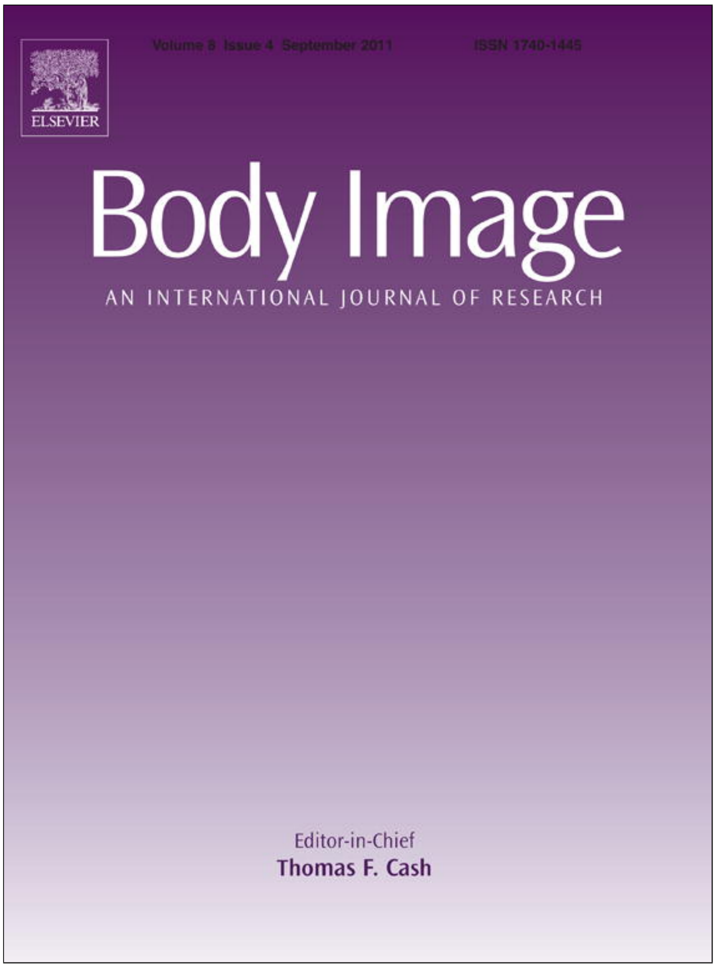
Volume 8 Issue 4 September 2011
ISSN 1740-1445
ELSEVIER
Body Image
AN INTERNATIONAL JOURNAL OF RESEARCH
Editor-in-Chief
Thomas F. Cash

Body Image 8 (2011) 404–410

Contents lists available at ScienceDirect

Body Image

journal homepage: www.elsevier.com/locate/bodyimage

Cross-validation of the very short form of the Physical Self-Inventory (PSI-VS): Invariance across genders, age groups, ethnicities and weight statuses

Alexandre J.S. Morin[a,b,*,1], Christophe Maïano[c,1]

[a] Department of Psychology, University of Sherbrooke, Sherbrooke, Québec, Canada
[b] Educational Excellence and Equity (E3) Research Program, Centre for Educational Research (University of Western Sydney), Sydney, Australia
[c] Institute of Movement Sciences Etienne-Jules Marey (UMR 6233), CNRS-University of Aix-Marseille II, Marseille, France

ARTICLE INFO

Article history:
Received 10 February 2011
Received in revised form 25 May 2011
Accepted 16 June 2011

Keywords:
Physical self-concept
Physical Self-Inventory
Very short form
Confirmatory factor analysis
Convergent validity
Reliability
Measurement invariance

ABSTRACT

In a recent review of various physical self-concept instruments, Marsh and Cheng (in press) noted that the very short 12-item version of the French Physical Self-Inventory (PSI-VS) represents an important contribution to applied research but that further research was needed to investigate the robustness of its psychometric properties in new and diversified samples. The present study was designed to answer these questions based on a sample of 1103 normally achieving French adolescents. The results show that the PSI-VS measurement model is quite robust and fully invariant across subgroups of students formed according to gender, weight, age and ethnicity. The results also confirm the convergent validity and scale score reliability of the PSI-VS subscales.

