

Psychometric properties of the greek exercise imagery Questionnaire-Aerobic Version

KATERINA DASKALAKI¹, MANOS STEFANAKIS², STELIOS HADJISAVVAS³, MICHALIS
EFSTATHIOU⁴, PARASKEVI MALLIOU⁵

^{1,5} Department of Physical Education and Sport Science, School of Physical Education and Sport Science,
Democritus University of Thrace, GREECE

^{2,3,4,5} Department of Health Sciences, School of Life and Health Sciences, University of Nicosia, NICOSIA
CYPRUS

Published online: December 30, 2024

Accepted for publication: December 15, 2024

DOI:10.7752/jpes.2024.12311

Abstract

Background: The Exercise Imagery Questionnaire-Aerobic Version (EIQ-AV) is a widely used tool for assessing exercise imagery (EI) and has considerably influenced EI research. Although it has been translated into Greek, its psychometric properties remain unevaluated, limiting its applicability for studying EI in the Greek population. This study aimed to evaluate the construct validity (encompassing factorial, convergent, and known-groups validity), criterion validity (specifically concurrent validity), and reliability of the Greek EIQ-AV. **Materials and Methods:** A total of 221 exercisers completed the Greek EIQ-AV, the Godin–Shephard Leisure–Time Exercise Questionnaire (GLTEQ), and a questionnaire on demographic and exercise-related characteristics. Data from 108 participants were analyzed using principal component analysis (PCA), 113 were used for confirmatory factor analysis (CFA), and the full sample was used to assess known-groups and concurrent validity. **Results:** After testing various models based on prior theory and PCA results, we concluded to the final Greek EIQ-AV. Similar to the English version, it comprises three factors: “energy”, “appearance”, and “technique”, differing only by the removal of one item. Factorial validity was confirmed using CFA, demonstrating an excellent model fit, with most fit indices—such as the Standardized Root Mean Square Residual, Comparative Fit Index, and Tucker–Lewis Index—showing excellent values. Convergent validity was supported by acceptable results (AVE > 0.5). Known-groups validity was confirmed because younger, more frequent, and more active exercisers reported significantly higher use of all EI types compared to older, less frequent, and less active exercisers. Additionally, all EIQ-AV subscale scores showed positive correlations with strenuous exercise and the total GLTEQ score and negative associations with age, supporting concurrent validity. Furthermore, the Greek EIQ-AV demonstrated adequate internal consistency and composite reliability. **Conclusions:** The Greek EIQ-AV demonstrates strong psychometric properties and is expected to advance knowledge in EI research.

Keywords: Validity, reliability, mental imagery, technique, appearance imagery

Introduction

Exercise Imagery Questionnaire-Aerobic Version (EIQ-AV) is a questionnaire developed by Hausenblas, Hall, Munroe, & Rodgers (1999). It is the first questionnaire that was created in order to assess exercise imagery (EI) used by exercisers (i.e. what do exercisers imagine regarding exercise). It was called EIQ-AV because it was developed with the participation of exercisers who attended aerobic lessons. EIQ-AV consists of three factors which measure how often exercisers form exercise images in relation to “energy” (images that exercise relieves their stress or get them energized, etc.), “technique” (images about correct technical execution of exercises) and “appearance” (images about losing weight, toning up, or becoming healthy by exercising). Therefore, the original questionnaire indicates that EI is multidimensional (Hausenblas et al., 1999; Kossert & Munroe-Chandler, 2007).

EIQ-AV has been already used widely either intact or with adaptations. Modified versions of it are the “Weight Lifting Imagery Questionnaire” (Munroe-Chandler, Kim, & Gammage, 2004) and the Exercise Imagery Questionnaire (EIQ) (Gammage, Hall, & Rodgers, 2000). A confounding issue seen in the literature is that researchers that have actually used EIQ (Pérez-Fabello, & Campos, 2020; Thøgersen-Ntoumani, Cumming, Ntoumanis, & Nikitaras, 2012, etc.) often cite Hausenblas et al. (1999), giving the impression that they have used EIQ-AV. EIQ was developed by Rodgers, Hall, Blanchard, & Munroe (1999) (as cited in Gammage et al., 2000) and although is based on EIQ-AV, it is not the same tool. In both questionnaires a three-factor structure is seen (i.e. “energy”, “technique” and “appearance”) as well as 9 items. The greatest difference between EIQ-AV and EIQ is that in the former there is a question about imagining health outcomes (included in the appearance sub-scale) which in the latter has been substituted with another question. Additionally, in EIQ, a question about “technique” is added and a question about “energy” is eliminated, while the wording of all its questions is

somewhat different from those included in EIQ-AV. Nevertheless, these two questionnaires share conceptually the same three factors and are very close regarding their items, therefore comparisons regarding the use of EI in studies that have used either one of them are absolutely feasible and already made in many studies.

Research that has used these questionnaires found that exercisers who trained more hours or more times weekly used all three types of EI significantly more than those who trained less (Gammage et al., 2000; Hausenblas et al., 1999), indicating that EI use has a relationship with exercise frequency and exercise intensity. It has also been suggested that age and sex may mediate EI use. For example, appearance imagery is more frequently used by younger exercisers (Milne, Burke, Hall, Nederhof, & Gammage, 2005) and women (Cumming, 2008; Gammage et al., 2000). Imagery about appearance has been also found to be a predictor of intentions to exercise (Rodgers, Munroe & Hall, 2002) and a motivation to maintain exercise behavior (Kim & Giaccobi, 2009). More differences on the preference of imagery type according to age, sex, as well as type of exercise activity have been found in different inquiries (Gammage et al., 2000; Kosteli, Williams, & Cumming, 2019; Thøgersen-Ntoumani et al., 2012). Regarding age and discussing imagery in a broader context, it has been argued that the generation and manipulation of mental images seem to decline with age (Malouin, Richards, & Durand, 2010), while younger adults have been found to form more often mental images than older ones (Daskalaki, Stefanakis, Hadjisavvas, Papadimitriou, & Malliou, 2024b). The exploration of the use of EI by different individuals is important due to its impact on the self-regulation of exercise behaviour (Thøgersen-Ntoumani et al., 2012), its proven effectiveness in enhancing motivation to exercise (Kosteli et al., 2019) and the positive outcomes that are observed in imagery interventions (Daskalaki et al., 2024a; Khairunnisa et al., 2024; Welis et al., 2024). Therefore, the translation and validation of EI questionnaires into other languages would permit a deeper understanding of the particularities of EI in different samples, making possible the sharing of knowledge in an international level.

In 2020, EIQ-AV of Hausenblas et al. (1999) has been translated into Greek through the recommended procedures (i.e. forward and backward translation and pilot research) (Daskalaki et al., 2020). Apart from indications of content validity regarding comprehensibility, i.e. that the questions were clear and understandable, other psychometric properties had not been tested. Further procedures to establish its content validity were not undertaken because rigorous content validation processes demand the control of parameters that could lead to addition or elimination of items (Daskalaki, Stefanakis, Efsthathiou, Beneka, & Malliou, 2024). As it is also discussed by researchers who focus on instrument development, although minor modifications of an instrument (i.e. elimination 2 out of 15 items) may not affect the theoretical background of a model, large re-specifications of an instrument compromise its theoretical integrity, therefore the instrument developed must be seen as a new one (Cid et al., 2022). On the other hand, the advantage of translating a questionnaire over constructing a new one is that it permits intercultural comparisons and facilitates the conduction of meta-analyses (Arafat, Chowdhury, Shalahuddin Qusar, & Hafez, 2016; Epstein, Osborne, Elsworth, Beaton, & Guillemin, 2015).

Taking into account the above, the aims of this study were to: a) produce a valid and reliable Greek version of the EIQ-AV which would be as close as possible to the original, taking advantage of principal component analysis (PCA) and CFA results, b) examine its construct (factorial, convergent and known-groups) as well as its criterion (concurrent) validity, c) investigate its reliability (internal consistency and composite reliability), d) explore if participating in specific exercise programs may differentiate exercisers in EI use.

