Cross-cultural invariance of the basic psychological needs in exercise scale and need satisfaction latent mean differences among Greek, Spanish, Portuguese and Turkish samples

Symeon P. Vlachopoulos [a,*], F. Hulya Asci [b], Luís Cid [c], Gozde Ersoz [d], David González-Cutre [e], Juan Antonio Moreno-Murcia [e], João Moutão [c]

[a] Aristotle University of Thessaloniki, Department of Physical Education and Sport Science at Serres, Laboratory of Social Research on Physical Activity, Greece
[b] Baskent University, Faculty of Health Science, Department of Sport Sciences, Ankara, Turkey
[c] Sports Sciences School of Rio Maior (ESDRM) and Research Centre in Sports, Health Sciences and Human Development (CIDESD), Portugal
[d] Suleyman Demirel University, Faculty of Health Science, Department of Sport Sciences, Isparta, Turkey
[e] Miguel Hernández University of Elche, Sports Research Center, Spain

ARTICLE INFO

ABSTRACT

Objectives: The present study examined the extent to which scores from the Basic Psychological Needs in Exercise Scale are measurement invariant across Greek, Spanish, Portuguese, and Turkish samples of exercise participants.
Design: A cross-sectional design wherein responses were analyzed from 504 Greek participants, 518 Spanish participants, 989 Portuguese participants, and 686 Turkish participants.
Methods: Data on exercise participants' need satisfaction were collected using the Basic Psychological Needs in Exercise Scale (Vlachopoulos & Michailidou, 2006) after back translation for the languages involved in the study.
Results: Multi-sample confirmatory factor analyses showed that BPNES scores were largely invariant at the configural, metric, and strong invariance levels, but not at the strict invariance level for Spanish, Portuguese, and Turkish samples compared to the Greek sample. Portuguese participants reported higher levels of satisfaction for all three needs compared to Greek participants, while for Spanish participants this was the case only for competence and relatedness. Turkish participants did not differ from Greek participants.
Conclusions: The present findings support valid cross-cultural comparisons at the latent variance/covariance level and the latent/observed mean level of the constructs of autonomy, competence, and relatedness using the BPNES.

Regular physical activity has proven to be effective in the primary and secondary prevention of several chronic diseases (e.g., cardiovascular disease, diabetes, hypertension, osteoporosis) (Warburton, Nicol, & Bredin, 2006). Despite such evidence, two thirds of the adult population in the European Union, exercise less than the recommended quantity (Cavill, Kahlmeier, & Racioppi, 2006). Quality of motivation to exercise has emerged as an important determinant of objectively assessed exercise participation (Standage, Sebire, & Loney, 2008). This fact highlights the need

to better understand factors that contribute to motivation for exercise behavior to reverse physical inactivity trends. Among theories of motivated behavior, self-determination theory (SDT; Deci & Ryan, 1985, 2008a; Ryan & Deci, 2002) has been used by researchers to provide a better understanding of factors that account for individual differences in exercise behavior (Ryan, Williams, Patrick, & Deci, 2009; Vlachopoulos, 2009; Wilson, Mack, & Grattan, 2008). In SDT, personal and social-environmental factors are outlined that may determine motivated behavior. The theory consists of the six mini-theories of basic needs theory (BNT), cognitive evaluation theory (CET), organismic integration theory (OIT), causality orientations theory (COT), goal content theory (GCT), and the theory of vitality (Ryan et al., 2009; Vansteenkiste, Niemiec, & Soenens, 2010).

* Corresponding author.
E-mail address: vlachop@phed-sr.auth.gr (S.P. Vlachopoulos).

Sub-theories of self-determination theory

In BNT the existence of three innate psychological needs for autonomy, competence, and relatedness, deemed to be necessary for psychological, physical, and social health is addressed (Vansteekiste et al., 2010). Satisfaction of the need for autonomy reflects a sense of volition and self-endorsement in one's behavior. When autonomy is satisfied, behavior is experienced as choiceful and owned, emanating from the self, and in accord with abiding interests and values (Vansteekiste et al., 2010). When the need for competence is satisfied, individuals feel that they interact effectively with one's environment and experience opportunities to express or develop their capacities. When the need for relatedness is satisfied, individuals feel that they are authentically associated with significant others and experience a sense of belonging (Ryan & Deci, 2002). In SDT, the dimensions of the social environment that support the three needs are also discussed. These dimensions are labeled autonomy support, competence support, and relatedness support (Vansteenkiste et al., 2010). In CET, the factors are discussed that either undermine or support intrinsic motivation. This may happen when external events either thwart the psychological needs for autonomy and competence or support them, respectively (Deci & Ryan, 1985). In OIT, the process of internalization is described and the endorsement of the value of extrinsically motivated behaviors is explained. Extrinsically motivated behaviors are those behaviors where the activity is perceived as a means to a separable outcome while interest and enjoyment are absent (Vansteekiste et al., 2010). Four types of extrinsic motivation have been distinguished in OIT. These are external regulation where behavior is energized by forces external to the individual; introjected regulation where behavior is energized by self-imposed pressure; identified regulation that reflects a sense of choice in enacting the behavior; and integrated regulation where behavior is enacted because it has become part of individuals' identity.

Identified and integrated regulations represent types of self-determined extrinsic motivation that are enhanced via the satisfaction of the needs for autonomy, competence, and relatedness (Vallerand, 2001). In COT, the existence of three types of orientations that partly determine functioning at the personality level is addressed. These are the autonomy orientation, the control orientation, and the impersonal orientation. These orientations develop as a result of received need support in conjunction with a range of genetic and biological factors (Vansteenkiste et al., 2010). In GCT the concepts of intrinsic and extrinsic life goals or aspirations that people pursue are delineated, and how they differentially affect well-being. Intrinsic goals tend to contribute to the satisfaction of the basic needs for competence, autonomy, and relatedness while extrinsic goals are likely unrelated to them. Further, the sub-theory of vitality is a theory of subjective energy addressing factors that may enhance it or deplete it (Ryan et al., 2009). Vitality is assumed to be heavily affected by psychological factors in the sense that thwarting of the psychological needs for autonomy, competence, and relatedness may diminish vitality while need satisfactions may enhance it or maintain it (Ryan et al., 2009).

SDT has provided the genesis of origin for the Hierarchical Model of Intrinsic and Extrinsic Motivation (Vallerand, 2001) where satisfaction of the needs for autonomy, competence, and relatedness, are assumed to lead to greater levels of autonomous motivation, and lower levels of controlled motivation, and amotivation. Further, psychological needs are held to be the mediators of the effects of socio-contextual factors (e.g., fitness instructors' need supportive behaviors) on levels of self-determination in one's behavior (Vallerand, 2001). Ryan (1995) and Vallerand (1997) have emphasized the importance of conducting domain-specific SDT research claiming that use of domain-specific scales is necessary in

explaining and predicting behavior accurately across various life domains. Following this assertion, much SDT-related exercise research has been conducted showing that psychological need satisfaction has been positively associated with global and contextual well being and positive affect (Vlachopoulos & Karavani, 2009; Wilson, Longley, Muon, Rodgers, & Murray, 2006; Wilson, Mack, Blanchard, & Grey, 2009), enjoyment/interest, attitude toward exercise, intention to exercise, internal locus of control and frequency of exercise participation (Vlachopoulos, 2007; Vlachopoulos & Michailidou, 2006), life satisfaction among overweight/obese exercise participants (Edmunds, Ntoumanis, & Duda, 2007) and exercise behavior (Edmunds, Ntoumanis, & Duda, 2006; Vlachopoulos & Neikou, 2007).