Introduction

In their classic review, Shavelson, Hubner, and Stanton (1976) represented self-concept as a pyramid, with global self-esteem at the apex and more specific constructs at the next-lower level, such as the academic self, the social self and the physical self. Fox and Corbin (1989) extended this conception to the physical self-concept. In their model, the upper level is occupied by a generic construct representing Global Self-Worth (GSW) and referring to the positive or negative way people feel about themselves as a whole. The next level (the domain level) is occupied by global Physical Self-Worth (PSW; satisfaction and pride in the physical self). Finally, the lower level (the subdomain level) is occupied by four constructs: Sport Competence (SC: ability to learn sports, etc.), Physical Condition (PC: stamina, fitness, etc.), Physical Attractiveness (PA: ability to maintain an attractive body over time, etc.) and Physical Strength (PS: perceived strength, etc.).

From this model, Fox and Corbin (1989) developed the Physical Self-Perception Profile (PSPP). Since then, the original PSPP, cross-cultural adaptations of this instrument, and multiple additional instruments have been validated to assess physical self-concept (for a review, see Marsh & Cheng, in press). From the PSPP and attempting to circumvent problems typically associated with the PSPP, Ninot, Delignières, and Fortes (2000) developed the French Physical Self-Inventory (PSI) for adults. Maïano et al. (2008) recently developed a very short form of the PSI (PSI-VS; 12 items, with two items per dimension) for use with adolescents in the context of in-depth longitudinal or idiographic studies (e.g., Stone & Shiffman, 1994). The factor validity and reliability of the PSI-VS was verified and supported in a total sample of 1338 French adolescents. In addition, Maïano et al.'s (2008) results supported the factorial invariance of the PSI-VS measurement model across genders but revealed latent mean differences showing that girls presented a lower level on most PSI dimensions. They also reported elevated latent factor correlations between the PSI-VS subscales, although these correlations still provided evidence of discriminant validity. No other attempt was made to replicate these results on new samples of "normal" adolescents. This is worrisome since a single study is insufficient to reach clear conclusions regarding the psychometric properties of an instrument. Moreover, the limitations mentioned by Maïano et al. (2008) stress the need for cross-validation efforts.

First, Maïano et al.'s (2008) study was based on a sample of normal-weight adolescents. It thus remains unknown whether the observed psychometric properties could generalize to youths with different weight statuses. Since overweight adolescents (or skinny

* Corresponding author at: University of Sherbrooke, Department of Psychology, 2500 boulevard de l'Université, Sherbrooke, QC, J1K 2R1 Canada.
E-mail address: alexandre.morin@usherbrooke.ca (A.J.S. Morin).

[1] Since both authors (A.J.S.M. and C.M.) contributed equally to the preparation of this paper, their order of appearance was determined at random: both should be considered first authors.

1740-1445/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bodyim.2011.06.005

ones) present a higher risk of being discriminated against on the basis of their weight, the resulting stigmatization may strongly influence their general and physical self-concepts (e.g., Puhl & Latner, 2007). Therefore, when overweight and obese adolescents are compared with normal-weight peers, they tend to present significantly lower levels of physical self-perceptions (e.g., Hau, Sung, Yu, Marsh, & Lau, 2005; Marsh, Hau, Sung, & Yu, 2007). However, the validity of this conclusion relies on the often-untested assumption that the measurement model used to assess physical self-concept is invariant across weight categories; whereas it is possible that overweight or obesity may modify the way physical self-concept is organized. To our knowledge, this assumption was only verified in a sample of Chinese children (Hau et al., 2005).