Given that the meaning of terms regarding validity may vary in different articles (Cid et al., 2022), it should be clarified that, in the present research, “construct validity” refers to: factorial validity [i.e., the assessment of the factor structure of the questionnaire through the values of model fit after CFA; convergent validity at the construct level (Baharum et al. 2023; Cid et al., 2022); and known-groups validity (Ouzouni & Nakakis, 2011)]. For criterion-validity, concurrent validity would be examined through the associations of EIQ-AV with other variables and questionnaires (Cid et al. 2022), investigating if EI use is associated with age and intensity of exercise as this last is measured through the Godin-Shephard Leisure-Time Exercise Questionnaire (GLTEQ) (Godin & Shephard, 1985). It is also important to mention that in the original EIQ-AV, item 9 (i.e. When I imagine exercising, it gets me energized) loaded on both the “Energy” and “Technique” sub-scales, but as content seems more relevant to “Energy”, while it differs conceptually from the two other items of “Technique”. Therefore, in case item 9 was found to load only on “Energy” (which would leave the technique sub-scale with only two items), that would not be considered as a deviation from the original model.

The hypotheses of the study were the following: regarding construct validity, the Greek EIQ-AV would demonstrate factorial and convergent validity after slight modifications of the original model; for known-groups validity, it was hypothesized that younger or more frequent or more active exercisers would report more frequent EI use in all subscales of EIQ-AV in comparison with older, less frequent or less active exercisers. For concurrent validity, our hypothesis was that the scores of the three subscales of the Greek version of EIQ-AV would be positively correlated with strenuous exercise and total score of GLTEQ and negatively correlated with age. No clear hypothesis was formulated about if implication in specific types of programs relate to EI use, due to lack of previous comparative data for exercisers who participate in various exercise activities. It was also hypothesized that the Greek EIQ-AV would have adequate internal consistency reliability and composite reliability.

Materials & Methods

Participants

The data were collected digitally, through Google Forms. Inclusion criteria were to be an adult (i.e. above 18 years old) and have an active subscription at a gym or similar fitness facility (e.g. Fitness Studio, Pilates Studio etc.). The research was announced in 4 Greek gyms as well as through the social media. In the research participated, after providing informed consent, 225 exercisers who were split randomly (with the function “approximately 50%” of SPSS) to form two independent samples. The first sample (n=109) would be used for PCA and the second one (n=116) for CFA. The sample size was adequate, since even 100 cases are sufficient when the variables are few, while the participants to variables ratio exceeded the 1:10 recommendation (Hooper, 2012). For both samples, skewness and kurtosis indexes of the 9 items of Greek EIQ-AV were either close to zero and/or less than 2, confirming univariate normality (Bollen & Long, 1993; Pérez-Fabello & Campos, 2020). Multivariate normality was tested by means of the Mahalanobis’ distance (Zygmunt & Smith, 2014) and values that were found to be .001 were removed as outliers (Pérez-Fabello & Campos, 2020). Thus, one case was removed from sample 1 and three cases were removed from sample 2, concluding in 108 and 113 eligible participants respectively. In order to examine known-groups and concurrent validity, the eligible participants of both samples (i.e. sample 3) would be used. The demographic, anthropometric and exercise-related characteristics of all samples are displayed in Table 1.

Table 1. Characteristics of the participants

	Sample 1 (n=108)	Sample 2 (n=113)	Sample 3 (n=221)
Sex	Women =100 (92.6%); Men=8 (7.4%)	Women=103 (91.2%); Men=10 (8.8%)	Women= 203 (91.9%); Men= 18 (8.1%)
Age (range, mean±SD)	22-69 years (45.96±10.82)	20-74 years (46.60±11.21)	20-74 years (46.28±11)
BMI	17.26-35.43 (22.91±3.54)	17.97-33.97 (22.83±3.09)	17.26-33.97 (22.87±3.31)
Years of exercising in a gym	1 month - 50 years (10.19±10.65 years)	2 months-46 years (11.41±11.25)	1 month-50 years (10.81±10.95)
Hours of training per week	1 to 14 hours (4.29±2.60)	1 to 16 hours (4.24±2.74)	1 to 16 hours (4.26±2.66)
Weight Training	n=31	n=45	n=76
Cardiovascular machines	n=29	n=39	n=68
Pilates	n=79	n=72	n=151
Yoga	n=26	n=23	n=49
Group Strength and Conditioning Programs	n=39	n=45	n=84
Group Choreography Programs	n=17	n=14	n=31
Suspension Training	n=16	n=13	n=29
CrossFit/Cross Training	n=17	n=8	n=25

Note. Group Strength and Conditioning Programs include programs that are realized in a group, using weights, elastic bands etc. Cardiovascular machines include: treadmill, elliptical, stationary bike/indoor cycling, etc. Group Choreography Programs include: Aerobic, Dance Aerobic, Step Aerobic, and Zumba.

Measurements

Greek Translation of EIQ-AV

EIQ-AV (Hausenblas et al., 1999) in its Greek translation (Daskalaki et al., 2020) was used. It consists of instructions, 9 items and 3 subscales (i.e. “energy”, “appearance” and “technique”). Each item is scored in a 9-point Likert Scale, from 1 (never) to 9 (always). The scores in all individual scales are summed and then the mean score for every dimension is calculated. Its validity and reliability have been established in the research of Hausenblas et al. (1999). The Greek translation of the questionnaire is provided in Table 2. Throughout the manuscript, for facilitating the readers, the items are written in English. The items are numbered identically in both the English and the Greek versions.

Table 2. Instructions and items of the Greek Translation of EIQ-AV

Νοερή απεικόνιση σημαίνει να βλέπεις με τη φαντασία σου τον εαυτό σου να γυμνάζεται. Η εικόνα στο μυαλό σου πρέπει να πλησιάζει την πραγματική σωματική δραστηριότητα όσο το δυνατόν περισσότερο. Η νοερή απεικόνιση μπορεί να περιλαμβάνει αισθήσεις, όπως το να ακούς τη μουσική του προγράμματος και να νιώθεις τον εαυτό σου να κινείται μέσα στις ασκήσεις. Η νοερή απεικόνιση μπορεί επίσης να σχετίζεται με συναισθήματα (π.χ. να εμψυχώνεσαι ή να ενεργοποιείσαι), με το να παραμένεις επικεντρωμένος (π.χ. να συγκεντρώνεσαι στο μάθημα της αεροβικής γυμναστικής και να μην αφαιρείσαι), με το να κάνεις σχέδια/να θέτεις στόχους σε σχέση με την άσκηση (π.χ. να φαντάζεσαι τον στόχο να χάσεις βάρος) κτλ.
Βαθμολογήστε από το 1 έως το 9 πόσο συχνά χρησιμοποιείτε τους παρακάτω τύπους νοερής απεικόνισης όπου 1=ποτέ δεν χρησιμοποιώ αυτόν τον τύπο απεικόνισης και 9=πάντα χρησιμοποιώ αυτόν τον τύπο απεικόνισης.
1.Όταν φαντάζομαι ότι κάνω άσκηση, με βοηθάει να συνεχίσω την ημέρα μου.

2. Φαντάζομαι ότι αποκτώ ένα πιο «γραμμωμένο» σώμα μέσω της άσκησης.
3. Όταν φαντάζομαι ότι κάνω άσκηση, χαλαρώνω από το άγχος.
4. Όταν μαθαίνω μια νέα σειρά ασκήσεων, φαντάζομαι ότι κάνω τις κινήσεις τέλεια.
5. Φαντάζομαι τη μορφή/τη θέση του σώματός μου.
6. Φαντάζομαι ότι γίνομαι πιο υγιής μέσω της άσκησης.
7. Όταν φαντάζομαι ότι κάνω άσκηση, ξεφεύγει το μυαλό μου από τη δουλειά.
8. Φαντάζομαι ότι χάνω βάρος μέσω της άσκησης.
9. Όταν φαντάζομαι ότι κάνω άσκηση, ενεργοποιούμαι.