Basic psychological need satisfaction measurement in exercise

The effects of psychological need fulfillment in exercise have been investigated in a number of studies (Kirkland, Karlin, Stellino, & Pulos, 2011; Vlachopoulos, 2007, 2008; Vlachopoulos, Kaperoni, & Moustaka, 2011; Vlachopoulos & Karavani, 2009; Vlachopoulos & Michailidou, 2006; Vlachopoulos & Neikou, 2007; Vlachopoulos, Ntoumanis, & Smith., 2010) using either the Basic Psychological Needs in Exercise Scale (BPNES: Vlachopoulos & Michailidou, 2006) or the Psychological Need Satisfaction in Exercise Scale (PNSE: Wilson et al., 2009; Wilson & Rogers, 2008; Wilson, Rodgers, Blanchard, & Gessell, 2003; Wilson, Rogers, Rodgers, & Wild, 2006). The instruments have been designed by different research groups attending to a number of construct validation steps advocated by measurement experts (Messick, 1995) and appear promising in the quest to capture variation in perceived need satisfaction in exercise (Wilson, Mack, Gunnell, Oster, & Gregson, 2008). The PNSE is an 18-item scale with six items per need factor. It was developed using responses of Canadian exercise participants while the BPNES is a 12-item scale with four items per factor developed using responses from Greek exercise participants. For both instruments, evidence has emerged to support their structural validity; this includes gender invariance and criterion-related validity based on a number of variables residing within the SDT nomological network (Wilson, Mack, Gunnell, et al., 2008).

Cross-cultural generalizability and measurement invariance

According to Duda and Allison (1990), scientific inquiry in the area of sport and exercise psychology requires the incorporation of cross-cultural analyses in order to avoid the propagation of theories that may not be generalizable. This is because culture and ethnicity may be important variables in explaining the variability in cognitions, affect, and behavior in psychological theories of sport and exercise (Duda & Allison, 1990; Duda & Hayashi, 1998). SDT is a particularly good example of a theory that has been studied with reference to its cross-cultural applicability. A central assertion of SDT has been that the basic psychological needs for autonomy, competence, and relatedness, are universal in their importance and their effects. This means that need satisfaction promotes psychological health, independent of cultural context (Deci & Ryan, 2008b). Such a claim stands in contrast to the view held by cultural-relativists (e.g., Markus, Kitayama, & Heiman, 1996) who claim that needs are learned within cultures and that autonomy, for instance, represents a western ideal taught in western cultures that emphasize individualism rather than collectivism and interdependence. This means that autonomy may not be so important in terms of people's well being in eastern cultures and may not play an important role in the lives of these individuals (Chirkov, 2012). SDT holds that cultures do influence people in important ways but that all humans have certain psychological needs that need to get

satisfied to experience optimal well-being independent of culture (Deci & Ryan, 2008b). However, the practices used to satisfy these needs may differ across cultures. Consistent evidence has now emerged supportive of the universality of autonomy and its relationship with positive psychological health in various domains including the academic domain (Chirkov, 2009; Chirkov & Ryan, 2001), subjective well being (Rudy, Sheldon, Awong, & Tan, 2007; Schmuck, Kasser, & Ryan, 2000; Sheldon et al., 2004), and work (Deci et al., 2001).

Cross-cultural comparisons testing the equivalence of scores derived from research instruments are necessary in testing the cross-cultural applicability of theories and models (Sue, 1999). Measurement invariance reveals the extent to which item responses preserve their meaning across groups (Millsap & Kwok, 2004). This is essential when psychological measures are used in group comparison. Comparing groups using a non-invariant measure is worthless or, put differently, is analogous to comparing apples and oranges. In such a case, there may be potential confounding and misinterpreting of mean differences (Hoyle & Smith, 1994). When measurement invariance does not hold, it may be that differences in instrument scores may not reflect true differences on the construct being measured, rather differential functioning of the instrument across the groups (Millsap & Kwok, 2004). Two types of measurement invariance have been identified (Little, 1997). Category 1 invariance refers to the psychometric properties of the scale scores such as configural invariance indicating that the groups associate the same items with the same constructs, invariance of item loadings (i.e., metric invariance), invariance of both item loadings and item intercepts (i.e., strong invariance), and concurrent invariance of item loadings, item intercepts, and item error variances (i.e., strict invariance; see Gregorich, 2006). Category 2 invariance refers to between-group differences in variances, covariances, and latent means. According to Cheung and Rensvold (2002), Category 1 invariance should hold before interpreting Category 2 differences that usually reflect more substantive research questions. For instance strong invariance should hold before latent mean differences are examined across groups.

The present study

Taking into account the line of research that addresses the cross-cultural applicability of SDT principles and the need to further understanding of the cross-cultural generalizability of SDT tenets in exercise, Vlachopoulos, Ntoumanis, and Smith (2010) examined the extent of measurement equivalence of the BPNES scores across Greek and British exercise participants. They provided initial evidence of generalizability of the scale scores. However, expanding the generalizability evidence base of SDT theoretical tenets in the exercise domain would require evidence of item meaning equivalence across samples reflecting cultural diversity. These cultures may differ on their individualistic vs. collectivistic emphasis; hence, following suit with the line of cross-cultural research generated in the broader SDT literature (e.g., Chirkov, 2009).

The primary purpose of the present study was to examine the extent of measurement invariance of the BPNES scores in cultural groups differing in their emphasis on individualistic and collectivistic values. Greek, Spanish, Portuguese, and Turkish samples of exercise participants were studied using responses of Greek exercise participants as the reference point. This was the case as the BPNES has been originally developed with Greek exercise participants. Research examining differences in cultural orientations across nations has shown that individuals tend to vary in the degree to which they endorse individualist and collectivist values (Hofstede, 1983). Individuals of Western European cultures such as Spain tend to more strongly endorse individualist values and are

hypothesized to describe themselves as autonomous, independent, self-contained, and more differentiated from others. In contrast, individuals in North and Eastern European nations such as Estonia, Sweden, and Turkey tend to more strongly endorse collectivist values, and perceive themselves as more interdependent and socially sensitive (Markus & Kitayama, 1991). Findings of meta-analyses of cross-cultural research (Oyserman, Coon, & Kemmelmeier, 2002) have provided support in favor of Hofstede's work showing relatively stronger individualism effects in Western European nations such as Spain and France in comparison to countries of the Eastern and Northern Europe (e.g., Estonia, Turkey, Finland, Norway). Moreover, stronger effects for collectivism were obtained in Northern European nations in comparison to Western European nations such as Spain and Eastern European nations such as Turkey. This research demonstrates the diversity of the samples presently studied in terms of their emphasis on individualist and collectivist values. In the present study, it was hypothesized that BPNES scores would emerge as invariant at the configural, metric, strong, and strict invariance levels (Gregorich, 2006) across the culturally diverse samples studied.

A secondary aim was to explore possible need satisfaction mean differences across samples from these cultures. In relation to the secondary aim of the study, it was hypothesized that exercise samples from Western countries in which individualistic values are strongly endorsed, such as Spain and Portugal, would report higher levels of autonomy, competence, and relatedness need satisfaction when compared to non-Western exercise samples such as those from Greece and Turkey.

Method

Participants

Greek participants

For Greek exercise participants we used the validation sample employed by Vlachopoulos and Michailidou (2006). The sample consisted of 504 individuals comprising 246 men (48.8%) and 258 women (51.2%). Participants' age ranged from 18 to 65 years (M_{age} = 28.92 yr, SD = 8.45 yr). Their exercise experience ranged between a few months and 45 years (M = 6.88 yr, SD = 6.91 yr) while weekly frequency of exercise attendance ranged between 1 and 7 times per week (M = 3.74, SD = 1.07). There were 109 participants involved in group-based exercise activities (21.6%), 280 involved in resistance training activities (55.6%), and 108 participants (21.4%) involved in a combination of the aforementioned activities.

