. A., Hartman, M. E., & Ekkekakis, P. (2017). Affect-based exercise prescription. *ACSM's Health and Fitness Journal, 21*(5), 10–15. https://doi.org/10.1249/FIT.0000000000000332.

Lochbaum, M., Stevenson, S., & Hilario, D. (2009). Achievement goals, thoughts about intense physical activity, and exerted effort: A Mediational analysis. *Journal of Sport Behavior, 32*(1), 53–68.

Loehr, V. G., & Baldwin, A. S. (2014). Affective forecasting error in exercise: Differences between physically active and inactive individuals. *Sport, Exercise, and Performance Psychology, 3*(3), 177–183. https://doi.org/10.1037/spy0000006.

Lutz, R., Lochbaum, M., Carson, T., Jackson, S., Greenwood, M., & Byars, A. (2008). Are we underestimating the affective benefits of exercise? An experience sampling study of university aerobics participants. *Journal of Sport Behavior, 31*(2), 166–186.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of Golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal, 11*(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2.

McAuley, E., & Courneya, K. S. (1994). The subjective exercise experiences scale (SEES): Development and preliminary validation. *Journal of Sport and Exercise Psychology, 16*(2), 163–177. https://doi.org/10.1123/jsep.16.2.163.

McAuley, E., Blissmer, B., Katula, J., & Duncan, T. E. (2000). Exercise environment, self-efficacy, and affective responses to acute exercise in older adults. *Psychology & Health, 15*(3), 341–355. https://doi.org/10.1080/08870440008401997.

Messick, S. (1988). The once and future issues of validity: Assessing the meaning and consequences of measurement. In H. Wainer & H. I. Braun (Eds.), *Test validity (p. 33–48)*. Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203056905-11.

Miller, J. C., & Krizan, Z. (2016). Walking facilitates positive affect (even when expecting the opposite). *Emotion, 16*(5), 775–785. https://doi.org/10.1037/a0040270.

Moutão, J. M., Alves, S. M., & Cid, L. (2013). Traducción y validación de la subjective vitality scale en una muestra de practicantes de ejercicio Portugueses. *Revista Latinoamericana de Psicologia, 45*(2), 223–230. https://doi.org/10.14349/rlp.v45i2.842.

Muchinsky, P. M. (1996). The correction for attenuation. *Educational and Psychological Measurement, 56*(1), 63–75. https://doi.org/10.1177/0013164496056001004.

Nevitt, J., & Hancock, G. R. (2001). Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modeling. *Structural Equation Modeling, 8*, 353–377. https://doi.org/10.1207/S15328007SEM0803_2.

Palmeira, A. L. (2006). *Análise Factorial Confirmatória da Escala da Experiência Subjectiva com o Exercício—Versão Portuguesa*. In *7° Congresso Nacional de Educação Física, Braga*.

Park, S., Park, K., Hancox, J., Castaneda-Gameros, D., & Koo, K. C. (2019). Physical activity and subjective vitality in older adults from community-and assisted-living settings. *Asian Nursing Research, 13*(5), 313–317. https://doi.org/10.1016/j.anr.2019.11.004.

Pedersen, B. K., & Saltin, B. (2015). Exercise as medicine - evidence for prescribing exercise as therapy in 26 different chronic diseases. *Scandinavian Journal of Medicine and Science in Sports, 25*, 1–72. https://doi.org/10.1111/sms.12581.

Physical Activity Guidelines Advisory Committee. (2018). *2018 physical activity guidelines advisory committee scientific report*. Washington, DC: U.S.Department of Health and Human Services.

Radel, R., Pelletier, L., Pjevac, D., & Cheval, B. (2017). The links between self-determined motivations and behavioral automaticity in a variety of real-life behaviors. *Motivation and Emotion, 41*(4), 443–454. https://doi.org/10.1007/s11031-017-9618-6.

Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence for fixed congeneric components. *Multivariate Behavioral Research, 32*, 329–354.

Raykov, T., Gabler, S., & Dimitrov, D. M. (2015). Maximal reliability and composite reliability: Examining their difference for multicomponent measuring instruments using latent variable modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 23*(3), 384–391. https://doi.org/10.1080/10705511.2014.966369.

Rejeski, W. J., Hardy, C. J., & Shaw, J. (1991). Psychometric confounds of assessing state anxiety in conjunction with acute bouts of vigorous exercise. *Journal of Sport & Exercise Psychology, 13*(1), 65–74. https://doi.org/10.1123/jsep.13.1.65.

Rhodes, R. E., Fiala, B., & Conner, M. (2009). A review and meta-analysis of affective judgments and physical activity in adult populations. *Annals of Behavioral Medicine, 38*, 180–204. https://doi.org/10.1007/s12160-009-9147-y.

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.