Note. For all items, a 9-point Likert scale was provided.

Godin-Shephard Leisure-Time Exercise Questionnaire

The GLTEQ (Godin & Shephard, 1985) in Greek (Theodorakis & Hassandra, 2005) was used to measure the exercise behaviour of the participants. The respondents were asked to answer on the question: “Considering a typical week, how many times on the average do you do the following kinds of exercise for more than 15 times per week during your free time”. Since the form was digital, they had to tick a box (from 0 to 7 times per week) they were implicating in: a) strenuous exercise (heart beats rapidly)- e.g. fast running, long-distance jogging, vigorous long-distance bicycling, aerobics etc. b) moderate exercise (not exhausting)- e.g. fast walking, easy bicycling, light weights etc., c) mild exercise (minimal effort)- e.g. easy walking, yoga, etc. The examples provided for each intensity were activities that may be seen in a gym. For each type of activity, the score in metabolic equivalents (METs) was calculated by multiplying “times per week” with 3 for mild, 5 for moderate, and 9 for strenuous activity. For the whole score, METs were summed up, e.g. (5 times X 3) + (3 times X 5) + (1 time X 9) = 39 units.

Demographic, anthropometric, and exercise-related variables

Sex, age, weight, height, years of exercising in a gym, hours of training per week, and programs attended at the gym were recorded through relative questions in the digital form. Body Mass Index (BMI) was calculated later by weight and height through the formula: $BMI = (\text{weight} \times \text{weight}) / \text{height}$.

Analysis of the data

Analysis of the data was realized using the SPSS Version 29.0.0.0 (241). The Jamovi software (Version 2.3.) was used for CFA. There were not missing data in the questions of EIQ-AV and GLTEQ, since the forms were digital and responses were mandatory. In very few cases, there were missing or ambiguous data regarding age, years of exercising, weight and height (because these fields demanded a short text answer and not just ticking a box), and these cases were deleted pairwise. Significance was set at $p < 0.05$.

Principal Components Analysis (PCA)

Sampling Adequacy was evaluated with the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) (Watkins, 2018; Kaiser, 1974) and factorability of the data with Bartlett's Test of Sphericity. To determine the number of components to extract, the criteria of eigenvalue >1 and the Cattell's scree test were used (Williams, Onsman, & Brown, 2010), with the extraction method of PCA, and using a varimax rotation, replicating the method used by the originators of the questionnaire (Hausenblas et al., 1999). Corresponding to prior theory and research (Watkins, 2018), also a fixed number of 3 factors to extract was decided, since PCA in this study was undertaken with the main purpose to examine if the Greek translation corresponded to the relations of the questions/items with the components seen in the English EIQ-AV and to obtain information for the existence of possible alternative models that would be used when conducting CFA. It should be noted that Hausenblas et al. (1999) only claimed that “criteria for retention of an item on a factor was set at .50” and not provided other information on the criteria used for determining the number of components to extract. So, the whole process of PCA was aimed to produce a three component structure and to identify if there were cross loading items that could possibly cause inadequate fit in the subsequent CFA (Farell & Rudd, 2009). Useful information to be taken into account was the following: a) for communalities to considered acceptable, average values should be at least between 0.5 and 0.6 as it is recommended for samples with 100-200 participants (Shrestha, 2021), b) loadings above 0.70 are considered excellent (Brown, 2009; Comrey & Lee, 1992), c) a cross loading of an item on a secondary component/factor of more than .30 complicate things (Sass & Schmitt, 2010) and may suggest that the item is unreliable (Hooper, 2012), d) final decisions for elimination of items would be made only during CFA.

Confirmatory Factor Analysis (CFA)

The original model of Hausenblas et al. (1999) as well as the models suggested by PCA were tested through CFA. Since in Hausenblas et al. (1999) item 9 was loading on two factors, two versions would be checked. In “Original Version 1”, item 9 would be included in “technique”, while in “Original Version 2” in “energy”. In all tested models, items would be allowed to load on just one factor, factors would be allowed to freely correlate, while error terms (residuals) would not be allowed to correlate (Chung, & Dong Liu, 2012). Therefore, we would not add paths or correlations between items (Carter, 2006), because that would increase errors in the model (Hermida, 2015). The criteria explored for model fit were: Chi-square (χ^2), the χ^2/df ratio, the Root Mean Squared Error of Approximation (RMSEA), the Standardized Root Mean Square Residual (SRMR), the Comparative Fit Index (CFI), the Tucker Lewis Index (TLI). Values for indices is suggested to be as follows: The χ^2 is desirable to be nonsignificant (Hinkin, 1998). The χ^2/df ratio can be acceptable with values up to 5.0, a

value of 3 is also considered reasonable, however a good fit is indicated when the value is lower than 2 (Cid et al., 2022). A RMSEA between 0.5 and 0.10 suggest “acceptable fit” (Lai & Green, 2016), and between 0.08 and 0.01 “mediocre fit” (De Moor, Distel, Trull, & Boomsma, 2009). SRMR should be <.05 (Ennigkeit & Hänsel, 2018). For CFI and TLI values close to .95 or above indicate good, acceptable fit (Hu & Bentler, 1999; Lai & Green, 2016).

Convergent validity

When a CFA is performed, it is important to examine convergent validity, which according to Hair, Black, Babin, Anderson and Tatham (2019) “measures the extent to which the indicators of a construct converge, thereby explaining the variance of the item” (as cited in Cid et al., 2022). A measure of convergent validity is the average variance extracted (AVE) which stands for the average (i.e. mean) of the squared loadings of all items belonging to a particular construct. It should take values equal or more than 0.5 (Cid et al., 2022; Shrestha, 2021). AVE was calculated manually for each factor, using the output of CFA, where the “standardized estimates” of factor loadings (λ) are demonstrated. In the literature, two formulas are found, which is a bit misleading, since they lead to different results. The first is: $AVE = \Sigma \lambda^2 / n$, where “ Σ ” stands for “sum”, “ λ ” for the factor loadings (i.e. the standardized estimates) and “ n ” the number of items in each factor. So, AVE is the sum (Σ) of the squared standardized factor loadings ($\Sigma \lambda^2$) divided by the number of items (n) (Afework et al., 2021; Muhamad Safiih & Azreen, 2016; Shrestha, 2021; Yusoff, 2011). The second formula, in its simplified form, is: $AVE = \Sigma \lambda^2 / [\Sigma \lambda^2 + \Sigma (1 - \lambda^2)]$ (Achjari, 2004; Carter, 2016). In order to be able to provide full information and permit comparisons with future studies, both equations were used for the calculation of AVE.

Known-groups and concurrent validity

To examine known-groups validity, first “extreme groups” (i.e. groups that differ much in a variable) were formed from the data (Hausenblas et al., 1999). The lowest and highest 25% of the distribution were obtained through splitting the observations for the variables in question in quartiles and the respective groups were as follows: for age, the two groups were “younger exercisers” (from 20 to 38 years old) and “older exercisers” (from 52 to 74 years old). For exercise frequency, “less frequent exercisers” were exercising up to 2 hours weekly, while “more frequent exercisers” from 5 hours and more. For exercise intensity, “less active exercisers” had a score in GLTEQ up to 25 units, while “more active exercisers” from 52 units and more. The scores of these different groups were compared across the three sub-scales as well as in the total score of the Greek version of EIQ-AV. For concurrent validity, it was examined if age, strenuous exercise and total score of GLTEQ (intensity of exercise) were correlated with the three subscales as well as with the total score of the Greek version of EIQ-AV.

Reliability (composite and internal consistency reliability)

Composite reliability (CR) refers to the estimation of internal consistency reliability when conducting a CFA (Cid et al., 2022), therefore it was calculated only for the second sample. It was calculated manually, using the output of CFA. The formula is: $CR = (\Sigma \lambda)^2 / [(\Sigma \lambda)^2 + \Sigma (1 - \lambda^2)]$ (Cheung, Cooper-Thomas, Lau & Wang, 2024; Muhamad Safiih, & Azreen, 2016). Ideally, its values must be above 0.70, however, also a value of >0.60 is accepted (Bagozzi & Yi, 1988; Baharum et al., 2023; Husain & Aziz, 2022; Sharma, & Hussain, 2019).