Spanish participants

There were 518 Spanish exercise participants ranging in age between 18 and 60 years (M_{age} = 29.96 yr, SD = 8.83) comprising 268 men (51.7%) and 246 women (47.5%) with four individuals not providing details of their gender (.8%). There were 222 participants involved in instructor-led activities (42.9%), 114 participants involved in self-paced programs that were designed but not led by an instructor (22%), and 182 participants practising individual activities without supervision of any kind (e.g., squash, free swimming; 32.1%). Participants exercised between 1 and 3 times per week (M = 2.47, SD = .59) from 20 to 260 min (M = 77.57, SD = 31.4 min) per visit to the fitness centre.

Portuguese participants

There were 989 Portuguese exercise participants including 451 men (45.6%) and 538 women (54.4%) aged between 18 and 70 years (M_{age} = 32.67 years, SD = 10.76). Of them, 270 (27.3%) exercised in group-type classes, 275 in strength training activities (27.8%), and

444 (44.9%) participants in cardio-fitness activities. They have exercised in total between 1 and 300 months ($M = 25.88$ months, $SD = 35.85$) and between 1 and 8 times per week ($M = 3.18$, $SD = 1.17$).

Turkish participants

There were 686 Turkish participants including 384 men (56%) and 302 (44%) women ranging in age between 18 and 60 years ($M_{age} = 30.78$ years, $SD = 8.64$). There were 236 (34.4%) participants involved in group-based classes and 450 participants (65.6%) involved in individual type activities. Participants exercised between 1 and 240 months ($M = 25.74$, $SD = 22.29$) and between 1 and 7 times per week ($M = 3.15$, $SD = 1.07$) for a duration of 30 min–180 min per visit to the fitness center ($M = 67.87$, $SD = 20.55$). Means, standard deviations, Cronbach's α values, and bivariate correlations between scale items for each sample separately can be consulted as Supplementary information linked to the present article.

Instruments

Basic psychological needs in exercise

The Basic Psychological Needs in Exercise Scale (BPNES; Vlachopoulos & Michailidou, 2006) was used to measure the extent to which the psychological needs for autonomy, competence, and relatedness are fulfilled in the exercise context. The BPNES consists of 12 items measuring each of the need constructs using four items. Participants are requested to indicate the degree of their agreement with each of the items on a 5-point Likert scale anchored by 1 (I don't agree at all) and 5 (I completely agree). Validity and reliability evidence has been provided in support of the scale's factor structure, internal consistency, stability over four weeks, correlations with external variables, and measurement invariance across gender and ethnicity using responses provided by British exercise participants (Vlachopoulos, 2007, 2008; Vlachopoulos et al., 2011; Vlachopoulos & Karavani, 2009; Vlachopoulos & Michailidou, 2006; Vlachopoulos & Neikou, 2007; Vlachopoulos et al., 2010). Further support of the scale's psychometric properties have been derived from Portuguese (Moutão, Cid, Alves, Leitão, & Vlachopoulos, 2012) and Spanish (Sánchez & Núñez, 2007) samples.

For the Turkish, Portuguese, and Spanish samples back translation techniques were employed to develop language-specific versions of the BPNES. The translation techniques followed a standardized procedure suggested by Brislin (1986) in which the inventory items and scale were translated from English into the target language by a bilingual researcher. Thereafter, the translated inventory was back-translated by a jury of independent academics at the institutions of each of the present authors; these academics were all bilingual. The back-translated versions were then compared with the original English version and any inconsistencies, errors, biases, and indications of incongruence were highlighted before further item wording modifications took place to derive the final version.

Procedures

Data were collected from non-probability samples of exercise participants obtained from cultural groups that varied in the characteristics of individualism and collectivism. Participants were recruited at the reception area of the fitness centers by research assistants. The assistants explained the purpose of the study and requested that participants complete self-report questionnaires before embarking upon their exercise activity. Data used in the present study have been collected mainly for validation purposes of the BPNES scores in the languages involved along with other

questionnaires. The questionnaire took approximately 15 min to complete. Data were collected every day of the week, either morning or evening, under the close supervision of a research assistant. Following completion of each questionnaire, the research assistant checked for missing data. In the case of missing responses, the assistant prompted the participant to respond to the relevant item(s). Data collectors adhered to the American Psychological Association ethical guidelines. Participants signed informed consent forms and were assured that their responses would be anonymized and kept in confidence. At each testing site, the regulations of the associated university were adhered to.

Data analysis

Confirmatory factor analysis

The factor structure of the BPNES was examined for each cultural sample using confirmatory factor analysis (CFA) before we embarked on measurement invariance analyses. For each CFA model, the factor variances were fixed to 1.0, the factor covariances were allowed to be estimated and item error covariances were fixed to zero. Model fit was assessed using the chi-square value (χ^2), the Comparative Fit Index (CFI; Bentler, 1990), and the Root Mean Square Error of Approximation (RMSEA; Steiger & Lind, 1980) accompanied by its 90% confidence interval (90% CI). Given that the χ^2 value is sensitive to sample size (Byrne, 2006), model fit assessment was based mainly on the remaining fit indexes. CFI values close to .95 indicate an excellent fit to the data (Hu & Bentler, 1999) with values of .90 or greater indicating a reasonable fit (Kline, 2005; Marsh, Hau, & Wen, 2004). An RMSEA value less than .06 denotes a good model fit (Hu & Bentler, 1999) while values between .08 and .10 represent an adequate fit (Browne & Cudeck, 1993; Byrne, 2000; Kline, 2005). Scale reliability was calculated based on the composite reliability index for each BPNES subscale that reflects the proportion of shared variance to error variance in a construct (Li, Harmer, & Acock, 1996) with a value that should not be lower than .60 (Bagozzi & Yi, 1988) and Cronbach's (1951) alpha.

Multi-group invariance

According to Steenkamp and Baumgartner (1998), the most powerful approach to testing cross-national measurement invariance has been multi-group confirmatory factor analysis. Given the aforementioned distinction between Category 1 and Category 2 tests of measurement invariance (Cheung & Rensvold, 2002; Little, 1997), the present analyses focused mainly on examining the extent of Category 1 cross-national invariance of BPNES scores. The following types of invariance were examined (Gregorich, 2006): configural (without equality constraints; Model 1), metric invariance (constraining factor loadings equal; Model 2), strong invariance (constraining both factor loadings and item intercepts equal concurrently; Model 3), and strict invariance (constraining factor loadings, item intercepts, and item residuals concurrently equal; Model 4).

Model 1 (i.e., configural) was a multi-sample model based on the a priori correlated 3-factor CFA BPNES model without equality constraints imposed on any of the model parameters. The hypothesis tested was that the same number of common factors would be present in each group and each factor would be associated with the same set of items. In Model 2 (i.e., metric invariance), item loadings were set to be invariant across groups. Evidence in support of this model would imply that the BPNES factors have the same meaning across groups. For Model 3 (i.e., strong invariance), item loading constraints were retained and item intercept equality constraints were added to the model. This model examined whether mean comparison across groups was defensible. In Model

Table 1
Goodness-of-fit indexes of CFA model for cultural samples.

