Psychometric Properties of the Short Form of the Physical Self-Description Questionnaire in a

French Adolescent Sample

Christophe Maïano[a, b*], Alexandre J. S. Morin[b*], Nicolas Mascret[c]

[a] Cyberpsychology Laboratory, Department of Psychoeducation and Psychology, Université du Québec en Outaouais (UQO), Canada.

[b] Institute for Positive Psychology and Education, Australian Catholic University, Australia.

[c] Aix Marseille Université, CNRS, ISM UMR 7287, 13288, Marseille, France.

* The first two authors (C. M. and A. J. S. M) contributed equally to this article and their order was determined at random: both should be considered first authors.

Corresponding author: Christophe Maïano, Université du Québec en Outaouais, Campus de Saint-Jérôme, Départment de Psychoéducation et de Psychologie, 5 rue St-Joseph, Saint-Jérôme (Québec), J7Z 0B7, email: christophe.maiano@uqo.ca

Abstract

Recently, Marsh, Martin, and Jackson (2010) developed a short form of the Physical Self-Description Questionnaire (PSDQ-S). The objective of this study was to examine the construct validity and reliability of the PSDQ-S in a French adolescent sample. The sample used in this study included 587 adolescents (247 boys, 340 girls, M_{age} = 14.62). Confirmatory factor analyses (CFA) provided support for the factor validity, reliability, and convergent validity of the French version of the PSDQ-S, and the strict measurement invariance of PSDQ-S across sex, age, body mass index, and involvement or not in sport practice. However, the latent means of the PSDQ-S did not prove to be invariant across sex, body mass index, and involvement or not in sport practice. Our findings suggest that the French version of the PSDQ-S presents acceptable psychometric properties and may be confidently used in research or practice to assess the physical self-conceptions of French adolescents.

Keywords: sex; age; sport practice; short form; convergent validity; measurement invariance.

In their seminal work, Fox and Corbin (1989) adapted Shavelson, Hubner, and Stanton's (1976) multidimensional and hierarchical self-concept model to the physical self-concept area. In this model, the higher level is occupied by the global self-concept (i.e., overall positive or negative perception and/or assessment of oneself). The intermediate "domain" level is occupied by a global construct representing the global physical self-concept (i.e., positive or negative perception and/or assessment of oneself in the global physical area). Finally, the lower "subdomain" level is occupied by constructs representing more specific components of physical self-conceptions, such as sport competence, physical condition, physical strength, flexibility, coordination, and physical appearance.

To operationalize this model, Marsh, Richards, Johnson, Roche, and Tremayne (1994) developed and validated the Physical Self-Description Questionnaire (PSDQ). The PSDQ was initially developed for adolescents and included a total of 70 items assessing 11 dimensions: activity, appearance, body fat, endurance, coordination, flexibility, health, sport competence, strength, global physical self-concept, and global self-esteem. Confirmatory factor analyses (CFA) conducted on a sample of 710 Australian high school students supported the factorial validity and measurement invariance of the PSDQ across male and female students, as well as the convergent and discriminant validity of the PSDQ with other physical self-concepts instruments (Marsh et al., 1994). Subsequent analyses demonstrated that the subscales showed satisfactory scale score reliability coefficients ranging from $\alpha = .82$ to .96 across samples, and had sex-based mean-level differences, with boys tending to have higher scores on most PSDQ subscales compared to girls (Marsh et al., 1994). Additional studies conducted among three Australian adolescent samples revealed: (a) good convergent and discriminant validity of the PSDQ with external criteria (e.g., body composition, physical fitness tests; Marsh & Redmayne, 1994; Marsh, 1996b), and; (b) good test-retest stability of the 11 scales (with test-retest correlations ranging from $r = .70$ to .89 over a 3-month period, and between $r = .31$ and .82 over a 14-month period; Marsh, 1996a).

Although other instruments are currently available for the assessment of physical self-conceptions among young persons, the PSDQ remains by far the most comprehensive (covering 11 dimensions versus 6-7 for the other instruments) and the most widely-validated instrument available to date (for reviews see Marsh & Cheng, 2012; Sypsa & Simons, 2008). Indeed, the psychometric

properties of the PSDQ (i.e., factor validity and reliability, convergent and discriminant validity) have been extensively documented (for a meta-analysis see Schipke & Freund, 2012) across: (a) language versions (French, Dutch, German, Hebrew, Italian, Spanish, Turkish, etc.), (b) age groups (children, adolescents, adults, elderly), (c) type of sample (normal, psychiatric, athletes, bodily handicapped, etc.), and (d) context (research, diagnostic, counselling, etc.).

Among its few downsides, the PSDQ is much longer (70 items vs. 12-to 35 items) than alternatives, especially when used in conjunction with multiple other instruments in the context of comprehensive studies (Marsh, Martin, & Jackson, 2010). This limitation led Marsh, Martin et al. (2010) to develop and validate a 40-item short form of the PSDQ (PSDQ-S), balancing brevity and psychometric strength. The PSDQ-S was developed and validated using a normative archival sample ($N = 1605$) of Australian high school students (12-18 years old) and was cross-validated among five additional samples of Australian students ($N = 708$), Australian elite adolescent athletes ($N = 349$), Spanish adolescents ($N = 986$), Israeli university students ($N = 395$), and Australian older adults ($N = 760$). The PSDQ-S covers all subscales included in the original PSDQ, each of which is assessed using either three (Appearance, body fat, endurance, flexibility, global physical self-concept, sport competence, strength), four (Activity) or five (Coordination, global self-esteem, and health) items. A series of CFAs provided support for the (a) factorial validity and measurement invariance of the PSDQ-S across samples, sex-groups, age-groups, and versions (PSDQ vs. PSDQ-S); (b) convergent and discriminant validity of the PSDQ-S among samples of Australian adolescents ($N = 322$) and Israeli University students ($N = 395$); and (c) sex- and age-based differences, showing that boys and adolescents tend to present higher scores on most PSDQ subscales when compared with girls and older adults, respectively. Subsequent analyses demonstrated satisfactory scale score reliability ($\alpha = .77$ to .94, across the six samples) and test-retest reliability ($r = .57$ to .90 over 1-year among 212 Australian adolescents and 553 older adults, respectively) of the 11 subscales.

Adaptations to Other Languages

To date, several studies have provided tentative support to the psychometric properties of various linguistic adaptations of the PSDQ-S. For instance, Agarwal, Bhalla, Kaur, and Babbar (2013) adapted the PSDQ-S for Indian university students and reported acceptable levels of scale score

reliability for the various subscales (α ranging from .77 to .97). Likewise, starting from the German adaptation of the full PSDQ (Stiller & Alfermann, 2007), Jones and Stumbrys (2014) created a short form using the same items as Marsh, Martin et al. (2010), and administered it to a sample of 72 university students. Although these authors do not report information about the scale score reliability or factor validity of the PSDQ-S, their results revealed significant associations between these subscales, psychological well-being, and the frequency of lucid dreaming. In a more extensive study, Papaioannou et al. (2013; also see Duda et al., 2013) administered the global self-esteem subscale of the PSDQ-S to a total of 7,789 early adolescent soccer players from France, Greece, Norway, Spain, and England. Across countries/languages, their results revealed modest-to-acceptable scale score reliability estimates for this subscale (α ranging from .51 to .73) and significant relationships with subjective vitality and moderate vigorous physical activity. Similarly, Castonguay, Sabiston, Crocker, and Mack (2014) used two subscales (i.e., appearance and body fat) of the PSDQ-S among a mixed sample of 435 English- and French-speaking Canadian adults. Their results reveal that both subscales had acceptable scale score reliability (α = .86 and .92) and were significantly related to subscales of the Body and Appearance-related self-conscious Emotions Scale.