The internal consistency reliability using Cronbach’s alpha (α) coefficient was calculated with SPSS for the three subscales and for the whole scale. This procedure was undertaken in all three samples, since the reliability of a questionnaire must be checked in all validation processes (Tsang, Royse, & Terkawi, 2017). According to George and Mallery (2003), adequate α values can be described as follows: $\alpha \geq .9$ = excellent, 0.80–0.89 = good, 0.70–0.79 = acceptable (Flo, Landmark, Hatlevik, & Fagerström, 2018).

Results

Principal Component Analysis (with the first sample)

The KMO value of sampling adequacy was .914, which according to Kaiser (1974) is considered “marvellous” (Watkins, 2018). The Bartlett’s Test of Sphericity was significant ($\chi^2_{(36)} = 900.156$; $p < .001$) confirming the factorability of the data. After Varimax Rotation, a one factor solution emerged that explained 70.18% of the variance (Table 3).

Table 3. Loadings and communalities of items of the “one factor model”

Component Matrix	Component	Communalities
	1	
3. When I imagine exercising, it relieves my stress.	.911	.830
9. When I imagine exercising, it gets me energized.	.897	.804
7. When I imagine exercising, it takes my mind off my work.	.892	.795
1. When I imagine exercising, it keeps me going during the day.	.880	.774
6. I imagine becoming healthy by exercising.	.824	.679
5. I imagine form/body position.	.811	.659
2. I imagine toning up by exercising.	.789	.623
4. When learning a new routine, I imagine doing the moves perfectly.	.784	.614
8. I imagine losing weight by exercising.	.734	.539

After that, a “forced” extraction of three factors was tested (“forced PCA model”). The three components structure accounted for 85.848% of the variance. The first component accounted for 39.24% of the variance and included items 1, 3, 6, 7, 9 (corresponding to “energy”). The second component accounted for 23.82% of the variance and included items 4 and 5 (corresponding to “technique”). The third component accounted for 22.788% of the variance and included items 2 and 8 (corresponding to “appearance”) (Table 4). The names given to the components were those of Hausenblas et al. (1999). In this model, a difference from the original EIQ-AV was that in the latter item 6 loaded on “appearance”, while item 9 not only on “energy” but also on “technique”.

Table 4. Eigenvalues and Percentage of Variance before and after rotation of the “Forced PCA model”

Component	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6.317	70.185	70.185	3.532	39.240	39.240
2	.803	8.926	79.111	2.144	23.820	63.060
3	.606	6.737	85.848	2.051	22.788	85.848
4	.346	3.849	89.697			
5	.301	3.343	93.040			
6	.203	2.256	95.296			
7	.177	1.964	97.260			
8	.156	1.737	98.997			
9	.090	1.003	100.000			

The communalities were between 0.738 and 0.913, and therefore acceptable (Shrestha, 2021). The loadings of the items on their respective components were excellent, all above 0.70, except for item 1, which was also loading highly on two components (Table 5).

Table 5. Loadings and communalities of items of the “Forced PCA model” (rotated component matrix)

	Component			Communalities
	1	2	3	
7. When I imagine exercising, it takes my mind off my work.	.866			.905
3. When I imagine exercising. it relieves my stress.	.835			.903
9. When I imagine exercising, it gets me energized.	.801			.858
6. I imagine becoming healthy by exercising.	.730			.738
1. When I imagine exercising, it keeps me going during the day.	.672	.534		.807
4. When learning a new routine, I imagine doing the moves perfectly.		.822		.857
5. I imagine form/body position.	.418	.798		.857
8. I imagine losing weight by exercising.			.887	.913
2. I imagine toning up by exercising.			.822	.889

*Component 1 corresponds to “energy”, 2 to “technique”, and 3 to “appearance”.

Internal consistency reliability (for the first sample)

Internal consistency reliability of the whole scale was excellent ($\alpha=0.946$). For the “energy” sub-scale, it was also excellent ($\alpha=0.948$), but it would be even higher, if item 6 was deleted ($\alpha=0.951$). For “appearance” and “technique”, α values were good (0.887 and 0.841 respectively).

Hypothesized models according to the results of PCA and to prior theory

Considering that before the rotation only 1 component had eigenvalue >1 , a “one-factor model” would be controlled. The second model that would be checked was the one that corresponded to the three components structure (“forced PCA model”). The third and fourth model, were the two versions of the “original model” proposed by Hausenblas et al. (1999). It was hypothesized that item 1 would also create problems at CFA and it should be maybe eliminated.

CFA (with the second sample)

For CFA, five different models were tested before concluding to the best fitting model (Table 6). The “one-factor model” had unsatisfactory model fits. Similarly, the “forced PCA model” exhibited a poor fit. According to the modification indices, the item that created a problem was item 6 that was loading much on “appearance”. Then, the two versions of the “original model” proposed by Hausenblas et al. (1999) were tested. The first version (original model 1) with item 9 in “technique” did not fit well, and the modification indices implied that the biggest problems were ought to item 9 which was loading much on “energy”. In the second version (original model 2) with item 9 in “energy” the indices were better. Observing the modification indices, the item that was creating problems was item 1. In case we removed it, all values were acceptable. Therefore, in order to be as close as possible to the original questionnaire, we retained original model 2, eliminating item 1, and this was our “final model” which constituted the Greek Version of EIQ-AV. With this one modification the “final model” had good fits: χ^2 was not significant ($0.018 < 0.05$), the χ^2/df ratio was below 2 (1.84) and indicated a perfect fit of the final model to the data. CFI and TLI were excellent (0.979 and 0.966 respectively) and SRMR was below the cut-off of 0.05, obtaining a value of 0.0259. RMSEA was not ideal but was within the acceptable limits (0.0867).

Table 6. Model fit indices of the five models

Models	Test for Exact Fit				Fit Measures				RMSEA 90% CI	
	χ^2	df	χ^2/df	p (χ^2)	CFI	TLI	SRMR	RMSEA	Lower	Upper
1. One Factor	170	27	6.29	<.001	0.834	0.778	0.0602	0.216	0.186	0.248
2. Forced PCA	78.6	24	3.27	<.001	0.936	0.905	0.0357	0.142	0.108	0.177
3. Original Model 1	133	24	5.54	<.001	0.872	0.809	0.0989	0.201	0.168	0.235
4. Original Model 2	68.5	24	2.85	<.001	0.948	0.922	0.0311	0.128	0.0932	0.164
5. Final Model	31.4	17	1.84	0.018	0.979	0.966	0.0259	0.0867	0.0356	0.133

Original Model 1: Item 9 on Technique; Original Model 2: Item 9 on Energy

Final Model: Original Model 2 without item 1 (Greek Version of EIQ-AV)

All factor loadings were above the recommended 0.70 indicating that each factor explained more than the 50% variance of the item (Cheung et al., 2024) while they were all statistically significant (Table 7).

Table 7. Factor loadings, confidence intervals and statistical significance

Factors	Items	Estimate	SE	95% Confidence Interval		Z	p	Stand. Estimate
				Lower	Upper			
Energy	3	2.39	0.199	2.00	2.78	12.02	<.001	0.898
	7	2.29	0.195	1.91	2.67	11.76	<.001	0.885
	9	2.20	0.195	1.82	2.58	11.28	<.001	0.864
Appearance	2	1.93	0.193	1.55	2.30	9.99	<.001	0.806
	6	2.02	0.193	1.64	2.40	10.45	<.001	0.832
	8	1.68	0.204	1.28	2.08	8.20	<.001	0.701
Technique	4	2.16	0.191	1.79	2.53	11.32	<.001	0.879
	5	2.24	0.168	1.91	2.57	13.37	<.001	0.978

Convergent validity

AVE, using both equations, was far above the minimum recommended value of 0.50 (Table 8).