CFA model	χ^2	S–B χ^2	df	p	Robust CFI	Robust RMSEA	Robust RMSEA 90% CI	Standardized factor loadings	Composite reliability		
									Auton.	Comp.	Relat.
Greek sample	122.28	96.38	51	.001	.980	.042	.029–.055	.60–.89	.85	.80	.91
Spanish sample	298.53	218.36	51	.001	.906	.080	.069–.091	.53–.82	.69	.77	.85
Portuguese sample	264.26	199.63	51	.001	.954	.054	.046–.062	.53–.80	.73	.72	.83
Turkish sample	284.57	199.71	42	.001	.912	.074	.064–.084	.64–.76	.78	.73	.80

Note. Greek sample ($n = 504$; Vlachopoulos & Michailidou, 2006; validation sample), Spanish sample ($n = 518$), Portuguese sample ($n = 989$), Turkish sample ($n = 686$). S–B χ^2 = Satorra–Bentler scaled χ^2; df = degrees of freedom; CFI = Comparative Fit Index; RMSEA = Root Mean Squared Error of Approximation. [a] For the Turkish sample the Competence #1 item has been removed owing to a low item loading of .27 and the factor correlation between Autonomy and Competence was fixed at .85 to achieve CFA model convergence. Auton. = Autonomy; Comp. = Competence; Relat. = Relatedness. Cronbach's alpha values for autonomy, competence, and relatedness were: Greek sample = .84, .81, .92; Spanish sample: .69, .77, .85; Portuguese sample: .73, .72, .82; Turkish sample: .78, .72[a], .81.

4 (i.e., strict invariance), equality constraints were added on item residuals while retaining item loading and item intercept equality constraints. This model examined whether cross-group comparisons of observed item or composite score variances were meaningful (Gregorich, 2006). In line with Cheung and Rensvold (2002), determining invariance in multi-sample testing using the $\Delta \chi^2$ value has been considered unsatisfactory; therefore, evaluation of multi-group CFA nested models was based on (a) the degree of adequate model fit exhibited by the more constrained model, and (b) the ΔCFI value between the two nested models. Values not exceeding .01 would indicate invariance (Cheung & Rensvold, 2002).

Latent mean differences

Mean and covariance structure analyses were used to test for latent mean differences for each need satisfaction construct. The latent mean value for the Greek participants was always constrained to zero, while it was freely estimated for the other cultural sample. The z statistic was used to determine statistical significance between the latent means. A comparison between latent mean differences was possible only after support for the strong invariance multi-group model.

Results

There was no missing data for the BPNES items in any of the samples. The normalized estimate of Mardia's coefficient of multivariate kurtosis was greater than 5.0 in all samples indicating multivariate non-normality (Byrne, 2006). Therefore, the ML robust method was used with EQSWIN 6.1 (Bentler, 2003) that provides the Satorra–Bentler Scaled χ^2 (S–B χ^2) (Satorra & Bentler, 2001). S–B χ^2 is corrected for non-normality and may be accompanied by the robust CFI, RMSEA and its 90% CI, respectively. For the Turkish sample, Competence #1 item was removed due to a low item loading of .27 and the factor correlation between Autonomy and Competence was fixed at .85 to achieve CFA model convergence. Such alterations did not take place for the Greek sample, either for the CFA model testing or the invariance models testing. The indices corrected for non-normality were adequate for all four samples (Greek, Spanish, Portuguese, and Turkish) indicating that the correlated 3-factor model for BPNES responses was a good fit to the data in all four cultural groups. The goodness-of-fit-indexes along with composite reliability and Cronbach's alpha coefficients for each BPNES subscale are presented in Table 1.

In terms of the measurement invariance models, multi-group model comparison (e.g., configural model vs. metric invariance model) was based on the ΔCFI value which was expected to be .01 or less to indicate no difference between the models. A comparison of the responses between Greek and Spanish exercise participants revealed evidence of metric and strong invariance but not strict invariance. The comparison of responses between Greek and Portuguese exercise participants revealed that the Relatedness #1 and 2 items and the Autonomy #4 item were not associated with

Table 2
Overall fit indexes for the cross-cultural measurement invariance basic psychological needs in exercise scale models.

Model	χ^2	S–B scaled χ^2	df	Robust CFI	ΔCFI	Robust RMSEA	Robust RMSEA 90% CI
Greece - Spain							
Model 1: Configural invariance	420.83	319.29	102	.947	–	.046	.040–.051
Model 2: Metric invariance	454.29	348.35	114	.943	.004	.045	.039–.050
Model 3: Strong invariance	697.15	563.83	120	.942	.001	.060	.055–.065
Model 4: Strict invariance	1004.21	776.75	129	.897	.045	.070	.065–.075
Greece – Portugal							
Model 1: Configural invariance	386.55	298.21	102	.964	–	.036	.031–.041
Model 2: Metric invariance	690.14	518.72	114	.926	.038	.049	.045–.053
Model 2a: Metric invariance (omitting Relatedness #1)	541.37	410.24	113	.945	.019	.042	.038–.046
Model 2b: Metric invariance (omitting Relatedness #1 and 2)	496.86	377.75	112	.951	.013	.040	.035–.044
Model 2c: Metric invariance (omitting Relatedness #1 and 2, and Autonomy #4)	470.38	358.57	111	.955	.009	.039	.034–.043
Model 3: Strong invariance	591.33	473.85	116	.962	–	.045	.041–.050
Model 4: Strict invariance	685.14	529.78	123	.951	.011	.047	.043–.051
Greece – Turkey							
Model 1: Configural invariance	406.86	303.79	93	.949	–	.044	.038–.049
Model 2: Metric invariance	460.20	347.71	104	.941	.008	.044	.039–.050
Model 3: Strong invariance	474.80	372.03	109	.938	.003	.045	.040–.050
Model 4: Strict invariance	836.88	672.45	117	.873	.065	.063	.059–.068

Note. Greek sample ($n = 504$; Vlachopoulos & Michailidou, 2006; validation sample), Spanish sample ($n = 518$), Portuguese sample ($n = 989$), Turkish sample ($n = 686$). χ^2 = chi-square; S–B χ^2 = Satorra–Bentler scaled χ^2; df = degrees of freedom; CFI = Comparative Fit Index; RMSEA = Root Mean Squared Error of Approximation. For the Turkish sample the Competence #1 item has been removed owing to a low item loading of .27; the factor correlation between autonomy and competence for the Turkish sample was fixed at .85 to achieve model convergence.

Table 3
Differences between Greek and comparison country exercise participants on latent need satisfaction constructs.

Cultural sample comparison	Latent mean difference analyses								
	Autonomy			Competence			Relatedness		
	z	diff.	ES	z	diff.	ES	z	diff.	ES
Greece − Turkey	−.69	−.03	.06	.70	.03	.07	.78	.04	.05
Greece − Spain	−.10	−.00	.00	4.24	.19*	.47	5.78	.34*	.85
Greece − Portugal	7.57	.32*	1.13	8.32	.30*	1.06	5.56	.30*	1.06

Note. z = z statistic value; diff. = difference between the two latent means; ES = magnitude of effect size of the difference between the two latent means based on Hancock (2001). Effect sizes have been interpreted in accordance with Cohen's d values of .2, .5, and .8 for a small, medium, and large effect size, respectively. * Significant at $p < .05$.

invariant item loadings. After removing equality constraints for these items, strong invariance was attained but not strict invariance. Removing equality constraints means that a revised multi-group model is specified. In this new model, the item loadings for the Relatedness #1 and 2 items and the Autonomy #4 item are not restricted to be equal as in the previous metric invariance multi-group model tested. This means that the new model is bound to have a fit better than the fit of the previous metric invariance model, given that fewer requirements are posed now on the new model. This newly revised metric invariance model is pitted once more against the original configural invariance model. Regarding the comparison of responses between Greek and Turkish exercise participants, the findings supported both metric and strong factorial invariance, but not strict invariance (Table 2).