Although promising, these studies remain preliminary, as no study has yet systematically examined the psychometric properties of these linguistic adaptations. Thus, apart from initial attempts by Marsh, Martin et al. (2010) to validate the PSDQ-S among samples of Spanish and Israeli students, the cross-cultural generalizability of the PSDQ-S among samples of non-English adolescents remains an open question that must be addressed before the PSDQ-S can be confidently used in the context of cross-cultural comparison studies. In this study, we focus on the French version of the PSDQ-S.

Currently, two self-report questionnaires are available to measure the physical self-concept among French-speaking young persons: the long form of the PSDQ (Guérin, Marsh, & Famose, 2004) and the short and very short forms of the Physical Self-Inventory (PSI-S & PSI-VS; Maïano et al., 2008; Morin & Maïano, 2011a, 2011b). Despite the numerous advantages of the PSI scales (see Marsh & Cheng, 2012), their reduced length also makes them less comprehensive than the PSDQ. In turn, the length of the PSDQ makes it unwieldy in many research contexts where the number of items needs to be limited for practical reasons. Consequently, the validation of a French version of the

PSDQ-S will fill an important gap in providing a reasonably short, yet comprehensive, measure of the physical self-concept for the French-speaking research community. However, the usefulness of such a French version goes well beyond the French community in providing a valuable instrument on which to anchor cross-cultural comparison studies (Tomás, Marsh, González-Romá, Valls, & Nagengast, 2014). In this regard, the availability of a validated French version is particularly important given that French is the official or co-official language in 29 countries and territories worldwide and one of the most commonly used language in North Africa.

Measurement Invariance and Latent Means Differences

A critical issue in the assessment of the psychometric properties of any measurement instrument is whether it can be used with individuals coming from different segments of the population and whether comparisons conducted across these subpopulations will be meaningful or reflect measurement biases. Measurement biases occur when an instrument behaves differently across distinct subgroups from the population and leads to the impossibility of comparing scores obtained on the instrument across these distinct subgroups. In practice, this verification is conducted via a sequence of tests of measurement invariance (e.g., Meredith, 1993; Millsap, 2011) where equality constraints are progressively added to different parameters from a measurement model (i.e., loadings, intercepts, and uniqueness) across subgroups of participants to systematically test whether these constraint hold in practice. The non-invariance of the factor loadings suggests that the instrument does not measure the same constructs across subgroups, and precludes any form of group-based comparison. The non-invariance of items' intercepts rather suggests that participants presenting the same true score on the construct of interest (e.g., physical self-conceptions) will still tend to score higher or lower on the measurement scales as a result of their membership in specific subgroups. Evidence of invariance of the factor loadings and item intercept is an important pre-requisite to validate group-based mean-level comparisons. The non-invariance of the items' uniquenesses finally suggests that the measurement errors differ as a function of group membership, and thus that the constructs are assessed with different levels of precision.

Although Marsh, Martin et al. (2010) reported evidence of measurement invariance across subsamples of boys and girls, adolescents and older adults, and typical adolescents versus elite

athletes, they did not examine the measurement invariance of the PSDQ-S across samples of early and late adolescents, adolescents involved or not in sport practice at a more normative level[1], and as a function of body mass index (BMI) levels. Furthermore, it remains necessary to verify whether evidence of sex-based measurement invariance will generalize to the French PSDQ-S. These observations are worrisome as the PSDQ – and by extension the PSDQ-S – is frequently used to compare physical self-conceptions as a function of BMI levels and subgroups of boys and girls, early and late adolescents, and youth involved or not in sport practice.

Pending evidence of measurement invariance, it then becomes possible to conduct more advanced tests of the invariance of the latent variances, covariances, and means across these same subgroups of participants, providing a direct test of the discriminant validity of the French PSDQ-S. Indeed, if the psychometric properties of the original PSDQ and PSDQ-S are maintained in the French version, then group-based mean-level differences obtained with the French PSDQ-S should replicate those observed in previous investigations of physical self-conceptions among adolescent populations. Generally, previous studies of physical self-conceptions conducted among samples of adolescents have revealed that (e.g., Bowker, 2006; Findlay & Bowker, 2007; Marsh, 1998; Marsh, Hau, Sung, & Hu, 2007; Marsh, Hey, Roche, & Perry, 1997; Schmalz & Davison, 2006): (a) boys tend to have higher levels of physical self-conceptions than girls across most dimensions of the physical self-concept that were considered; (b) early adolescents tend to have higher levels of physical self-conceptions (across most dimensions) than middle, and late adolescents; (c) adolescents with the highest BMIs tend to have the lower levels of physical self-conceptions (except for the strength subscale on which they presented higher levels); and (d) adolescents athletes tended to present higher levels of physical self-conceptions (across most dimensions) than nonathletes.

Therefore, the main objective of the present study was to examine the psychometric properties of the French PSDQ-S among a sample of adolescents. More specifically, this study examined the: (a) factorial validity and scale score reliability of the PSDQ-S among a sample of French adolescents; (b) measurement and latent mean invariance of the PSDQ-S across sex-groups (boys versus girls), age-groups (early versus late adolescence), sport practice involvement (i.e., involved versus not involved), and BMI; and (c) convergent validity of the PSDQ-S with the PSI-S, and its invariance across sex-

groups, age-groups, and sport practice involvement.

Method

Sample and Procedures

A sample of 587 adolescents (age$_{range =}$ 11-18 years; M_{age} = 14.62 years, SD = 1.80) attending seven middle and high schools located in Southern France form the sample was used in this study. Of those: (a) 247 were boys (42.1%) and 340 were girls (57.9%); (b) 262 (44.6%) were early adolescents (age$_{range}$: 11-14 years) and 340 (55.4%) were late adolescents (age$_{range}$: 15-18 years); and (c) 218 (37.1%) were involved in physical education (PE) and did not practice sport outside school and 369 (62.9%) were involved in PE and practiced sport outside school. Only adolescents who returned consent forms signed by themselves and their parents were included in the study. All adolescents that were included in this study completed the questionnaires during PE classes. This project was approved by the French Advisory Committee on Information Processing in Material Research in the Field of Health and by the Chief Education Officer of the Académie of Aix-Marseille.

Measures

Demographics. Participants completed a questionnaire in which they were asked to report their sex, age, height, weight, and involvement in sport practice outside of the school context (i.e., "Do you practice a sport outside of your physical education classes and of the school sport association?"). Self-reported height and weight were used to compute participant BMI (weight/height^2).

Physical self-conceptions. Multidimensional physical self-conceptions were measured using a French version of the PSDQ-S and of the PSI-S (Maïano et al., 2008; Morin & Maïano, 2011a). The French version of the PSDQ was developed and validated by Guérin et al. (2004) among a sample of 1,040 French adolescents. In two studies, these authors provided support for the 11-factor structure, reliability (i.e., scale score reliability and test-retest stability), and convergent validity (with 13 physical fitness criteria) of the PSDQ. This French version includes 70 items that are rated on a 6-point scale ranging from 1 (*False*) to 6 (*True*). To build the short version used in the current study, we retained from these 70 items the same 40 items selected by Marsh, Martin et al. (2010) to develop the English PSDQ-S.