Reliability (cronbach's α) and composite reliability of the Final Model

In the second sample, both internal consistency reliability and composite reliability (CR) were at the indicated levels. For the sub-scales “energy” and “technique” the internal consistency reliability (cronbach's α) was excellent, and for “appearance” it was good. For “appearance”, in case we deleted item 6, the reliability would fall to 0.715, which also gave credit to our choice to include it in this factor, as originally was. The whole scale had excellent reliability (0.927). Because “technique” consisted of only two items, the inter-item correlations of item 4 and 5 were also controlled: They were highly inter-correlated (.859), while their correlations with other items were below .628. CR was also acceptable and above the indicated levels for all sub-scales. All values as well as the way they were calculated are demonstrated in Table 8.

Table 8. Values used in the formulas, AVE, composite reliability and cronbach's α

FACTORS	λ	λ^2	(1- λ^2)	AVE 1	AVE 2	CR	α
ENERGY							
3. When I imagine exercising, it relieves my stress.	0.898	0.806	0.194				
7. When I imagine exercising, it takes my mind off my work.	0.885	0.783	0.102				
9. When I imagine exercising, it gets me energized.	0.864	0.746	0.118				
	$\Sigma\lambda = 2.647$ ($\Sigma\lambda$) ² =7	$\Sigma\lambda^2 = 2.335$	$\Sigma = .414$	0.778	0.849	0.94	0.914
TECHNIQUE							
4. When learning a new routine, I imagine doing the moves perfectly	0.879	0.772	0.228				
5. I imagine form/body position.	0.978	0.956	0.044				
	$\Sigma\lambda = 1.857$ ($\Sigma\lambda$) ² =3.45	$\Sigma\lambda^2 = 1.678$	$\Sigma = .332$	0.839	0.83	0.91	0.923
APPEARANCE							
2. I imagine toning up by exercising.	0.806	0.649	0.351				
6. I imagine becoming healthy by exercising.	0.832	0.692	0.308				
8. I imagine losing weight by exercising.	0.701	0.491	0.509				
	$\Sigma\lambda = 2.33$ ($\Sigma\lambda$) ² =5.47	$\Sigma\lambda^2 = 1.832$	$\Sigma = 1.168$	0.61	0.61	0.82	0.822

λ : Stand. Estimate (stand.factor loading); AVE 1 is the result according to the formula: $AVE = \Sigma\lambda^2/n$; AVE 2 is the result according to the formula: $AVE = \Sigma\lambda^2 / [\Sigma\lambda^2 + \Sigma (1-\lambda^2)]$; CR: Composite reliability; α : cronbach's alpha

Concurrent and known-groups validity (third sample)

Concurrent validity

The data were not normally distributed, therefore correlations between the variables were examined with Spearman's correlations. Total score of GLTEQ was positively related with all subscales of the Greek version of EIQ-AV (energy: $r = 0.232$, $p < .001 < 0.05$; technique: $r = 0.217$, $p < .001 < 0.05$; appearance: $r = 0.263$, $p < .001 < 0.05$) and with the total score of EIQ-AV ($r = 0.269$, $p < .001 < 0.05$) (one-tailed). Strenuous exercise was positively correlated with "energy" score ($r = 0.236$, $p = 0.001 < 0.05$), "technique" score ($r = 0.257$, $p = 0.001 < 0.05$), "appearance" score ($r = 0.344$, $p = 0.001 < 0.05$), and total score of EIQ-AV ($r = 0.316$, $p = 0.001 < 0.05$). Age was negatively associated with all subscales (energy: $r = -0.151$, $p = 0.014 < 0.05$; appearance: $r = -0.291$, $p < .001 < 0.05$; technique: $r = -0.133$, $p = 0.026 < 0.05$) and total score of EIQ-AV ($r = -0.214$, $p = 0.001 < 0.05$) (one tailed).

Known groups validity

The data were not normally distributed, therefore the non-parametric Independent Samples Mann-Whitney U test was used. The independent (grouping) variables were: age, exercise frequency, and exercise intensity (scores in GLTEQ). The dependent variable was the scores in the three subscales as well as in the whole questionnaire. Regarding age, there were statistically significant differences between younger ($n=55$) and older exercisers ($n=56$) in all subscales and in total score of the Greek EIQ-AV (more frequent use of EI by younger exercisers) ($p < 0.05$). Regarding exercise frequency, there were statistically significant differences between less frequent ($n=50$) and more frequent exercisers ($n=77$) in all subscales and in total score of the Greek EIQ-AV ($p < 0.05$) (more frequent use of EI by more frequent exercisers). Regarding exercise intensity, there were statistically significant differences between less active ($n=64$) and more active exercisers ($n=56$) in all subscales and in total score of the Greek EIQ-AV ($p < 0.05$) (more frequent use of EI by more active exercisers) (Table 9).

Table 9. Mean Ranks and means of EI use (three subscales and total score of EIQ-AV) according to age, exercise frequency and exercise intensity

Exercisers (age)		Mean Rank	Mann-Whitney U	Sig.	Mean
Energy	Younger	63.58	1123	.007	6.64±2.21
	Older	48.55			5.46±2.54
Technique	Younger	61.58	1233	.035	5.26±1.97
	Older	50.52			4.56±2.23
Appearance	Younger	72.61	626	<.001	6.77±1.54
	Older	39.69			4.68±1.97
EIQ-AV Total	Younger	67.58	903	<.001	18.67±5.72
	Older	44.63			14.7±6.74
Exercisers (Frequency)					
Energy	Less Frequent	53.39	2455	0.0045	5.25±2.53
	More Frequent	70.89			6.42±2.40
Technique	Less Frequent	56.28	2311	0.029	4.57±2.15
	More Frequent	69.01			5.35±2.36
Appearance	Less Frequent	55.86	2332	0.022	5.11±2.33
	More Frequent	69.29			5.97±1.99
EIQ-AV Total	Less Frequent	54.04	2423	0.007	14.93±7.01
	More Frequent	70.47			17.74±6.75
Exercisers (Intensity)					
Energy	Less active	51.69	2356	0.001	5.36±2.43
	More active	70.57			6.64±2.47
Technique	Less active	50.71	2418.5	<.001	4.47±1.89
	More active	71.69			5.75±2.34
Appearance	Less active	51.38	2376	0.001	5.34±2.06
	More active	70.93			6.5±2.16
EIQ-AV Total	Less active	50.23	2449	<.001	15.17±6.38
	More active	77.24			18.89±6.97

*Means±SD are displayed for descriptive reasons and are calculated by dividing the score of the respective scale by the items. P is one-tailed.

Regarding imagery use by program (i.e. Pilates, Yoga, Cardiovascular Machines, Weight Training, Group Strength and Conditioning Programs, Group Choreography Programs, Suspension Training and CrossFit/Cross Training), statistically significant differences were found only between exercisers who attended or not attended Weight Training and Group Choreography Programs (henceforth aerobic exercisers). Weight-trainers made significantly more use of appearance imagery (mean rank= 122.79) than those who did not practice weight-training (mean rank=104.95), with $U=4591$, $p = .049 < 0.05$. Aerobic exercisers made significantly more use of appearance imagery (mean rank= 132.85) than non-aerobic exercisers (mean rank=106.23) with $U=2205.5$, $p = .030 < 0.05$. Additionally, aerobic exercisers made significantly more use of imagery (mean rank= 132.00) than non-aerobic exercisers (mean rank=106.37) with $U=2232$, $p = .037 < 0.05$.

Reliability (cronbach's α) (third sample) Internal consistency reliability (α) for the "energy" sub-scale was .943, for "technique" .884, for "appearance" .842, and for the whole scale .927. The values would not be improved if items were deleted.

Discussion

The purpose of this study was to assess the psychometric properties of the Greek version of EIQ-AV. First, a preliminary PCA was carried out. The results of PCA informed the subsequent CFA, during which 5 different models were tested, concluding to a valid and reliable questionnaire which was very close to the original.