Latent mean differences analysis showed that only Portuguese exercise participants reported significantly higher levels of autonomy need satisfaction compared to Greek exercise participants. Greek individuals did not differ from Turkish and Spanish exercise participants. In regard to competence need satisfaction, it was Spanish and Portuguese participants that reported significantly greater levels of need satisfaction compared to Greek individuals. Turkish exercise participants did not differ from Greek participants. In terms of relatedness, Spanish and Portuguese exercise participants reported greater levels of need satisfaction compared to Greek participants. Greek participants did not differ from Turkey participants in terms of the satisfaction of the relatedness need in exercise (Table 3).

Discussion

The present study aimed to expand the evidence base concerning generalizability of the Basic Psychological Needs in Exercise Scale scores (BPNES: Vlachopoulos & Michailidou, 2006) across diverse cultural samples that we assumed to differ on the dimensions of individualism and collectivism. The study was conducted in the context of recent debates about the universality of the beneficial impact of autonomy and autonomy support on psychological health across nations and cultures (Deci & Ryan, 2008b). Opponents of this theoretical position hold the cultural deterministic view that the construct of autonomy along with other cultural values such as individualism, liberalism, independence, and self-reliance are "Western constructions". According to this view, such constructions may not be applicable to societies that are described as less individualistic and more collectivistic or group-oriented (Markus et al., 1996). That is, the meaning of autonomy may be differentially negotiated in various socio-cultural contexts.

A number of studies that have directly assessed the statistical invariance of SDT-based instruments of autonomous motivation across cultures, supported the linguistic meaningfulness, and

applicability of these SDT-based operationalizations of autonomy to individuals from various nations, societies, and ethno-linguistic groups (Chirkov, 2009). Additional evidence has also supported the universally beneficial role of autonomy in academic learning and human functioning, in studies conducted among Western and non-Western nations, varying in collectivism and other cultural dimensions, including Chinese learners, students in South Korea, Taiwan, USA, Belgium, Britain, Canada, German, and Italy (see Chirkov, 2009, 2012 for overviews). Following this SDT-based research perspective, the aim of the present study was to examine whether statistical invariance would hold for the BPNES, a short measure of need satisfaction in exercise, across cultural samples of exercise participants, varying in the characteristics of individualism and collectivism. Evidence in support of equivalent item meaning interpretation across these cultural groups, would provide grounds for continued cross-cultural research examining the SDT perspective of the universality of the positive relationship of autonomy and basic need satisfaction with markers of psychological health and well being in the exercise domain.

The samples of exercise participants presently examined included Spanish, Portuguese, and Turkish individuals using responses from Greek exercise participants as the reference point given that the scale has been originally developed in Greek. Initial evidence on the generalizability of the BPNES responses across cultural samples using Greek exercise participants as the reference point has been provided by Vlachopoulos et al. (2010) who compared the Greek responses to those of British exercise participants. These researchers provided evidence of metric invariance (equality of factor loadings) and strong invariance (equal item loadings and item intercepts) after removing in both cases the Relatedness #2 item equality constraint, respectively. Further, strict invariance was demonstrated when the Relatedness #2, 4 and Autonomy #4 error variance equality constraints were also removed.

Configural invariance

Given the need to examine the generalizability of theories and test their applicability in various cultural groups, the present study built upon the initial evidence derived from British exercise participants to expand upon other cultural groups such as Spanish, Portuguese, and Turkish. In general, the findings did support equivalence of the BPNES item parameters for Spanish, Portuguese, and Turkish participants using responses of Greek participants as the reference point. Specifically, for all three samples, evidence emerged in support of configural invariance. That is, the constructs of autonomy, competence, and relatedness need satisfaction in exercise, were conceptualized in the same way across the groups (Cheung & Rensvold, 2002). Configural invariance may fail when the cultural context differentially affects participants' perceptions about the construct being measured, and participants attach different meanings to the construct (Millsap & Everson, 1991; Riordan & Vandenberg, 1994; Tayeb, 1994) or when errors are made in translation (Cheung & Rensvold, 2002). Therefore, the present support obtained in favor of configural invariance leads to the conclusion that participants of different cultural groups may have defined the construct in the same way.

Metric invariance

In terms of equivalence of factor loadings, metric invariance was supported for all three cultural groups except in the case of the Portuguese sample. This means that Portuguese participants are likely to interpret item meaning differently regarding Relatedness #1, ("I feel comfortable with the people I exercise with"),

Relatedness #2 ("My relationships with the people I exercise with are very friendly"), and Autonomy #4 ("I feel that I have the opportunity to make choices with regard to the way I exercise") items. For instance, it may be that exercise participants from different cultural groups hold a differential definition of the meaning of positive relationships between exercise participants. By way of illustration of this, Portuguese individuals may place a stronger emphasis on individualist values when compared to Greeks. It may be that such an orientation leads to different criteria that are applied to judging the positive nature of relationships among exercise participants when compared to Greek individuals who may place greater value on collectivism and interdependence. That is, the degree of positivity of personal relationships among Portuguese exercise participants may not be an important criterion for a satisfying exercise session when compared to Greeks, who may place a higher value on interdependence. Clearly, figuring out the nuances of such a difference would require delving into the deeper meanings of this phenomenon using qualitative techniques. These allow for such information to be extracted and a deeper understanding of this behavioral domain to be achieved. Invariance of item loadings corresponds to a stronger test of the equivalence of item meaning across groups, and is the minimum prerequisite for meaningful group comparisons of any type (Bollen, 1989). Metric invariance supports valid group comparisons of latent factor variances, covariances, and regression coefficients between latent variables (Gregorich, 2006).

Strong invariance

Strong measurement invariance (i.e., invariance of item loadings and item intercepts concurrently; Gregorich, 2006) was also attained for all three cultural groups with Greek responses as the reference point. This type of invariance supports valid cross-group latent mean and observed mean comparisons (Gregorich). Lack of strong invariance means that variables unrelated to the common factors measured by the instrument may lead to lower or higher item responses in one group compared to the other (Gregorich). When strong invariance holds, observed group means are held to be unbiased estimates of group differences in the latent factor means (Gregorich). Therefore, under strong factorial invariance, meaningful cross-group comparisons between either observed group means or latent factor means may be performed. This is the case for Spanish and Turkish exercise participants while for the Portuguese participants the Relatedness #1 and 2, and the Autonomy #4 item may not be appropriate for inclusion in such a comparison. Nonetheless, the decision is dependent on the measurement properties of the items based on the particular samples under investigation. The fact that three items out of the 12, were not found invariant at the metric level, and as a consequence at next levels of the invariance sequence (i.e., strong invariance), does not invalidate the use of the scale in measuring need satisfaction, both within each cultural sample and across cultural samples. For instance, in the case of metric non-invariance, two factor loadings may emerge as significantly different. In the present data, a closer examination of the magnitude of the difference, points to a very small difference. Such a finding indicates that the difference is trivial despite statistical significance. Clearly, in order to decide on the utility of particular BPNES items in performing valid cross-group comparisons, systematic and substantial evidence should firstly be collected using a number of samples.