The PSI-S is a French adaptation of the PSPP (Fox & Corbin, 1989), which has been

extensively validated and cross-validated with a sample of 3,047 French adolescents in two studies,
which supported the factor structure, scale score reliability, test-retest reliability, and convergent
validity of the PSI-S (Maïano et al., 2008; Morin & Maïano, 2011a). The PSI-S includes 18 items
rated on a 6-point scale ranging from 1 (*Not at all*) to 6 (*Entirely*)] and assessing 6 dimensions (global
self-worth, physical self-worth, physical condition, sport competence, physical attractiveness, and
physical strength) of the physical self-concept.

Analyses

The analyses were conducted with Mplus 7.11 (Muthén & Muthén, 2013). Because of
significant multivariate non-normality of the data (Mardia's normalized estimate = 348.20), analyses
were performed using the robust maximum likelihood estimator (MLR). In addition, full-information
estimation was used to correct for the small amounts of missing data present at the item level
(Missing$_{range}$ = 3.24% to 6.47%; $M_{missing}$ = 4.43%; Enders, 2010). In the first stage of the analyses, the
a priori 11-factor model of the PSDQ-S was examined using CFA (Model 1-1 in Table 1),
hypothesizing that: (a) youths' responses to the PSDQ-S would be explained by 11 correlated factors;
(b) each item would have a non-zero loading on the factor it was designed to measure and zero
loadings on other factors; and (c) uniquenesses would not be correlated. Moreover, because one-
fourth of the items of the PSDQ-S are negatively worded (i.e., five items from the health subscale,
three items from the body fat subscale, and two items from the global self-esteem subscale), an
additional model was examined in order to control for the methodological artifact introduced by the
wording of these items. Indeed, evidence is rapidly accumulating that negatively worded items tend to
be accompanied by methodological artifacts that need to be explicitly taken into account in statistical
models (e.g., DiStefano & Motl, 2006; Marsh, Scalas, & Nagengast, 2010) in order to "purge" the
estimated latent factors from "contaminating method effects" (Marsh, Scalas et al., 2010, p. 369) due
to construct-irrelevant variability (see also Morin, Arens, & Marsh, 2014). This model (Model 1-2)
used an additional orthogonal latent method factor (LMF) associated with the negatively worded
items. It should be noted, however, that over-and-above the impact of this LMF on model fit,
parameter estimates remained relatively unchanged by this inclusion in the present study.

The measurement invariance and latent mean differences of the PSDQ-S was then examined

across age-groups (11-14 years vs. 15-18 years; see models 2-1 to 2-6), sex-groups (girls vs. boys; see models 3-1 to 3-7), and sport practice involvement (involved vs. not involved; see models 4-1 to 4-6). These invariance tests were performed in the following sequence (Meredith, 1993; Millsap, 2011): (a) configural invariance, (b) weak (loadings) invariance, (c) strong (loadings, intercepts) invariance, (d) strict (loading, intercepts, uniquenesses) invariance, (v) variance/covariance (loading, intercepts, uniquenesses, variances, covariances) invariance, and (vi) latent means (loading, intercepts, uniquenesses, variances, covariances, latent means) invariance.

The sample sizes within each of the BMI-derived categories (e.g., underweight, overweight, obese) were too small to conduct complete tests of measurement invariances across BMI subgroups. Thus, we relied on multiple indicators multiple causes models (MIMIC; e.g., Marsh, Ellis, Parada, Richards, & Heubeck, 2005; Marsh, Tracey, & Craven, 2006; Morin, Marsh, & Nagengast, 2013; Marsh, Nagengast, & Morin, 2013) to investigate the invariance of the intercepts and latent means of the PSDQ-S measurement model as a function of BMI levels treated as a continuous predictor (see models 5-1 to 5-3). These MIMIC invariance tests as a function of BMI levels were performed in the following sequence (Morin et al., 2013; Marsh et al., 2013): (a) MIMIC null effect model (i.e., the paths from BMI to the latent factors and item intercepts are constrained to be zero); (b) MIMIC saturated model (i.e., the paths from BMI to the item intercepts are freed); and (c) MIMIC invariant intercept model (i.e., the paths from BMI to the latent factors are freed, but the paths from BMI to the item intercepts are constrained to be zero). The goodness-of-fit indices of the MIMIC null effect model can be compared with those of the MIMIC saturated and invariant intercept models to examine whether the BMI has a significant effect on the items intercept or latent mean of the PSDQ-S, respectively (Morin et al., 2013; Marsh et al., 2013). Additionally, the comparison of the MIMIC saturated and invariant intercept models indicates whether the item intercepts of the PSDQ-S are invariant or not according to the BMI of the adolescents (Morin et al., 2013; Marsh et al., 2013).

Given the known oversensitivity of the chi-square test of exact fit to sample size and minor model misspecifications (e.g., Marsh, Hau, & Grayson, 2005), we relied on common fit indices to describe the fit of the alternative models (e.g., Hu & Bentler, 1999): the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA), and the 90%

confidence interval of the RMSEA. Values greater than .90 and .95 for the CFI and TLI indicate, respectively, acceptable and good model fit, while values smaller than .06 and .08 for the RMSEA, respectively, support good and acceptable model fit. Because chi-square difference tests present problems similar to those of the chi-square itself, robust χ^2 difference tests ($\Delta R\chi^2$; Satorra, 2000; Satorra & Bentler, 2001) were also complemented by the examination of changes in fit indices. According to Cheung and Rensvold (2002) and Chen (2007) a ΔCFI of .01 or less and a ΔRMSEA of .015 or less between a more restricted model and the preceding one in the invariance sequence suggest that the invariance hypothesis should not be rejected. Since indices incorporating a penalty for parsimony (TLI and RMSEA) can also improve in more restricted models, ΔTLIs are also inspected (Marsh et al., 2005). Furthermore, although the efficacy of the proposed guidelines for the comparison of nested invariance models has been validated in CFA for tests of weak, strong, and strict measurement invariance (Chen, 2007; Cheung & Rensvold, 2002), they appear to be of questionable efficacy for tests of latent mean invariance (Fan & Sivo, 2009). In addition, these indices still appear to show sensitivity to design conditions and model complexity (e.g., Fan & Sivo, 2005, 2007). Overall, these guidelines should not be treated as golden rules, but only as rough guidelines for descriptive model evaluation and comparison that should also take into account parameters estimates, statistical conformity, and theoretical adequacy (Fan & Sivo, 2009; Marsh et al., 2005).

Scale score reliability was computed from the CFA standardized parameter estimates, using McDonald's (1970) $\omega = (\Sigma|\lambda i|)^2 / ([\Sigma|\lambda i|]^2 + \Sigma\delta ii)$ where λi are the factor loadings and δii, the error variances, which has the advantage of taking into account the strength of association between each items, and the latent construct it is purported to assess, as well as item-specific measurement error (Sijtsma, 2009). Finally, the convergent validity of the PSDQ-S subscales with the PSI-S subscales was also examined among the overall sample using latent variable correlations between all subscales from both instruments. In addition, the invariance of these latent variable correlations was also examined across age- and sex-groups, as well as sport practice involvement. These tests all started from a model of configural invariance of the CFA measurement models of both instruments (i.e., loading, intercepts, uniquenesses, and variance-covariances of each instrument were invariant, while the covariances between the PSDQ-S and PSI-S were freely estimated). The goodness-of-fit of this

model was then compared to the fit of a model where the covariances between the PSDQ-S and PSI-S were constrained to invariance.