Rodrigues, F., Teixeira, D. S., Cid, L., & Monteiro, D. (2019). Have you been exercising lately? Testing the role of past behavior on exercise adherence. *Journal of Health Psychology.*, 135910531987824. https://doi.org/10.1177/1359105319878243.

Rodrigues, F., Teixeira, D. S., Neiva, H. P., Cid, L., & Monteiro, D. (2020). Understanding exercise adherence: The predictability of past experience and motivational determinants. *Brain Sciences, 10*(2), 98. https://doi.org/10.3390/brainsci10020098.

Rose, E. A., & Parfitt, G. (2007). A quantitative analysis and qualitative explanation of the individual differences in affective responses to prescribed and self-selected exercise intensities. *Journal of Sport and Exercise Psychology, 29*(3), 281–309. https://doi.org/10.1123/jsep.29.3.281.

Ruby, M. B., Dunn, E. W., Perrino, A., Gillis, R., & Viel, S. (2011). The invisible benefits of exercise. *Health Psychology, 30*(1), 67–74. https://doi.org/10.1037/a0021859.

Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and Eudaimonic well-being. *Annual Review of Psychology, 52*(1), 141–166. https://doi.org/10.1146/annurev.psych.52.1.141.

Ryan, R. M., & Deci, E. L. (2017). Self-determination theory. In *Basic psychological needs in motivation, development, and wellness*. New York: Gilford Press.

Ryan, R. M., & Frederick, C. (1997). On energy, personality, and health: Subjective vitality as a dynamic reflection of well-being. *Journal of Personality, 65*(3), 529–565. https://doi.org/10.1111/j.1467-6494.1997.tb00326.x.

Sass, D. A. (2011). Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *Journal of Psychoeducational Assessment, 29*, 347–363.

Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods, 1*(2), 199–223. https://doi.org/10.1037/1082-989X.1.2.199.

Sjögren, T., Nissinen, K. J., Järvenpää, S. K., Ojanen, M. T., Vanharanta, H., & Mälkiä, E. A. (2006). Effects of a physical exercise intervention on subjective physical well-being, psychosocial functioning and general well-being among office workers: A cluster randomized-controlled cross-over design. *Scandinavian Journal of Medicine & Science in Sports, 16*, 381–390. https://doi.org/10.1111/j.1600-0838.2005.00516.x.

Smirmaul, B. P. C., Ekkekakis, P., Teixeira, I. P., Nakamura, P. M., & Kokubun, E. (2015). Questionário de Preferência e Tolerância da Intensidade de Exercício: Versão em português do Brasil. *Revista Brasileira de Cineantropometria e Desempenho Humano*, (5), *17*, 550–564. https://doi.org/10.5007/1980-0037.2015v17n5p550.

Sperandei, S., Vieira, M. C., & Reis, A. C. (2016). Adherence to physical activity in an unsupervised setting: Explanatory variables for high attrition rates among fitness center members. *Journal of Science and Medicine in Sport, 19*(11), 916–920. https://doi.org/10.1016/j.jsams.2015.12.522.

Stamatakis, E., Ekelund, U., Ding, D., Hamer, M., Bauman, A. E., & Lee, I. M. (2019). Is the time right for quantitative public health guidelines on sitting? A narrative review of sedentary behaviour research paradigms and findings. *British Journal of Sports Medicine, 53*, 377–382. https://doi.org/10.1136/bjsports-2018-099131.

Teixeira, D. S., & Palmeira, A. L. (2016). Needs satisfaction effect on exercise emotional response: A serial mediation analysis with motivational regulations and exercise intensity. *Motriz. Revista de Educacao Física, 22*(4), 368–375. https://doi.org/10.1590/S1980-6574201600040023.

Vandoni, M., Codrons, E., Marin, L., Correale, L., Bigliassi, M., & Buzzachera, C. (2016). Psychophysiological responses to group exercise training sessions: Does exercise intensity matter*? PLoS One, 11*, e0149997.

Warburton, D. E. R., & Bredin, S. S. D. (2017). Health benefits of physical activity: A systematic review of current systematic reviews. *Current Opinion in Cardiology, 32*, 541–556. https://doi.org/10.1097/HCO.0000000000000437.

Widenfelt, B., Treffers, P., Beurs, E., Siebelink, B., & Koudijs, E. (2005). Translation and cross-cultural Adaptationof assessment instruments used in PsychologicalResearch with children and families. *Clinical Child and Family Psychology Review, 8*(2), 135–147. https://doi.org/10.1007/s10567-005-4752-1.

Williams, D. M., Dunsiger, S., Ciccolo, J. T., Lewis, B. A., Albrecht, A. E., & Marcus, B. H. (2008). Acute affective response to a moderate-intensity exercise stimulus predicts physical activity participation 6 and 12 months later. *Psychology of Sport and Exercise, 9*(3), 231–245. https://doi.org/10.1016/j.psychsport.2007.04.002.