Strict invariance

In terms of strict factorial invariance (i.e., concurrently constrained item loadings, item intercepts, and item error variances)

the hypothesis was not supported in any of the cultural groups. Strict invariance supports cross-group comparisons of observed item or composite score variances and covariances. This is the case because such comparisons should entirely correspond with common factor variation and not be confounded by differences in residual variance (Gregorich, 2006). Lack of strict factorial invariance may be attributed to differences in vocabulary, grammar, idioms or syntax (Malpass, 1977). Based on the present findings, cross-group comparisons of observed variances between responses obtained from Greek and either Spanish, Portuguese, or Turkish participants are not warranted. However, this is not considered a serious problem given that group comparisons of observed variances are not really useful as they do not take measurement error into account.

Implications for construct validity and theory development

Measurement experts (Marsh, 1997; Messick, 1995) have maintained that construct validation is at the heart of the measurement process recommending the implementation of systematic programs of construct validation that combine applications of relevant theory (in this case BNT which is a subtheory of SDT) with available statistical evidence to inform score interpretations. In terms of measurement of basic psychological need satisfaction in exercise using the BPNES, the present study expands the currently minimal available evidence attesting to the invariance of the BPNES scores (Wilson, Mack, Gunnell, et al., 2008). Indeed, in a systematic review of the evidence supporting construct validity for the two instruments capturing variation in need satisfaction in exercise (i.e., BPNES: Vlachopoulos & Michailidou, 2006; PNSE: Wilson, Longley, et al., 2006; Wilson, Rogers, et al., 2006), Wilson et al. concluded that systematic evidence in support of the generalizability of the BPNES and PNSE scores is lacking.

It was also noteworthy that two items of the relatedness subscale and one item of the autonomy subscale were found metrically non-invariant for the Portuguese sample, but not for the Spanish and Turkish samples. Relatedness #2 item scores have also been found to be metrically non-invariant when comparing Greek BPNES responses to responses of British exercise participants (Vlachopoulos et al., 2010). In the present study, such differences in item interpretation may be due to cultural differences that affect the meaning attached to the positivity of relationships between exercise participants in the Portuguese society. Hence, qualitative work aiming to further understanding of the meaning attached to particular items that function differentially across cultures is recommended. Omitting items when testing for invariance of structural parameters (e.g., a regression coefficient) may have implications for the generalizability of this parameter depending on the items used. Nonetheless, keeping only the items with equivalent item loadings is more advantageous as it reduces bias in assessing whether a regression coefficient is not different across the groups. Further, removing items when testing for statistical invariance does not have any influence on the conceptual integrity of the scale, because these items are removed only during invariance testing. Further, Competence #1 item had a low factor loading for the Turkish sample and was dropped from the present analyses, hence not allowing a comparison with the respective factor loading in the Greek sample. Clearly, further scrutiny of this particular item would be required to determine its effectiveness as an indicator of the construct that it is purported to represent.

Further a high correlation emerged between the autonomy and competence factors for the Turkish sample, leading to the use of a constraint to achieve CFA model convergence. High correlations between these factors have, in general, been observed in the need satisfaction measurement literature in exercise both across different nationalities (Greek and Canadian exercise participants)

and using different need satisfaction instruments (either the BPNES or the PNSE). Hence, the generalizable nature of the magnitude of this correlation may require further investigation, both from a theoretical and a methodological viewpoint.

A certain extent of variability was also observed in the composite reliability and Cronbach's alpha coefficients of the BPNES across cultural samples. Despite this variability, all coefficients attained acceptable values demonstrating satisfactory levels of reliability. It may be possible that such variability in item coherence is also a function of differences in the cultural context under which responses have been provided. Scale responses may be influenced by person characteristics interacting with the context/environment under which responses are provided. Hence, future research might examine the possible sources of variance in the psychometric properties of the BPNES scores associated with individualistic/ collectivistic cultures. For instance, the reliability generalization method may be used to explore variability in score reliability estimates and pinpoint the possible sources of this variance (Vacha-Haase, Henson, & Caruso, 2002).

Overall, evidence has emerged for the statistical invariance of the BPNES responses in exercise across the diverse cultural samples that have been examined. The present results concur with those of similar studies in the broader cross-cultural SDT literature. Such literature addresses the statistical and linguistic invariance of scales measuring autonomy support in various samples differing in the cultural dimensions of individualism and collectivism (Chirkov, 2009). The present findings open new avenues in SDT-based cross-cultural research in relation to the effects of perceived autonomy support on psychological health and human functioning in the exercise domain. The demonstrated cross-cultural equivalence of BPNES item meaning may facilitate further tests of the universality of the effects of psychological need satisfaction on psychological health and well-being. Such research may be conducted across cultures differing in the norms and values they endorse.

Need satisfaction latent mean differences among cultural groups

The hypothesis advanced for the secondary aim of the present study was that the Western countries such as Spain and Portugal, would evidence significantly higher mean scores on autonomy, competence, and relatedness need satisfaction in exercise compared to the non-Western countries of Greece and Turkey. Such a hypothesis was advanced because people in collectivist cultures may already satisfy psychological needs to a certain extent in other important life domains that may compensate for need satisfaction in the exercise domain. Given the systematic nature of these effects, it may be that in Western societies that value interdependence and social connectedness less, individuals have a greater need to experience competence, autonomy, and a social bonding with others, simply because these needs are indeed "nutriments" for psychological health (Deci & Ryan, 2008a). It was found that Portuguese participants reported greater satisfaction for all three needs compared to the Greek individuals, and this was also the case for the Spanish in reference to competence and relatedness. No significant differences emerged between Greek and Turkish samples for any of the psychological needs. However, these hypotheses are tentative and a more systematic approach would be required to uncover the complex interactions between culture, psychological need satisfaction, psychological health, and health-related behavior such as exercise participation.

Limitations and future directions

The present findings are limited to exercise participants belonging to the cultural groups under investigation; namely, Greek, Spanish, Portuguese, and Turkish. The findings cannot be considered representative and be generalized to the populations studied, at a national level. Such generalization would have necessitated the use of random sampling procedures. A second limitation of the present findings relates to the nature of the activity that participants were engaged in; that is, exercise in the context of a fitness center. Individuals exercising in different exercise contexts (e.g., at home, in parkland, on a beach, etc.) have not been included in the study. Third, the samples comprised healthy adults excluding other categories of participants such as older individuals or individuals with chronic disease. Fourth, latent mean differences results should be interpreted in the context of an indirect rather than a direct assessment of participants' cultural orientation. Presently, determination of the cultural orientation of the samples studied was based on Hofstede's (1983) cultural assessments of the respective countries from which the samples have been drawn. However, given inter-individual variability in cultural orientation within a country's population, future research should employ a direct assessment of participants' cultural orientation. Rather than being based on an average estimation of the participants' cultural orientation, as was the case in the present investigation, a direct assessment would clearly be beneficial in facilitating a better understanding of this complex set of interrelationships.

The extent of measurement equivalence found in BPNES item scores allows for the use of the BPNES with the above mentioned types of exercise participants. Such use may involve examining potential equivalence in psychological processes posited within SDT that involves the satisfaction of the needs for autonomy, competence, and relatedness in exercise. More specifically, cross-cultural comparisons may be performed to examine cross-group equivalence of latent factor variances, covariances, and regression coefficients between latent variables, thus testing more substantive hypotheses, elsewhere labeled as Category 2 invariance hypotheses (Little, 1997). In addition, meaningful cross-cultural group comparisons followed by detailed theory-driven hypotheses may be performed on need satisfaction constructs either at the latent mean level or the observed mean level, given the attainment of strong invariance for BPNES items. The lack of strict invariance for all of the present cultural group comparisons precludes the valid examination of group differences in observed variances and covariances given inequalities in the item error variances across the samples. Also, future studies examining the relationship of culture (i.e., individualism and collectivism) with need satisfaction, would be well advised to take direct measures of the cultural orientation of the sample, given inter-individual variability within cultural samples; this would determine participants' cultural orientation more accurately. Overall, the present findings support the use of the BPNES in exercise-related SDT-based research involving cross-cultural group comparisons of the constructs of autonomy, competence, and relatedness. Such comparisons may facilitate a cross-cultural examination of broader psychological processes that are germane to SDT and posited to influence exercise behavior.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.psychsport.2013.03.002.