Results

Factor Validity and Reliability

The goodness-of-fit statistics of the CFA models are presented in Table 1. The results showed a satisfactory degree of fit to the data (i.e., CFI and/or TLI > .90; RMSEA < .05) for the two CFA models (11-factor model without and with a LMF). However, changes in goodness-of-fit indices revealed that the 11-factor model with LMF (Model 1-2) provided a significantly better level of fit to the data than the traditional 11-factor model (Model 1-1). The standardized parameter estimates of this model are presented in Table 2, while the latent factors correlations and scale score reliability estimates are reported in Table 3. These results reveal that all factor loadings were substantial and significant ($M_\lambda = .75$), except for one negatively worded item from the global self-esteem subscale (item GSE 5 "Nothing I ever do seems to turn out right" – "*Dans la vie, rien de ce que je fais ne semble aboutir*") that was associated with a suboptimal loading ($\lambda = .153$). The results further showed that most of the latent factor correlations were significant ($r_{range} = -.15$ to $.80$; $M_{r =} .41$), except for those involving the health subscale, which proved not to be significantly correlated with the other PSDQ-S subscales. In addition, the results also showed strong correlations ($r < .70$) between (a) the coordination and flexibility subscales, (b) the sport competence and four other subscales (i.e., activity, endurance, global physical self-concept, and strength subscales), (c) the global physical self-concept and two other subscales (i.e., global self-esteem and appearance), and (d) the endurance and two other subscales (i.e, activity and strength). Finally, scale score reliability coefficients were acceptable for all subscales ($\omega_{range} = .74$ to $.91$; $M_\omega = .86$).

Multiple Group Measurement Invariance Tests

The measurement invariance of this final, retained model was then verified across age groups (Models 2-1 to 2-6), sex (Models 3-1 to 3-7), and sport practice involvement (Models 4-1 to 4-6). The results from the measurement invariance tests conducted across age groups and sport practice involvement were very similar and showed that (a) all of the χ^2 and most of the $\Delta R\chi^2$ were significant; (b) the CFI, TLI and RMSEA indicated adequate model fit at all steps; (c) the $\Delta RMSEA$ never

showed an increase greater than .015; and (d) the ΔCFI and ΔTLI never showed a decrease greater

than .01, except for the latent mean invariance across subgroups formed based on sport practice

involvement (Model 4-6). These results confirmed the full measurement invariance of the PSDQ-S

across age groups and sport practice involvement, but also highlighted significant latent mean

differences between adolescents involved or not in sport practice. Results showed that, when the

latent means of adolescents not involved in sport practice were fixed to zero (to serve as the first

comparison group) for identification purposes, the latent means (expressed in standard deviation

units) of those involved in sport practice were significantly ($p \leq .05$) higher on the coordination

(0.60), strength (0.60), flexibility (0.30), endurance (0.79), activity (1.57), sport competence (0.69),

appearance (0.21), global physical self-concept (0.31), and global self-esteem (0.34) subscales. No

significant differences were found for the health (0.13) and body fat (0.02) subscales.

Additionally, the measurement invariance tests (Models 3-1 to 3-7) conducted across sex

groups showed that (a) all of the χ^2 and most of the $\Delta R\chi^2$ were significant; (b) the CFI, TLI and

RMSEA indicated adequate model fit at all steps (except for steps 3-5 and 3-7); (c) the ΔRMSEA

never showed an increase greater than .015; and (d) the ΔCFI and ΔTLI remained below .01 for the

first three steps, confirming the strict (i.e., configural, loadings, intercepts, and uniquenesses)

measurement invariance of the model. However, the ΔCFI was greater than .01 at the fifth step

(Model 3-5), suggesting a lack of invariance for the latent variance-covariance. This lack of

invariance was due to a higher level of variability in the health and activity subscales among girls

relative to boys. When the invariance constraints were relaxed for the variances of these factors

(Model 3-6), the results supported the partial invariance of the variance-covariance matrix. Finally,

the ΔCFI and ΔTLI were also greater than .01 for the last step (Model 3-7), providing evidence of

latent means differences across sex. Probing these differences revealed that, when girls' latent means

were fixed to zero (to serve as the first comparison group) for identification purposes, boys' latent

means (expressed in standard deviation units) were significantly ($p \leq .05$) higher on the coordination

(0.58), strength (1.13), endurance (1.06), health (0.92), activity (0.68), body fat (0.62), sport

competence (0.87), appearance (0.45), global physical self-concept (0.68), and global self-esteem

(0.64) subscales. No significant differences were found on the flexibility (0.07) subscale.

MIMIC Models

The results from the MIMIC null effect model (see Model 5-1 in Table 1) provided a satisfactory fit to the data (CFI and TLI > .90; RMSEA < .05). Additionally, the MIMIC saturated model (see Model 5-2 in Table 1) provided a satisfactory fit to the data (CFI and TLI > .90; RMSEA < .05). Results also showed that the MIMIC saturated model provided a substantially better fit to the data than the MIMIC null effect model (i.e., ΔCFI and ΔTLI increase of .014 and .011, respectively). The third model (i.e., the MIMIC invariant intercept model, see Model 5-3 in Table 1) also provided a satisfactory fit to the data (CFI and TLI > .90; RMSEA < .05) and also provided a better fit to the data than MIMIC null effect model. Finally the results showed that the fit of the MIMIC saturated model was not substantially better (i.e., ΔCFI, ΔTLI, and ΔRMSEA are all under .01) than the fit of the MIMIC invariant intercept model. This result shows that the effects of BMI levels on the items are fully explained by the effects of BMI levels on the latent means, thus supporting the invariance of the items intercepts as a function of BMI levels. Detailed results from the MIMIC invariant intercept model revealed that adolescents with higher levels of BMI tended to present higher latent means on the strength subscale ($\hat{\beta} = .162$) and lower latent mean on the body fat ($\hat{\beta} = -.456$; also see endnote 2) subscale. However, no significant differences were found for the other PSDQ-S subscales.

Convergent Validity

The convergent validity of the PSDQ-S was assessed in relation to another validated measure of physical self-concept (PSI-S) using latent correlations. The goodness-of-fit statistics of the CFA model including both instruments are fully satisfactory (see Model 1-3 in Table 1), and the scale score reliability coefficients of the PSI-S (see Table 3) were satisfactory for most subscales ($\omega_{range} = .70$ to .91, $M_\omega = .84$), with the exception of the physical attractiveness subscale for which the scale score reliability proved unsatisfactory ($\omega = .47$; for discussion of problems associated with this scale, see Morin & Maïano, 2011). Thus, this subscale was not analyzed further.

When the latent covariances between the subscales from both instruments were constrained to be invariant across subgroups formed on the basis of age ($\Delta\chi^2 = 108.11$; $\Delta df = 66$; $p < .001$; $\Delta CFI = -0.002$; $\Delta TLI < 0.000$; $\Delta RMSEA < 0.000$), sex ($\Delta\chi^2 = 103.12$; $\Delta df = 66$; $p < .001$; $\Delta CFI = -0.002$; ΔTLI

< 0.000; ΔRMSEA < 0.000), and sport practice involvement (Δχ² = 79.99; Δdf = 66; p > .05; ΔCFI = -0.001; ΔTLI = +0.002 ΔRMSEA < 0.000), the results fully supported the invariance of these latent covariances across subgroups. These results thus support the equivalence of the convergent validity correlations across subgroups. For this reason, we only focus on the latent correlations calculated on the basis of the complete sample.

The latent variables correlations between both instruments are reported in Table 3. Results showed that the global self-esteem (r = .75), global physical self-concept (r = .85), endurance (r = .92), strength (r = .94), and sport competence (r = .88) subscales of the PSDQ-S were strongly correlated to their respective PSI-S subscales. Furthermore, the global physical self-concept subscale of the PSDQ-S was also strongly significantly correlated to the global self-worth (r = .87) subscale of the PSI-S.