PCA was performed by following the recommended procedures, but also with forced extraction of 3 factors. Although, by doing the latter, the usual statistical criteria were not addressed, this approach was followed in order to simulate as much as possible the original EIQ-AV developed by Hausenblas et al. (1999). Useful information gained by this procedure was that item 1 (i.e. When I imagine exercising, it keeps me going during the day) was loading heavily on two components, while item 6 (i.e. I imagine becoming healthy by exercising) was loading on "energy" in difference with the original questionnaire in which it is included in "appearance". However, all other items were loading on their established components. By following this procedure, EIQ-AV was also explored in a new population, before testing the original model.

When conducting CFA, comparisons of different models to the same data and of different combinations till the best fit is achieved is a tactic found in the literature. It is also underlined that apart from statistical criteria, theoretical and conceptual criteria must be taken into account when a study builds upon previous theory and knowledge (Bagozzi & Yi, 1988). In order to be consistent with recommended statistical procedures, and since the data of PCA (i.e. the initial eigenvalues) indicated that the Greek EIQ-AV could be unidimensional, we first tested a one-factor model which, however, was not fitting well to the data. Apart from that, researchers caution against the usage of solutions with less than three factors (i.e. just one or two), as they may provide inaccurate representations of the structure (Beavers et al., 2013). Then, the 3-factor model produced by PCA was tested, whose model fits were poor. The biggest problem was created by item 6, which according to the modification indices loaded much on "appearance", i.e. on the factor originally proposed by Hausenblas et al. (1999). Indeed, when testing the second version of the original model, with the item in its original factor, the model fits improved. However, the elimination of item 1 seemed to be unavoidable, since, when retaining it, the model fits were poor, while it was also loading heavily on two factors in PCA. Therefore, the only major change done was the removal of this one item. This would be considered as a minor modification which does not compromise the theoretical integrity of the model (Cid et al., 2022), while it was also double-checked by PCA and CFA in different samples.

As for "technique", the choice to retain item 9 in "energy" and keep "technique" with only two items (as was indicated by PCA and CFA results), not only is consistent with the original model, in which this item was anyway loading also on "energy", but it has also refined the structure of the questionnaire, making it clear how to calculate the scores of its subscales. Additionally, this clarification seems logical, since the remaining two items describe much more consistently "technique", and as it is reported, it is important that theoretical considerations are taken into account in model re-specification (Wipulanusat, Panuwatwanich, & Stewart, 2017). Although often it is recommended that a factor should have no less than three variables, it is also discussed that "this depends on the design of the study" (Yong & Pearce, 2013). The inclusion of only two variables in a factor is acceptable by the literature (Williams et al., 2010) when there is a clear justification for such a decision (Schreiber, 2021). It is also sustained that a two-item factor could be reliable if there is a high correlation between the two items ($r > .70$) but not with other variables (Yong & Pearce, 2013). Indeed, the inter-item correlation of items 4 and 5 was above 0.70 and the highest in comparison with all other items.

The Greek EIQ-AV was found to possess construct validity since factorial, convergent and known-groups validity were established. Specifically, regarding factorial validity, the model fits of the final model were in most cases excellent, and in one case (i.e. RMSEA) at least acceptable, which is very satisfactory when considering that no correlated errors were permitted in the model (Negarandeh, Sarkhani, & Heydari, 2024). Convergent validity at the construct level was confirmed since the values of AVE for all subscales were above 0.5. Known-groups validity was also verified since younger, more frequent and more active exercisers were using significantly more all types of imagery than older, less frequent and less active ones. These findings are in accordance with previous research that indicated that exercisers who worked out more times (Hausenblas et al., 1999) or more hours per week (Gammage et al., 2000) used all types of EI more than less avid exercisers. Also, younger exercisers used not only appearance imagery more than older ones as it was suggested in previous research (Thøgersen-Ntoumani et al., 2012) but all types of EI, which was consistent with our initial hypothesis, based on relative data (Daskalaki et al., 2024b; Malouin et al., 2010). All these results demonstrate that the Greek Version of EIQ-AV is capable of differentiate groups with different characteristics. Regarding criterion (concurrent) validity, it was found that the scores of all subscales of EIQ-AV were positively correlated with strenuous exercise as well as with the total score of GLTEQ, and negatively associated with age, as we had initially hypothesized.

An interesting finding was that differences in types of imagery used by program was found only in exercisers who practiced weight training or attended choreography programs, since these exercisers were using more appearance imagery than those who did not participate in such programs. However, this finding must be seen with caution and should be further investigated because the exercisers of our sample did not practice only one but more types of programs. Gammage et al. (2000) had also investigated imagery use by type of activity and included in their analyses only exercisers who participated exclusively in weight training, cardiovascular equipment, aerobics and running. They found that there were not differences in the “energy” subscale between those exercisers, and this was also the case in our participants. However, for technique imagery, in the aforementioned study there were differences according to program attended (weight-trainers used more technique imagery than exercisers implicating in cardiovascular equipment training or running). This was not the case in the present research, possibly because participants engaged in various group training activities. As a result, they were likely accustomed to receiving feedback from their trainers on technique (Daskalaki, Papadimitriou, & Malliou, 2024; Daskalaki, et al., 2024a; Daskalaki et al., 2024b) and used technique imagery regardless of activity type. Gammage et al. (2000) also reported that runners were using less appearance imagery than the other groups. The fact that the trainees of our study only used more appearance imagery than others when engaged in choreography programs or weight training may be due to the association of these activities with weight loss and muscle toning, respectively (D’Abundo, 2009; Howley, 2018). The fact that this result was not seen in CrossFit exercisers may seem strange, however, “appearance” ideals differ greatly to different CrossFit exercisers, since some of them may prioritize strength over appearance while others pay special attention to their looks (Coyne & Woodruff, 2020; Podmore & Ogle, 2018). Obviously, more research is needed in EI use by activity type, with bigger samples. Therefore, our relative results can be judged as useful but preliminary. Finally, it was also proved that the Greek version of EIQ-AV is a reliable measurement tool, as it was shown in all samples.

Conclusions

The Greek version of EIQ-AV was found to be a valid and reliable measurement tool to be used in Greek exercisers. Its only difference from the original English questionnaire was the deletion of item 1, which however was a choice double-checked through PCA and CFA. Its construct validity was shown by the excellent model fits attained in CFA, the convergent validity at the factor level, and the known groups validity. Regarding the latter, it was established that younger, more frequent and more active exercisers were using significantly more all types of EI in comparison with older, less frequent and less active ones. Additionally, the scores of all subscales of EIQ-AV were positively correlated with strenuous exercise as well as with the total score of GLTEQ, and negatively associated with age, which supports criterion (concurrent) validity. It was also observed that exercisers who included in their activities weight training or choreography classes were using appearance imagery significantly more than those who did not practice such activities. Further, it was found that the Greek EIQ-AV is a reliable measurement tool.

Future researches could further examine its psychometric properties and apply it in persons with different characteristics (e.g. sedentary individuals). In this case, the suitability of item 1 could be re-explored. A limitation of this study was that differences by sex were not explored, due to limited participation of men in the research. Also, the investigation of EI use by activity type was based on only few observations, thus more studies are warranted. On the other hand, the delivery of a reliable and valid Greek translation of EIQ-AV will facilitate the investigation of EI use in the Greek-speaking population, permitting to fill all gaps previously identified.

Disclosure statement No potential conflicts of interest were reported by the authors.