References

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science, 16*, 74–94. http://dx.doi.org/10.1007/BF02723327.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238–246.

Bentler, P. M. (2003). *EQS 6.1 for Windows [Computer software]*. Encino, CA: Multivariate, Software.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Brislin, R. W. (1986). The wording and translation of research instruments. In W. J. Lonner, & J. W. Berry (Eds.), *Field methods in educational research* (pp. 137–164). Newbury Park, CA: Sage.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & S. J. Long (Eds.), *Testing structural equation models* (pp. 445–455). Newbury Park, CA: Sage.

Byrne, B. M. (2000). *Structural equation modeling with AMOS: Basic concepts, applications, and, programming*. Mahwah, NJ: Lawrence Erlbaum Associates.

Byrne, B. M. (2006). *Structural equation modeling with EQS: Basic concepts, applications, and, programming* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Cavill, N., Kahlmeier, S., & Racioppi, F. (Eds.). (2006). *Physical activity and health in Europe: Evidence for action*). Copenhagen, Denmark: World Health Organization Regional Office for Europe. Retrieved July 25, 2012 from: http://www.euro.who.int/__data/assets/pdf_file/0011/87545/E89490.pdf.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indices for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*, 233–255. http://dx.doi.org/10.1207/S15328007SEM0902_5.

Chirkov, V. I. (2009). A cross-cultural analysis of autonomy in education: a self-determination theory perspective. *Theory and Research in Education, 7*, 253–262. http://dx.doi.org/10.1177/1477878509104330.

Chirkov, V. I. (2012). Dialectical relationships among human autonomy, the brain, and culture. In A. Delle Fave (Series Ed.) & V. I. Chirkov, R. M. Ryan, & K. M. Sheldon (Vol. Eds.), *Human autonomy in cross-cultural context: Perspectives on the psychology of agency, freedom, and well-being: Vol. 1. Cross-cultural advancements in positive psychology* (pp. 65–91). New York: Springer.

Chirkov, V. I., & Ryan, R. M. (2001). Parent and teacher autonomy-support in Russian and U.S., adolescents: common effects on well-being and academic motivation. *Journal of Cross-Cultural Psychology, 32*, 618–635.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297–334.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Deci, E. L., & Ryan, R. M. (2008a). Self-determination theory: a macrotheory of human motivation, development, and health. *Canadian Psychology, 49*, 182–185. http://dx.doi.org/10.1037/a0012801.

Deci, E. L., & Ryan, R. M. (2008b). Facilitating optimal motivation and psychological well-being across life's domains. *Canadian Psychology, 49*, 14–23. http://dx.doi.org/10.1037/0708-5591.49.1.14.

Deci, E. L., Ryan, R. M., Gagne, M., Leone, D. R., Usunov, J., & Kornazheva, B. P. (2001). Need, satisfaction, motivation, and well-being in the work organizations of a former eastern bloc country: a cross-cultural study of self-determination. *Personality and Social Psychology Bulletin, 27*, 930–942.

Duda, J. L., & Allison, M. T. (1990). Cross-cultural analysis in exercise and sport psychology: a void in the field. *Journal of Sport and Exercise Psychology, 12*, 114–131.

Duda, J. L., & Hayashi, C. (1998). Measurement issues in cross-cultural research within sport and exercise psychology. In J. L. Duda (Ed.), *Advances in sport and exercise psychology measurement* (pp. 471–483). Morgantown: Fitness Technology.

Edmunds, J., Ntoumanis, N., & Duda, J. L. (2006). A test of self-determination theory in the exercise domain. *Journal of Applied Social Psychology, 36*, 2240–2265. http://dx.doi.org/10.1207/S15328007SEM0902_5.

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*, 722–740. http://dx.doi.org/10.1016/j.psychsport.2006.07.006.

Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care, 44*, S78–S94. http://dx.doi.org/10.1097/01.mlr.0000245454.12228.8f.

Hancock, G. R. (2001). Effect size, power, and sample size determination for structured means modeling and MIMIC approaches to between-group hypothesis testing of means on a single latent construct. *Psychometrika, 66*, 373–388.

Hofstede, G. (1983). Dimensions of national cultures in fifty countries and three regions. In J. B. Deregowski, S. Dziuraciec, & R. C. Annis (Eds.), *Explications in cross-cultural psychology* (pp. 335–355). Lisse, Netherlands: Swets & Zeitlinger.

Hoyle, R. H., & Smith, G. T. (1994). Formulating clinical research hypotheses as structural equation models: a conceptual overview. *Journal of Consulting and Clinical Psychology, 62*, 429–440. http://dx.doi.org/10.1037/0022-006X.62.3.429.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.

Kirkland, R. A., Karlin, N. J., Stellino, M. B., & Pulos, S. (2011). Basic psychological needs, satisfaction, motivation, and exercise in older adults. *Activities, Adaptation and Aging, 35*, 181–196. http://dx.doi.org/10.1080/01924788.2011.596764.

Kline, R. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: The, Guilford Press.

Li, F., Harmer, P., & Acock, A. (1996). The task and ego orientation in sport questionnaire: construct equivalence and mean differences across gender. *Research Quarterly for Exercise and Sport, 68*, 228–238.

Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: practical and theoretical issues. *Multivariate Behavioral Research, 32*, 53–76.

Malpass, R. S. (1977). Theory and method in cross-cultural psychology. *American Psychologist, 32*, 1069–1079.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion and, motivation. *Psychological Review, 98*, 224–253.

Markus, H. R., Kitayama, S., & Heiman, R. J. (1996). Culture and basic psychological principles. In E. T. Higgins, & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 857–913). New York: Guilford Press.

Marsh, H. W. (1997). The measurement of physical self-concept: a construct validation approach. In K. R. Fox (Ed.), *The physical self: From motivation to well-being* (pp. 27–58). Champaign, Il: Human Kinetics.

Marsh, H. W., Hau, K., & 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, 11*, 320–341.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*, 741–749.

Millsap, R. E., & Everson, H. (1991). Confirmatory measurement model comparisons using latent means. *Multivariate Behavioral Research, 26*, 479–497.

Millsap, R. E., & Kwok, O. (2004). Evaluating the impact of partial factorial invariance on selection in two populations. *Psychological Methods, 9*, 93–115. http://dx.doi.org/10.1037/1082-989X.9.1.93.

Moutão, J., Cid, L., Alves, J., Leitão, J., & Vlachopoulos, S. P. (2012). Validation of the basic psychological needs in exercise scale in a Portuguese sample. *Spanish Journal of Psychology, 15*, 399–409.

Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin, 128*, 3–52. http://dx.doi.org/10.1037//0033-2909.128.1.3.

Riordan, C. M., & Vandenberg, R. J. (1994). A central question in cross-cultural research: do employees of different cultures interpret work-related measures in an equivalent manner? *Journal of Management, 20*, 643–671.

Rudy, D., Sheldon, K. M., Awong, T., & Tan, H. H. (2007). Autonomy, culture, and well-being: the benefits of inclusive autonomy. *Journal of Research in Personality, 41*, 983–1007. http://dx.doi.org/10.1016/j.jrp.2006.11.004.