However, the results also revealed some high correlations (r > .70) between non-corresponding subscales from both instruments. For instance: (a) the strength subscale of the PSDQ-S was highly correlated with the physical self-worth (r = .72) and sport competence (r = .73) subscales of the PSI-S, (b) the sport competence subscale of the PSDQ-S was highly correlated with the physical self-worth (r = .77) subscale of the PSI-S, (d) the global self-esteem subscale of the PSDQ-S was highly correlated with the physical self-worth (r = .71) subscale of the PSI-S, (e) the global physical self-concept subscale of the PSDQ-S was highly correlated with the global self-worth (r = .87) subscale of the PSI-S. Supporting the convergent validity of the PSDQ-S subscales, it is interesting to note that these high correlations are in line with those observed between matching PSDQ-S subscales. For instance, the correlation between the strength and global physical self-concept subscales of the PSDQ-S (r = .62) was similar to the correlation between the strength subscale of the PSDQ-S and physical self-worth scale of the PSI-S (r = .72). These correlations are also in line with those previously reported among matching PSI-S subscales (e.g., Morin & Maïano, 2011a). Finally, it must be noted that the correlations involving the global self-esteem and the global physical self-concept subscales were in line with the proposed hierarchical nature of the physical self-concept (e.g., Fox & Corbin, 1989), suggesting strong levels of associations between hierarchically order self-components.

Discussion

The first objective of this study was to evaluate the factor validity and reliability of the PSDQ-S among a sample of French adolescents. The results showed that the *a priori* 11-factor model provided a satisfactory degree of fit to the data and that the incorporation of a LMF related to the negatively worded items further increased the fit of the a priori PSDQ-S model. The superiority of a model including a LMF related to the negatively worded items confirms numerous previous recommendations regarding the need to control for item wording effects in measurement models (e.g., DiStefano & Motl, 2006; Dunbar et al., 2000; Marsh, Scalas et al., 2010; Tomás & Oliver, 1999).

The results showed that this final measurement model presented good scale score reliability coefficients ($M_\omega = .86$) and that most of the PSDQ-S subscales were moderately correlated ($M_{r =} .41$), with the exception of the Health subscale, which presented non-significant correlations with all other PSDQ-S subscales. These results are mostly consistent with previous results showing either non-significant, or at least lower, correlations involving the Health subscale (e.g., Agarwal et al., 2013; Marsh, Martin et al., 2010), which is also consistent with the fact that this specific subscale taps into a somewhat different facet of physical conceptions than the rest of the PSDQ-S subscales. Overall, the results from the factorial validity and reliability of the PSDQ-S are consistent with those recently reported by Marsh, Martin et al. (2010).

The results also show that one negatively worded item (out of a total of 40 items included in the PSDQ-S) from the global self-esteem subscale was suboptimal. This result is consistent with recent findings reported by Martin and Whalen (2013) among American athletes with physical disabilities, which also demonstrated that the two negatively worded items of the global self-esteem scale of the PSDQ-S were suboptimal. Although these authors used this observation to support the elimination of these items, we recommend that future studies using both the French and original English version of the PSDQ-S should devote special attention to the performance of this item. Furthermore, it may also be interesting for future research using both the English and French version of the PSDQ-S to investigate the addition of a positively-worded version of this item (i.e., "Everything I do seems to turn out right" – "*Dans la vie, tout ce que je fais semble bien tourner*"). This will help to determine whether the psychometric properties of the current French version of the PSDQ can be preserved or improved with the proposed reformulation of this item, while at the same time ensuring that the same

set of items is retained across linguistic versions.

The second objective was to examine the measurement and latent mean invariance of the factor structure of the PSDQ-S across various subgroups (i.e., sex, age, BMI, and sport practice involvement) of adolescents. Analyses provided strong support for the strict measurement invariance (i.e., invariance of the factor loadings, items' intercepts, and items' uniquenesses) of the PSDQ-S across all of these subgroups. These results are not only consistent with those from Marsh, Martin et al. (2010) for the sex and age subgroups, but also demonstrate for the first time the invariance of the PSDQ-S across subgroups formed based on their involvement, or lack thereof, in normative levels of sport practice.

Additionally, results revealed significant latent mean differences across sex and sport practice subgroups. These results are consistent with those reported from previous studies that have investigated mean-level differences in physical self-conceptions, thus supporting the validity of the French version of the PSDQ-S (e.g., Bowker, 2006; Findlay & Bowker, 2007; Marsh, 1998; Marsh et al., 1997; Marsh, Martin et al., 2010; Schmalz & Davison, 2006). More precisely, these results showed that boys and adolescents involved in sport practice tend to present higher levels on most (except for flexibility across sex, or health and body fat across sport practice involvement) of the PSDQ-S subscales. In addition, results also showed that adolescents with higher BMI levels tended to present higher latent means on the strength subscale and lower latent means on the body fat subscale. These results are consistent with Marsh et al. (2007). However, it must be kept in mind that the current study revealed far less significant differences than Marsh et al. (2007). One potential explanation for this difference in results may be attributed to the fact that in the present study the weight and heights measures were self-reported, whereas in the Marsh et al.'s (2007) study they were objectively measured, which may have induced some biases in the estimation of BMI levels. Finally, no significant latent mean differences were observed across age-groups. These findings contrast with those from previous studies of the PSDQ (e.g., Marsh, 1998; Marsh et al., 2007), which showed lower physical self-conceptions for older adolescents (compared with younger). This lack of latent mean differences across age-categories and as a function of BMI for most subscales may also be related to a loss of precision of the PSDQ-S to detect mean differences when compared to the longer original

instrument. Finally, this lack of latent mean differences across age-categories may also potentially be explained by cultural differences between French and Australian adolescents. However, this latter possibility should be taken with caution, as this is the first study to examine age- and BMI-related differences using the PSDQ-S. Clearly, age- and BMI-related differences (or lack thereof) on the PDSQ-S subscales should be more thoroughly examined in future cross-cultural studies.

The third objective of this study was to test the convergent validity of the PSDQ-S in relation to the PSI-S). Findings highlight that the PSDQ-S subscales are strongly related to their respective scales measured with the PSI-S and present a pattern of associations with PSI-S subscales that is consistent with those observed among PDSQ-S in this study of among the PSI-S subscales in the context of previous studies (e.g., Morin & Maïano, 2011a). In addition, results also revealed that these patterns of association are also invariant across age groups, sex groups, and adolescents involved or not in sport practice. This means that responses provided by the adolescents with both of these instruments were very similar in magnitude. Consequently, these results support the convergent validity of the French version of the PSDQ-S.

In addition to the previously noted limitations, an important limitation of the present investigation that needs to be kept in mind when interpreting these results is the need for cross-validation of the present results to additional and more diversified samples of French adolescents, as well as to samples from other cultural and linguistic groups in order to fully ascertain the possibility to rely on the PSDQ-S in the context of comprehensive cross-cultural studies. Another limitation of the present study is that we did not assess the test-retest reliability of the French version of the PSDQ-S, which should be examined in future research in order to obtain a fuller picture of the psychometric properties of this instrument. In a related way, it would be interesting to extend these examination of test-retest reliability to more extensive developmental research in order to systematically assess the degree of sensitivity of the PSDQ-S to meaningful developmental differences in physical self-conceptions – for instance to change occurring across the pubertal transition. In addition, a complete test of the psychometric properties of the French version of the PSDQ-S would require additional verifications of the convergent validity of this instrument in relation to more objective physical fitness tests. These issues should be addressed in future studies. Finally, another important limitation of this

study is that the sample size available here precluded more refined tests of measurement invariance and latent means differences across combinations of grouping variables. This is one area that should be more precisely investigated in future studies. An interesting extension of these tests would be to apply mixture distribution item response modeling (MIRT) in order to extract subgroups of participants for whom the instrument shows differential item functioning (for application of this method to the PSDQ, see Freund, Tietjens, & Strauss, 2013; Tietjens, Freund, Büsch, & Strauss, 2012).