References

- Achjari, D. (2004). Partial least squares: another method of structural equation modeling analysis. *Jurnal Ekonomi dan Bisnis Indonesia*, 19(3), 238-248.
- Afework, T., Wondimagegnehu, A., Alemayehu, N., Kantelhardt, E. J., & Addissie, A. (2021). Validity and reliability of the Amharic version of supportive care needs survey-short form 34 among cancer patients in Ethiopia. *BMC health services research*, 21(1), 484. <https://link.springer.com/article/10.1186/s12913-021-06512-2>
- Arafat, S. M., Chowdhury, H.R, Shalahuddin Qusar, M. M. A. & Hafez, M.A. (2016). Cross Cultural Adaptation and Psychometric Validation of Research Instruments: A Methodological Review. *Journal of Behavioral Health*, 5, 129-136. DOI: 10.5455/jbh.20160615121755
- Bagozzi, R. P., & Yi, Y. (1988). On the Evaluation of Structural Equation Models. *Journal of the Academy of Marketing Science*, 16(1), 74-94. <https://doi.org/10.1007/BF02723327>
- Baharum, H., Ismail, A., Awang, Z., McKenna, L., Ibrahim, R., Mohamed, Z., & Hassan, N. H. (2023). The study adapted instruments based on Confirmatory Factor Analysis (CFA) to validate measurement models of latent constructs. *International Journal of Environmental Research and Public Health*, 20(4), 2860. <https://doi.org/10.3390/ijerph20042860>

- Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J., & Esquivel, S. L. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research, and Evaluation*, 18(1): 6. <https://doi.org/10.7275/qv2q-rk76>
- Bollen, K. A., & Long, J. S. (1993). *Testing structural equation models*. Sage.
- Brown, J. (2009). Choosing the right number of components or factors in PCA and EFA. *JALT Testing & Evaluation SIG Newsletter*, 13(2). https://teval.jalt.org/test/bro_30.htm
- Carter, S. R. (2006). Using confirmatory factor analysis to manage discriminant validity issues in social pharmacy research. *International Journal of Clinical Pharmacy*, 38, 731–737. DOI 10.1007/s11096-016-0302-9
- Cheung, G. W., Cooper-Thomas, H. D., Lau, R. S., & Wang, L. C. (2024). Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations. *Asia Pacific Journal of Management*, 41(2), 745-783. <https://link.springer.com/article/10.1007/s10490-023-09871-y>
- Chung, P. K., & Dong Liu, J. (2012). Examination of the Psychometric Properties of the Chinese Translated Behavioral Regulation in Exercise Questionnaire-2. *Measurement in Physical Education and Exercise Science*, 16(4), 300–315. <https://doi.org/10.1080/1091367X.2012.693364>
- Cid, L., Monteiro, D., Teixeira, D. S., Evmenenko, A., Andrade, A., Bento, T., Vitorino, A., Couto, N., & Rodrigues, F. (2022). Assessment in sport and exercise psychology: Considerations and recommendations for translation and validation of questionnaires. *Frontiers in Psychology*, 13, 806176. <https://doi.org/10.3389/fpsyg.2022.806176>
- Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Lawrence Erlbaum Associates.
- Coyne, P., & Woodruff, S. J. (2020). Examining the influence of CrossFit participation on body image, self-esteem, and eating behaviours among women. *Journal of Physical Education and Sport*, 20(3), 1314-1325. DOI:10.7752/jpes.2020.03183
- Cumming, J. (2008). Investigating the relationship between exercise imagery, leisure time exercise behaviour, and exercise self-efficacy. *Journal of Applied Sport Psychology*, 20, 184-198. <https://doi.org/10.1080/10413200701810570>
- D'Abundo, M. L. (2009). Issues of health, appearance and physical activity in aerobic classes for women. *Sport, Education and Society*, 14(3), 301–319. <https://doi.org/10.1080/13573320903037655>
- Daskalaki, K., Kouri, E., Giannakou E., Beneka, A., Giftofidou, A., Bebetos, E., & Malliou, P. (2020). Greek Translation of the “Exercise Imagery Questionnaire- Aerobic Version”. In V. Malliou, & A. Giftofidou (Eds.). *ePROCEEDINGS Functional Rehabilitation of Injuries in Athletes and Trainees* (pp. 85-89). <https://leidiata.phyed.duth.gr/wp-content/uploads/2021/07/SHORT-PAPERS-2020-4.pdf>
- Daskalaki, K., Papadimitriou, K., Malliou P-V. (2024). Use of mental imagery techniques to transform your fitness classes into a fantastic experience. *ACSMs Health & Fitness Journal*, 29(2), 37-41. DOI: 10.1249/FIT.0000000000000947
- Daskalaki, K., Sakellari, P., Stefanakis, M., Hadjisavvas, S., Efstathiou, E., Mamoukari, P., & Malliou, P. (2024a). The effects of adding imagery to therapeutic exercises on pain reduction, flexibility, functional performance, and quality of life in women with chronic low back pain. *Journal of Physical Education and Sport*, 24(8), 1964-1977. DOI:10.7752/jpes.2024.08218
- Daskalaki, K., Stefanakis, M., Efstathiou, M., Beneka, A., Malliou, P. (2024). Developing the ‘Imagery during Exercise Questionnaire-Gym Version’ and assessing its content validity. *Journal of Physical education and Sport*, 24 (5), 1162 - 1173. DOI:10.7752/jpes.2024.05134. <http://www.efsupit.ro/images/stories/may2024/Art%20134.pdf>
- Daskalaki, K., Stefanakis, M., Hadjisavvas, S., Papadimitriou, K., & Malliou, V.P. (2024b). New findings on imagery use: What do exercisers imagine during gym workouts and why. *Acta Gymnica*, 54, Article e2024.009. <https://doi.org/10.5507/ag.2024.009>
- De Moor, M. H., Distel, M. A., Trull, T. J., & Boomsma, D. I. (2009). Assessment of borderline personality features in population samples: is the Personality Assessment Inventory-Borderline Features scale measurement invariant across sex and age? *Psychological assessment*, 21(1), 125-130. DOI: 10.1037/a0014502
- Ennigkeit, F., & Hänsel, F. (2018). Factorial and convergent validity of the Exercise Identity Scale in a German adult sample. *Measurement in Physical Education and Exercise Science*, 22(4), 343–355. <https://doi.org/10.1080/1091367X.2018.1474113>
- Epstein, J., Osborne R. H., Elsworth, G. R., Beaton, D.E., & Guillemin, F. (2015). Cross-cultural adaptation of the Health Education Impact Questionnaire: experimental study showed expert committee, not back-translation, added value. *Journal of Clinical Epidemiology*, 68(4), 360-369. <https://doi.org/10.1016/j.jclinepi.2013.07.013>
- Farrell, A. M., & Rudd, J. M. (2009). Factor analysis and discriminant validity: A brief review of some practical issues. In *Australia and New Zealand Marketing Academy Conference 2009*. Anzmac.