Ryan, R. M. (1995). Psychological needs and the facilitation of integrative processes. *Journal of, Personality, 63*, 397–427.

Ryan, R. M., & Deci, E. L. (2002). An overview of self-determination theory: an organismic- dialectical perspective. In E. L. Deci, & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 3–33). Rochester, NY: The University of Rochester Press.

Ryan, R. M., Williams, G. C., Patrick, H., & Deci, E. L. (2009). Self-determination theory and, physical activity: the dynamics of motivation in development and wellness. *Hellenic Journal of Psychology, 6*, 107–124.

Sánchez, J., & Núñez, J. (2007). Análisis preliminar de las propiedades psicométricas de la versión española de la Escala de Necesidades Psicológicas Básicas en el Ejercicio Físico [Preliminary analysis of the psychometric properties of a Spanish version of the Basic Psychological Needs in Exercise Scale]. *Revista Iberoamericana de Psicología del Ejercicio y el Deporte, 2*, 83–92.

Satorra, A., & Bentler, P. M. (2001). A scaled differences chi-square test statistic for moment structure analysis. *Psychometrika, 66*, 507–514.

Schmuck, P., Kasser, T., & Ryan, R. M. (2000). Intrinsic and extrinsic goals: their structure and relationship to well-being in German and U.S. college students. *Social Indicators Research, 50*, 225–241. http://dx.doi.org/10.1023/A:1007084005278.

Sheldon, K. M., Elliot, A. J., Ryan, R. M., Chirkov, V., Kim, Y., Wu, C., et al. (2004). Self-concordance and subjective well-being in four cultures. *Journal of Cross-Cultural Psychology, 35*, 209–223.

Standage, M., Sebire, S. J., & Loney, T. (2008). Does exercise motivation predict engagement in objectively assessed bouts of moderate-intensity exercise? A self-determination theory perspective. *Journal of Sport & Exercise Psychology, 30*, 337–352.

Steenkamp, J. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research, 25*, 78–90.

Steiger, J. H., & Lind, J. M. (1980, June). Statistically based tests for the number of common factors. In *Paper presented at the meeting of the Psychometric Society, Iowa City, IA.*

Sue, S. (1999). Science, ethnicity, and bias: where have we gone wrong? *American Psychologist, 54*, 1070–1077.

Tayeb, M. (1994). Organizations and national culture: methodology considered. *Organization, Studies, 15*, 429–446.

Vacha-Haase, T., Henson, R. K., & Caruso, J. C. (2002). Reliability generalization: moving toward improved understanding and use of score reliability. *Educational and Psychological Measurement, 62*, 562–569. http://dx.doi.org/10.1177/0013164402062004002.

Vallerand, R. J. (1997). Toward a hierarchical model of intrinsic and extrinsic motivation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 271–360). New York, NY: Academic Press.

Vallerand, R. J. (2001). A hierarchical model of intrinsic and extrinsic motivation in sport and exercise. In G. C. Roberts (Ed.), *Advances in motivation in sport and exercise* (pp. 263–319). Champaign, IL: Human Kinetics.

Vansteenkiste, M., Niemiec, C. P., & Soenens, B. (2010). The development of the five mini-theories of self-determination theory: an historical overview, emerging trends, and future directions. In T. C. Urdan, & S. A. Karabenick (Vol. Eds.). *Advances in motivation and achievement: Vol. 16. The decade ahead: Theoretical perspectives on motivation and achievement* (pp. 105–165). Emerald Group Publishing Company. http://dx.doi.org/10.1108/S0749-7423(2010)000016A007.

Vlachopoulos, S. P. (2007). Psychometric evaluation of the basic psychological needs in exercise scale in community exercise programs: a cross-validation approach. *Hellenic Journal of Psychology, 4*, 52–74.

Vlachopoulos, S. P. (2008). The basic psychological needs in exercise scale: measurement invariance over gender. *Structural Equation Modeling, 15*, 114–135. http://dx.doi.org/10.1080/10705510701758398.

Vlachopoulos, S. P. (Ed.). (2009). Self-determination theory, physical activity and well-being [special, issue]. *Hellenic Journal of Psychology, 6*(2).

Vlachopoulos, S. P., Kaperoni, M., & Moustaka, F. C. (2011). The relationship of self-determination theory variables to exercise identity. *Psychology of Sport and Exercise, 12*, 265–272. http://dx.doi.org/10.1016/j.psychsport.2010.11.006.

Vlachopoulos, S. P., & Karavani, E. (2009). Psychological needs and subjective vitality in exercise: a cross-gender situational test of the needs universality hypothesis. *Hellenic Journal of Psychology, 6*, 207–222.

Vlachopoulos, S. P., & Michailidou, S. (2006). Development and initial validation of a measure of autonomy, competence, and relatedness in exercise: the basic psychological needs in exercise scale. *Measurement in Physical Education and Exercise Science, 10*, 179–201. http://dx.doi.org/10.1207/s15327841mpee1003_4.

Vlachopoulos, S. P., & Neikou, E. (2007). A prospective study of the relationships of autonomy, competence, and relatedness with exercise attendance,

adherence, and dropout. *Journal of Sports Medicine and Physical Fitness, 47*, 475–482.

Vlachopoulos, S. P., Ntoumanis, N., & Smith, A. L. (2010). The basic psychological needs in exercise scale: translation and evidence for cross-cultural validity. *International Journal of Sport and Exercise Psychology, 8*, 394–412.

Warburton, D. E. R., Nicol, C. W., & Bredin, S. S. D. (2006). Health benefits of physical activity: the evidence. *Canadian Medical Association Journal, 174*, 801–809. http://dx.doi.org/10.1503/cmaj.051351.

Wilson, P. M., Longley, K., Muon, S., Rodgers, W. M., & Murray, T. C. (2006). Examining the contributions of perceived psychological need satisfaction to well-being in exercise. *Journal of Applied Biobehavioral Research, 11*, 243–264. http://dx.doi.org/10.1111/j.1751-9861.2007.00008.x.

Wilson, P. M., Mack, D. E., Blanchard, C. M., & Grey, C. E. (2009). The role of perceived psychological need satisfaction in exercise-related affect. *Hellenic Journal of Psychology, 6*, 183–206.

Wilson, P. M., Mack, D. E., & Grattan, K. P. (2008a). Understanding motivation for exercise: a self- determination theory perspective. *Canadian Psychology, 49*, 250–256. http://dx.doi.org/10.1037/a0012762.

Wilson, P. M., Mack, D. E., Gunnell, K., Oster, K., & Gregson, J. P. (2008b). Analyzing the measurement of psychological need satisfaction in exercise contexts: evidence, issues, and future directions. In M. P. Simmons, & L. A. Foster (Eds.), *Sport and exercise psychology, research advances* (pp. 361–391). Hauppauge, NY: Nova Science.

Wilson, P. M., Rodgers, W. M., Blanchard, C. M., & Gessell, J. (2003). The relationship between psychological needs, self-determined motivation, exercise attitudes, and physical fitness. *Journal of Applied Social Psychology, 33*, 2373–2392.

Wilson, P. M., & Rogers, W. T. (2008). Examining relationships between perceived psychological need satisfaction and behavioural regulations in exercise. *Journal of Applied Biobehavioral Research, 13*, 119–142. http://dx.doi.org/10.1111/j.1751-9861.2008.00031.x.

Wilson, P. M., Rogers, W. T., Rodgers, W. M., & Wild, T. C. (2006). The psychological need satisfaction in exercise scale. *Journal of Sport and Exercise Psychology, 28*, 231–251.