In conclusion, our findings suggest that the French PSDQ-S presents acceptable psychometric properties and may be confidently used in research or practice to assess the physical self-conceptions of French adolescents with similar characteristics and in the context of group-based comparisons based on sex, age, BMI, and sport practice involvement. Nevertheless, the present study showed that one global self-esteem item included in this French version may be suboptimal and thus we encourage future research to systematically investigated alternative operationalization of this subscale.

References

Agarwal, S., Bhalla, P., Kaur, S., & Babbar, R. (2013). Effect of body mass index on physical self

concept, cognition & academic performance of first year medical students. *The Indian Journal of

Medical Research*, *138*, 515-522.

Bowker, A. (2006). The relationship between sports participation and self-esteem during early

adolescence. *Canadian Journal of Behavioural Science*, 38, 214-229. doi: 10.1037/cjbs2006009

Castonguay, A. L., Sabiston, C. M., Crocker, P. R., & Mack, D. E. (2014). Development and

validation of the Body and Appearance Self-Conscious Emotions Scale (BASES). *Body Image*, *11*,

126-136. doi: 10.1016/j.bodyim.2013.12.006

Chen, F .F. (2007). Sensitivity of goodness of fit indexes to lack of measurement. *Structural Equation

Modeling*, *14*, 464-504. doi: 10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of fit indexes for testing

measurement invariance. *Structural Equation Modeling*, *9*, 233-255. doi:

10.1207/S15328007SEM0902_5

DiStefano, C., & Motl, R. W. (2006). Further investigating method effects with negatively worded

items on self-report surveys. *Structural Equation Modeling*, *13*, 440-464. doi:

10.1207/s15328007sem1303_6

Duda, J.L., Quested, E., Haug, E., Samdal, O., Wold, B., Balaguer, I., Castillo, I., Sarrazin, P.,

Papaioannou, A., Ronglan, L.T., Hall, H., & Cruz, J. (2013). Promoting Adolescent health through

an intervention aimed at improving the quality of their participation in Physical Activity (PAPA):

Background to the project and main trial protocol. *International Journal of Sport and Exercise

Psychology*, *11*, 319-327. doi: 10.1080/1612197X.2013.839413

Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford.

Fan, X., & Sivo, S.A. (2005). Sensitivity of fit indexes to misspecified model components: Rationale

of two-index strategy revisited. *Structural Equation Modeling, 12*, 343-367. doi:

10.1207/s15328007sem1203_1

Fan, X., & Sivo, S.A. (2007). Sensitivity of fit indices to model misspecification and model types.

Multivariate Behavioral Research, *42*, 509-529. doi: 10.1080/00273170701382864

Fan, X., & Sivo, S.A. (2009). Using goodness-of-fit indexes in assessing mean structure invariance. *Structural Equation Modeling, 16*, 54-69. doi: 10.1080/10705510802561311

Findlay, L. C., & Bowker, A. (2007). The link between competitive sport participation and self-concept in early adolescence: A consideration of gender and sport orientation. *Journal of Youth and Adolescence*, *38*, 29-40. doi: 10.1007/s10964-007-9244-9

Fox, K. R., & Corbin, C. B. (1989). The Physical Self-Perception Profile: Development and preliminary validation. *Journal of Sport and Exercise Psychology*, *11*, 408-430.

Freund, P. A., Tietjens, M., & Strauss, B. (2013). Using rating scales for the assessment

of physical self-concept: why the number of response categories matters. *Measurement in Physical Education and Exercise Science*, *17*, 249–263. doi: 10.1080/1091367X.2013.807265

Guérin, F., Marsh, H.W., & Famose, J.-P. (2004). Generalizability of the PSDQ and relationship to physical fitness: European French connection. *Journal of Sport and Exercise Psychology*, *26*, 19-38.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria vs new alternatives. *Structural Equation Modeling, 6,* 1-55. doi 10.1080/10705519909540118

Jones, S. M., & Stumbrys, T. (2014). Mental health, physical self and lucid dreaming: A correlational study in sport students. *International Journal of Dream Research*, *7*, 54-60. doi: 10.11588/ijodr.2014.1.12330

Maïano, C., Morin, A. J. S., Ninot, G., Monthuy-Blanc, J., Stephan, Y., Florent, J.-.F., & Vallée, P. (2008). A short and very short form of the physical self-inventory: Development and factor validity. *Psychology of Sport and Exercise*, *9*, 830-847. doi: 10.1016/j.psychsport.2007.10.003

Marsh, H. W. (1996a). Physical self-description questionnaire: Stability and discriminant validity. *Research Quarterly for Exercise and Sport*, *67*, 249-264. doi: 10.1080/02701367.1996.10607952

Marsh, H. W. (1996b). Construct validity of physical self-description questionnaire responses: Relations to external criteria. *Journal of Sport and Exercise Psychology*, *18*, 111-131.

Marsh, H. W. (1998). Age and gender effects in physical self-concepts for adolescent elite athletes and nonathletes: A multicohort-multioccasion design. *Journal of Sport and Exercise Psychology*,

20, 237-259.

Marsh, H. W., & Cheng, J. H. S. (2012). Physical self-concept. In G. Tenenbaum, R. Eklund, & A. Kamata (Eds). *Handbook of measurement in sport and exercise psychology* (pp. 215-226). Champaign, IL: Human Kinetics.

Marsh, H. W., Ellis, L., Parada, L., Richards, G., & Heubeck, B. G. (2005). A short version of the Self Description Questionnaire II: Operationalizing criteria for short-form evaluation with new applications of confirmatory factor analyses. *Psychological Assessment*, *17*, 81-102. doi: 10.1037/1040-3590.17.1.81

Marsh, H. W., Hau, K-T., & Grayson, D. (2005). Goodness of fit evaluation in structural equation modeling. In A. Maydeu-Olivares, & J. McArdle (Eds.), *Contemporary psychometrics* (pp. 275-340). Hillsdale, NJ: Erlbaum.

Marsh, H. W., Hau, K. T., Sung, R. Y. T, & Yu, C. W. (2007). Childhood obesity, gender, actual-ideal body image discrepancies, and physical self-concept in Hong Kong children: Cultural differences in the value of moderation. *Developmental Psychology*, *43*, 647-662. doi: 10.1037/0012-1649.43.3.647

Marsh, H. W., Hey, J., Roche, L. A., & Perry, C. (1997). The structure of physical self-concept: Elite athletes and physical education students. *Journal of Educational Psychology*, *89*, 369-380. doi: 10.1037/0022-0663.89.2.369

Marsh, H. W., Martin, A. J., & Jackson, S. (2010). Introducing a short version of the Physical Self-Description Questionnaire: New strategies, short-form evaluative criteria, and applications of factor analyses. *Journal of Sport and Exercise Psychology*, *32*, 438-482.

Marsh, H. W., Nagengast, B., & Morin, A. J. (2013). Measurement invariance of big-five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and la dolce vita effects. *Developmental Psychology*, *49*, 1194-1218. doi: 10.1037/a0026913

Marsh, H. W., & Redmayne, R. S. (1994). Multidimensional physical self-concept and its relations to multiple components of physical fitness. *Journal of Sport and Exercise Psychology*, *16*, 43-55.