-
- Flo, J., Landmark, B., Hatlevik, O. E., & Fagerström, L. (2018). Using a new interrater reliability method to test the modified Oulu Patient Classification instrument in home health care. *Nursing open*, 5(2), 167-175. <https://onlinelibrary.wiley.com/doi/10.1002/nop2.126>
- Gammage, K.L., Hall, C.R., & Rodgers, W.M. (2000). More about Exercise Imagery. *Sport Psychologist*, 14, 348-359.
- George, D. , & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference* (4th ed.). Allyn & Bacon.
- Godin, G., & Shephard, R. (1985). A simple method to assess exercise behavior in the community. *Canadian Journal of Applied Sport Sciences*, 10(3), 141-146.
- Hair, J., Black, W., Babin, B., Anderson, R., & Tatham, R. (2019). *Multivariate Data Analysis* (8th Ed.). Pearson Educational.
- Hausenblas, H. A., Hall, C. R., Munroe, K. J., & Rodgers, W. M. (1999). Exercise imagery: Its nature and measurement. *Journal of Applied Sport Psychology*, 11(2), 171-180. <https://doi.org/10.1080/10413209908404198>
- Hermida, R. (2015). The problem of allowing correlated errors in structural equation modeling: concerns and considerations. *Computational Methods in Social Sciences*, 3(1), 5-17. <https://www.cceol.com/search/article-detail?id=417198>
- Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational research methods*, 1(1), 104-121. <https://doi.org/10.1177/109442819800100106>
- Hooper, D. (2012). Exploratory Factor Analysis. In H. Chen (Ed.), *Approaches to Quantitative Research – Theory and its Practical Application: A Guide to Dissertation Students*. Oak Tree Press.
- Howley, E. T. (2018). ACSM's Top 10 Fitness Trends: What's Hot and What's Not: A Look Back. *ACSM's Health & Fitness Journal*, 22(6), 18-23. DOI: 10.1249/FIT.0000000000000427
- Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. DOI: 10.1080/10705519909540118
- Husain, H., & Aziz, H. (2022). Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) to measure the validity and reliability constructs of historical thinking skills, Tpack, and application of historical thinking skills. *International Journal of Education, Psychology, and Counseling*, 7(46), 608-623. DOI: 10.35631/IJEPC.746046
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39, 31-36. <https://link.springer.com/article/10.1007/bf02291575>
- Khairunnisa, A., Dinata, W. W., Rifki, M. S., Komaini, A., Igoresky, A., Asman, A., & Nazri, S. B. M. (2024). Enhancing free throw basketball shooting ability in female basketball players through imagery and concentration exercises. *Journal of Physical Education and Sport*, 24(6), 1426-1433. DOI:10.7752/jpes.2024.06161
- Kim, B. H. & Giacobbi, P. R. Jr. (2009). The Use of Exercise-Related Mental Imagery by Middle-Aged Adults. *Journal of Imagery Research in Sport and Physical Activity*, 4(1). DOI: 10.2202/1932-0191.1031
- Kossert, A. L., & Munroe-Chandler, K. (2007). Exercise imagery: A systematic review of the empirical literature. *Journal of Imagery Research in Sport and Physical Activity*, 2(1). <https://doi.org/10.2202/1932-0191.1015>
- Kosteli, M. C., Williams, S. E., & Cumming, J. (2019). Exploring imagery as a technique for promoting physical activity in older adults. *Imagination, Cognition and Personality*, 38(4), 405-424. <https://doi.org/10.1177/02762366187670>
- Lai, K., & Green, S. B. (2016). The Problem with Having Two Watches: Assessment of Fit When RMSEA and CFI Disagree. *Multivariate Behavioral Research*, 51(2-3), 220-239. <https://doi.org/10.1080/00273171.2015.1134306>
- Malouin, F., Richards, C. L., & Durand, A. (2010). Normal aging and motor imagery vividness: implications for mental practice training in rehabilitation. *Archives of physical medicine and rehabilitation*, 91(7), 1122-1127. <https://doi.org/10.1016/j.apmr.2010.03.007>
- Milne, M. I., Burke, S. M., Hall, C., Nederhof, E., & Gammage, K. L. (2005). Comparing the imagery use of older and younger exercisers. *Imagination, Cognition, and Personality*, 25, 59-67. DOI:10.2190/AJ75-VWT3-9XHM-UJDW.
- Muhamad Safiih, L., & Azreen, N. (2016). Confirmatory Factor Analysis Approach: A Case Study of Mathematics Students' Achievement in TIMSS. *Malaysian journal of mathematical sciences*, 10, 41-51.
- Munroe-Chandler, K. J., Kim, A. J., & Gammage, K. L. (2004). Using Imagery to Predict Weightlifting Dependency in Men. *International Journal of Men's Health*, 3(2), 129-139. <https://doi.org/10.3149/jmh.0302.129>

- Negarandeh, R., Sarkhani, N., & Heydari, M. (2024). Evaluation of cross-cultural adaptation and validation of the Persian version of the health literacy for school-aged children scale: Methodological study. *Health Science Reports*, 7(2), e1876. <https://doi.org/10.1002/hsr2.1876>
- Ouzouni, C. & Nakakis, K. (2011). Validity and Reliability of Measurement Instruments in Quantitative Studies, *Nosileftiki*, 50(2), 231-239.
- Pérez-Fabello, M. J., & Campos, A. (2020). Psychometric properties of the Spanish version of the Exercise Imagery Questionnaire (EIQ). *Cuadernos de Psicología del Deporte*, 20(3), 41-54. <https://www.redalyc.org/journal/2270/227064713004/227064713004.pdf>
- Podmore, M., & Ogle, J. P. (2018). The lived experience of CrossFit as a context for the development of women's body image and appearance management practices. *Fashion and Textiles*, 5, 1-23. <https://doi.org/10.1186/s40691-017-0116-y>
- Rodgers, W. M., Munroe, K. J., & Hall, C. R. (2002). Relations among exercise imagery, self-efficacy, exercise behaviour, and intentions. *Imagination, Cognition and Personality*, 21(1), 55-65.
- Rodgers, W. M., Hall, C. R., Blanchard, C. M., & Munroe, K. J. (1999). The relationship between imagery and obligatory exercise. In *Canadian Psychomotor Learning and Sport Psychology Conference, Edmonton, Canada*.
- Sass, D. A., & Schmitt, T. A. (2010). A Comparative Investigation of Rotation Criteria Within Exploratory Factor Analysis. *Multivariate Behavioral Research*, 45(1), 73-103. <https://doi.org/10.1080/00273170903504810>
- Schreiber, J. B. (2021). Issues and recommendations for exploratory factor analysis and principal component analysis. *Research in Social and Administrative Pharmacy*, 17(5), 1004-1011. <https://doi.org/10.1016/j.sapharm.2020.07.027>
- Sharma, S., & Hussain, M. (2019). Confirmatory Factor Analysis of Hindi Version SPADI Scale in Overhead Athletes with Shoulder Impingement Syndrome. *Journal of Clinical & Diagnostic Research*, 13(7), 12-17. DOI: 10.7860/JCDR/2019/41451.13006
- Shrestha, N. (2021). Factor analysis as a tool for survey analysis. *American journal of Applied Mathematics and statistics*, 9(1), 4-11. DOI: 10.12691/ajams-9-1-2.
- Theodorakis, Y. & Hassandra, M. (2005). Smoking and Exercise, Part II: Differences between Exercisers and Non-Exercisers Inquiries, *Inquiries in Sport & Physical Education*, 3(3), 239 – 248.
- Thøgersen-Ntoumani, C., Cumming, J., Ntoumanis, N., & Nikitaras, N. (2012). Exercise imagery and its correlates in older adults. *Psychology of Sport and Exercise*, 13(1), 19-25. <https://doi.org/10.1016/j.psychsport.2011.08.002>
- Tsang, S., Royse, C.F., & Terkawi, A.S. (2017). Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi Journal of Anaesthesia*, 11(Suppl 1), S80-S89. DOI: 10.4103/sja.SJA_203_17
- Watkins, M. W. (2018). Exploratory Factor Analysis: A Guide to Best Practice. *Journal of Black Psychology*, 44(3), 219-246. <https://doi.org/10.1177/0095798418771807>
- Welis, W., Febrinaldo, A., Ihsan, N., Mario, D. T., Komaini, A., Rifki, M. S., ... & Németh, Z. (2024). The influence of imagery training and self-confidence on shooting skills in futsal: A two-way ANOVA design. *Journal of Physical Education and Sport*, 24(7), 1767-1777. DOI:10.7752/jpes.2024.07197
- Williams, B., Onsman, A., & Brown, T (2010). Exploratory Factor Analysis: A Five-Step Guide for Novices. *Australasian Journal of Paramedicine*, 8, 1-13. DOI:10.33151/ajp.8.3.93
- Wipulanusat, W., Panuwatwanich, K., & Stewart, R. A. (2017). Workplace innovation: Exploratory and confirmatory factor analysis for construct validation. *Management and Production Engineering Review*, 8(2), 57-68. DOI: 10.1515/MPER-2017-0018
- Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in quantitative methods for psychology*, 9(2), 79-94.
- Yusoff, M. S. B. (2011). Psychometric Properties of the Learning Approach Inventory: A Confirmatory Factor Analysis. *Education in Medicine Journal*, 3(2), e24-e31. DOI:10.5959/eimj.3.2.2011.or3
- Zygmunt, C., & Smith, M. R. (2014). Robust factor analysis in the presence of normality violations, missing data, and outliers: Empirical questions and possible solutions. *The Quantitative Methods for Psychology*, 10(1), 40-55.