Marsh, H. W., Richards, G. E., Johnson, S., Roche, L., & Tremayne, P. (1994). Physical self-description questionnaire: Psychometric properties and a multitrait-multimethod analysis of

relations to existing instruments. *Journal of Sport and Exercise Psychology*, *16*, 270-305.

Marsh, H. W., Scalas, L .F., & Nagengast, B. (2010). Longitudinal tests of competing factor structures for the Rosenberg Self-Esteem Scale: Traits, ephemeral artifacts, and stable response styles. *Psychological Assessment*, *22*, 366-381. doi: 10.1037/a0019225

Marsh, H. W., Tracey, D. K., & Craven, R. G. (2006). Multidimensional self-concept structure for preadolescents with mild intellectual disabilities: A hybrid multigroup mimic approach to factorial invariance and latent mean differences. *Educational and Psychological Measurement*, *66*, 795-818. doi: 10.1177/0013164405285910

Martin, J. J., & Whalen, L. (2012). Self-concept and physical activity in athletes with physical disabilities. *Disability and Health Journal*, *5*, 197-200. doi: 10.1016/j.dhjo.2012.03.006

McDonald, R. P. (1970). Theoretical foundations of principal factor analysis, canonical factor analysis, and alpha factor analysis. *British Journal of Mathematical and Statistical Psychology*, *23*, 1-21. doi: 10.1111/j.2044-8317.1970.tb00432.x

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58,* 525-543.

Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York, NY: Routledge.

Morin, A. J. S., Arens, A. K., & Marsh, H. W. (in press). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling*.

Morin, A. J. S., Marsh, H. W., & Nagengast, B. (2013). Exploratory structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed., pp. 395-438). Charlotte, NC: Information Age Publishing, Inc.

Morin, A. J. S., & Maïano, C. (2011a). Cross-validation of the short form of the physical self-inventory (PSI-S) using exploratory structural equation modeling. *Psychology of Sport and Exercise, 12,* 540-554. doi: 10.1016/j.psychsport.2011.04.003

Morin, A. J. S., & Maïano, C. (2011b). Cross-validation of the very short form of the Physical Self-Inventory (PSI-VS): Invariance across genders, age groups, ethnicities and weight statuses. *Body Image*, *8*, 404-410. doi : 10.1016/j.bodyim.2011.06.005

Muthén, L. K., & Muthén, B. O (2013). *Mplus user's guide.* Los Angeles, CA: Muthén & Muthén.

Papaioannou, A. G., Appleton, P. R., Torregrosa, M., Jowett, G. E., Bosselut, G., Gonzalez, L., ... & Zourbanos, N. (2013). Moderate-to-vigorous physical activity and personal well-being in European youth soccer players: Invariance of physical activity, global self-esteem and vitality across five countries. *International Journal of Sport and Exercise Psychology*, *11*, 351-364. doi: 10.1080/1612197X.2013.830429

Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In R. D. Heijmans, D. S. Pollock, & A. Satorra (Eds.), *Innovations in multivariate statistics. A Festschrift for Heinz Neudecker* (pp. 233-247). London: Kluwer Academic Publishers.

Satorra, A., & Bentler, P. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507-514. doi: 10.1007/BF02296192

Schipke, D., & Freund, P. A. (2012). A meta-analytic reliability generalization of the Physical Self-Description Questionnaire (PSDQ). *Psychology of Sport and Exercise*, *13*, 789-797. doi: 10.1016/j.psychsport.2012.04.012

Schmalz, D. L., & Davison, K. K. (2006). Differences in physical self-concept among pre-adolescents participating in gender-typed and cross-gendered sports. *Journal of Sport Behavior*, *29*, 335-352.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, *46*, 407-411. doi: 10.2307/1170010

Sijtsma, K. (2009). On the use, misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*, 107-120. doi: 10.1007/s11336-008-9101-0

Stiller, J., & Alfermann, D. (2007). Die deutsche übersetzung des Physical Self-Description Questionnaire (PSDQ) [German translation of the Physical Self-Description Questionnaire (PSDQ)]. *Zeitschrift für Sportpsychologie*, *14*, 149-161. doi: 10.1026/1612-5010.14.4.149

Sypsa, C., & Simons, J. (2008). Questionnaires measuring the physical self children: A review. *European Psychomotricity Journal*, *1*, 61-72.

Tietjens, M., Freund, P. A., Büsch, D., & Strauss, B. (2012). Using mixture distribution models to test the construct validity of the Physical Self-Description Questionnaire. *Psychology of Sport and Exercise*, *13*, 598-605. doi:10.1016/j.psychsport.2012.02.009

Tomás, I., Marsh, H. W., González-Romá, V., Valls, V., & Nagengast, B. (2014). Testing

measurement invariance across Spanish and English versions of the Physical Self-Description

Questionnaire: An application of exploratory structural equation modeling. *Journal of Sport and

Exercise Psychology*, *36*, 179-188. doi: 10.1123/jsep.2013-0070

Endnotes

[1] We use "sport practice" to refer to practicing sports and fitness activities/exercises (fitness,

aerobics, steps, weight training, running, etc.) for recreational or competitive purposes.

[2] It should be noted here that higher scores on the body fat subscale reflect perceptions that one's

body is thinner, whereas lower scores reflect perceptions that one's body is fatter.

Table 1

Goodness-of-Fit Statistics of Confirmatory Factor Analytic (CFA) Models

Model	Description	$\chi^2(df)$	CFI	TLI	RMSEA	90% CI	CM	$\Delta R\chi^2(df)$	ΔCFI	ΔTLI	ΔRMSEA
CFA	1-1. CFA 11-factor.	1614.43(685)**	.920	.908	.048	.045-.051	-	-	-	-	-
	1-2. CFA-LMF 11-factor.	1411.96(675)**	.936	.926	.043	.040-.046	1-1	131.57(10)**	+.016	+.018	-.005
	1-3. CFA-PSDQ-S & PSI-S	2909.64(1449)**	.924	.913	.041	.039-.044	-				
CFA: Age	2-1. Configural invariance	2330.95(1350)**	.919	.907	.050	.046-.053	-	-	-	-	-
	2-2. λs invariant	2340.92(1388)**	.922	.912	.048	.045-.052	2-1.	27.18(38)	+.003	+.005	-.002
	2-3. λs, τs invariant	2357.87(1416)**	.923	.915	.048	.044-.051	2-2.	13.05(28)	+.001	+.003	.000
	2-4. λs, τs, δs invariant	2379.54(1456)**	.924	.919	.046	.043-.050	2-3.	38.46(40)	+.001	+.004	-.002
	2-5. λs, τs, δs, ξs/φs invariant	2455.98(1523)**	.923	.921	.046	.042-.049	2-4.	77.53(67)	-.001	+.002	.000
	2-6. λs, τs, δs, ξs/φs, ηs invariant	2480.78(1535)**	.922	.921	.046	.042-.049	2-5.	26.01(12)*	-.001	.000	.000
CFA: Sex	3-1. Configural invariance	2229.79(1350)**	.922	.910	.047	.044-.051	-	-	-	-	-
	3-2. λs invariant	2316.66(1388)**	.917	.907	.048	.044-.051	3-1.	86.37(38)**	-.005	-.003	+.001
	3-3. λs, τs invariant	2402.56(1416)**	.912	.903	.049	.045-.052	3-2.	85.34(28)**	-.005	-.004	+.001
	3-4. λs, τs, δs invariant	2458.93(1456)**	.911	.905	.048	.045-.052	3-3.	60.78(40)*	-.001	+.002	-.001
	3-5. λs, τs, δs, ξs/φs invariant	2662.27(1523)**	.899	.896	.050	.047-.054	3-4.	195.70(67)**	-.012	-.009	+.002
	3-6. λs, τs, δs, ξs/φs (ξ^{HE}, ξ^{AC}) invariant	2626.72(1521)**	.902	.899	.050	.047-.053	3-4.	161.70(65)**	-.009	-.006	+.002
	3-7. λs, τs, δs, ξs/φs (ξ^{HE}, ξ^{AC}), ηs invariant	2825.28(1533)**	.885	.883	.054	.050-.057	3-6.	283.68(12)**	-.017	-.016	+.004

Note. AC = activity; BMI = body-mass index; CFA = confirmatory factor analytic model; CFI = comparative fit index; CM = comparison model; *df* = degrees of freedom; HE = health; LMF = latent method factor; RMSEA = root mean square error of approximation; PSDQ-S = Physical Self-Description Questionnaire – short form; PSI-S = Physical Self-Inventory – short version; TLI = Tucker-Lewis index; 90% CI = 90% confidence interval of the RMSEA; λ = loading; τ = intercept; δ = uniquenesses; ξ = variance; φ = covariance; η = factor means; χ^2 = chi-square; $\Delta R\chi^2$ = robust chi-square difference tests (calculated from loglikelihoods for greater precision); Δ = change from previous model. * *p* < .05, ** *p* <.01.

Table 1 (continued)

Model	Description	$\chi^2(df)$	CFI	TLI	RMSEA	90% CI	CM	$\Delta R\chi^2(df)$	ΔCFI	ΔTLI	ΔRMSEA
CFA: Sport	4-1. Configural invariance	2276.09(1350)**	.919	.907	.048	.045-.052	-	-	-	-	-
	4-2. λs invariant	2327.70(1388)**	.918	.908	.048	.045-.051	4-1.	52.46(38)	-.001	+.001	.000
	4-3. λs, τs invariant	2369.78(1416)**	.917	.909	.048	.045-.051	4-2.	41.62(28)	-.001	+.001	.000
	4-4. λs, τs, δs invariant	2395.60(1456)**	.918	.912	.047	.044-.050	4-3.	40.88(40)	+.001	+.001	-.001
	4-5. λs, τs, δs, ξs/φs invariant	2552.44(1523)**	.910	.908	.048	.045-.051	4-4.	157.49(67)**	-.008	-.004	+.001
	4-6. λs, τs, δs, ξs/φs, ηs invariant	2763.97(1535)**	.893	.891	.052	.049-.055	4-5.	226.46(12)**	-.017	-.017	+.004
CFA including BMI	5-1 MIMIC null effect	1606.15(715)**	.924	.913	.046	.043-.049	-	-	-	-	-
	5-2 MIMIC saturated	1409.73(675)**	.938	.924	.043	.040-.046	5-1	206.55(40)**	+.014	+.011	-.003
	5-3 MIMIC invariant intercept	1454.41(704)**	.936	.926	.043	.040-.046	5-2	43.03(29)*	-.002	+.002	.000

Table 2

Standardized Parameters Estimates from the Confirmatory Factor Analytic Model with Latent Method Factor

Items	CO (λ)	ST (λ)	FL (λ)	EN (λ)	GSE (λ)	HE (λ)	AC (λ)	BF (λ)	SC (λ)	GPSC (λ)	AP (λ)	LMF (λ)	δ
CO1	.682												.535
CO2	.739												.454
CO3	.815												.336
CO4	.827												.316
CO5	.796												.366
ST1		.850											.278
ST2		.850											.278
ST3		.850											.278
FL1			.790										.376
FL2			.884										.219
FL3			.811										.342
EN1				.768									.410
EN2				.880									.225
EN3				.787									.381
GSE1					.727								.472
GSE2*					.327							.505	.638
GSE3					.761								.421
GSE4					.742								.449
GSE5*					.153							.654	.549
HE1*						.617						.371	.481
HE2*						.537						.594	.358
HE3*						.623						.505	.357
HE4*						.389						.556	.539
HE5*						.361						.660	.434
AC1							.734						.461
AC2							.800						.361

Items	CO (λ)	ST (λ)	FL (λ)	EN (λ)	GSE (λ)	HE (λ)	AC (λ)	BF (λ)	SC (λ)	GPSC (λ)	AP (λ)	LMF (λ)	δ
AC3							.845						.287
AC4							.817						.333
BF1*								.831				.312	.213
BF2*								.859				.352	.138
BF3*								.774				.210	.357
SC1									.855				.269
SC2									.873				.238
SC3									.896				.197
GPSC1										.860			.261
GPSC2										.827			.315
GPSC3										.871			.242
AP1											.847		.283
AP2											.743		.448
AP3											.894		.201

Note. * reversed items; λ = loading; δ = uniquenesses; AC = activity; AP = appearance; BF = body fat; CO = coordination; EN = endurance; FL = flexibility; GPSC = global physical self-concept; GSE = global self-esteem; HE = health; LMF = latent method factor; SC = sport competence; ST = strength. All loadings and uniquenesses significant at *p* < .01

Table 3

Latent Factor Correlations (Top) and Convergent Validity (Bottom) of the Confirmatory Factor Analytic (CFA) Model with Latent Method Factor

Factors	CO	ST	FL	EN	GSE	HE	AC	BF	SC	GPSC	AP	ω
					Latent factor correlations from the CFA model							
CO	-											.88
ST	.669**	-										.89
FL	.705**	.382**	-									.87
EN	.643**	.730**	.443**	-								.85
GSE	.626**	.619**	.405**	.629**	-							.74
HE	.033	-.021	-.148	-.069	-.035	-						.75
AC	.592**	.605**	.428**	.755**	.504**	.021	-					.88
BF	.150*	.109	.069	.248**	.255**	.072	.164**	-				.90
SC	.700**	.776**	.364**	.802**	.683**	-.030	.708**	.271**	-			.91
GPSC	.586**	.623**	.353**	.622**	.781**	-.009	.490**	.491**	.710**	-		.89
AP	.437**	.551**	.353**	.461**	.656**	-.086	.396**	.225**	.440**	.702**	-	.87
					Latent factor correlations from the convergent validity analyses							
PSI-S-GSW	.415**	.419**	.256**	.379**	.745**	.018	.272**	.504**	.469**	.874**	.696**	.70
PSI-S-PSW	.627**	.717**	.316**	.686**	.713**	.032	.571**	.384**	.773**	.847**	.581**	.86
PSI-S-PC	.511**	.642**	.301**	.922**	.471**	-.027	.570**	.234**	.680**	.500**	.394**	.88
PSI-S-SC	.594**	.728**	.315**	.688**	.619**	.014	.593**	.286**	.884**	.647**	.473**	.91
PSI-S-PA	.499**	.571**	.369**	.510**	.705**	-.064	.432**	.492**	.556**	.840**	.877**	.47
PSI-S-PS	.484**	.938**	.253**	.640**	.514**	-.038	.514**	.086	.665**	.533**	.454**	.83

Note. ω = McDonald (1970) scale score reliability coefficient; AC = activity; AP = appearance; BF = body fat; CO = coordination; EN= endurance; FL = flexibility; GPSC = global physical self-concept; GSE = global self-esteem; GSW = global self-worth; HE = health; PA = physical attractiveness; PC = physical condition; PS = physical strength; PSI-S = Physical Self-Inventory-short version; PSW = physical self-worth; SC = sport competence; ST = strength.*p < .05, **p < .01.