Second, although Maïano et al. (2008) confirmed the gender-based measurement invariance of the PSI-VS and observed significant mean-level differences favouring boys on most dimensions, these results need to be replicated. In addition, they did not examine the measurement invariance of the PSI-VS across age categories (i.e., early and late adolescence) and ethnicities. Adolescence is a period of multiple transformations that exert a determining impact on how youths perceive themselves generally and physically (Steinberg & Morris, 2001). Whereas early adolescence is characterized by efforts to cope with these multiple changes, late adolescence is a period where these changes are progressively integrated into a new self-definition. Gender and ethnicity are known to influence how these changes, particularly pubertal maturation, are integrated into adolescents' self-concepts (Morin, Maïano, Marsh, Janosz, & Nagengast, 2011; Siegel, Yancey, Aneshensel, & Schuler, 1999). For instance, whereas the Caucasian European/North American culture emphasizes lean "prepubertal" looks for girls, other cultures put less emphasis on leanness and more on the fuller forms emerging with puberty (Siegel et al., 1999). These observations suggest that ethnicity, age and gender may influence not only adolescents' average physical self-conceptions, but also the way physical self-concept is defined and organized in subgroups formed on the basis of these variables.

Finally, the convergent validity of the PSI-VS was never evaluated. Indeed, if the PSI-VS correctly measures what it is supposed to measure, it should be significantly related to constructs known to be related to physical self-concept, such as global self-esteem, physical self-image congruence, social physique anxiety, and disturbed eating attitudes and behaviours (e.g., Crocker et al., 2003; Hagger et al., 2010; Marsh et al., 2007).

In a recent review of physical self-concept instruments, Marsh and Cheng (in press) included a single non-English instrument, the PSI, and noted that the PSI-VS makes "*a potentially important contribution to applied research. However, further research is needed to more fully evaluate the robustness of support for construct validity*." In this study, we thus: (a) examine the factor validity the PSI-VS among subsamples of adolescents defined according to gender, age, ethnicity and weight; and (b) test the convergent validity of the PSI-VS with measures of global self-esteem, social physique anxiety, disturbed eating attitudes and behaviours and physical self-image congruence.

Method

Samples and Procedure

A first sample of 885 adolescents (aged 11–18; M_{age} = 15.71, SD_{age} = 1.66), composed of 341 boys and 544 girls, was recruited from eight middle and high schools located in Southern (n = 578) and Northern (n = 307) France. A cross-validation sample of 218 adolescents (M_{age} = 14.36 years, SD_{age} = 1.56), composed of 88 boys and 130 girls, was recruited from two middle schools located in Southern France. All participants were schooled in regular classes and, as such, presented no intellectual, motor or sensory disability according to French education policies. The questionnaires were administered in quiet classroom conditions and participants gave written informed consent. Both samples did not differ from one another on gender (χ^2 = 0.25; df = 1; p > .05) or weight category (χ^2 = 1.75; df = 2; p > .05), but Sample 2 was somewhat younger (t = 10.92; df = 1101; $p \leq .01$) and included fewer minority youths (χ^2 = 99.96; df = 1; $p \leq .01$). Supplementary analyses (see the online supplements) show that the results were fully replicated and invariant across samples and geographic areas (Northern versus Southern France). Thus, both samples were merged to ensure parsimony and to maximize sample size in each category. This combined sample included: (a) 1103 participants (M_{age} = 15.45, SD_{age} = 1.73); (b) 429 boys and 674 girls; (c) 343 early (aged 11–14) and 760 late adolescents (aged 15–18); (d) 930 participants whose parents are of European origin and 173 whose parents are of other origins (mostly from North Africa); and (e) 124 underweight, 877 normal-weight and 102 overweight students (including 12 obese). Weight categories were determined from the participants' body mass index (BMI) and gender-and age-specific cut-off scores (Cole, Bellizzi, Flegal, & Dietz, 2000; Cole, Flegal, Nicholls, & Jackson, 2007).

Measures

Participants' physical self-concept was assessed with the previously described PSI-VS (Maïano et al., 2008). Participants were also asked to report their height, weight, gender, age and parents' birth country (as a proxy for ethnicity). In order to assess the convergent validity of the PSI-VS, four instruments were used. First, participants' disturbed eating attitudes and behaviours were assessed with the Eating Attitudes Test (EAT: Garner, Olmstead, Bohr, & Garfinkel, 1982), which includes three subscales (all items are rated on a six-point scale): (a) Dieting (EAT-D; 13 items), (b) Bulimia and Food Preoccupation (EAT-BFP; six items) and (c) Oral Control (EAT-OC; seven items). Second, the seven items (rated on a five-point scale) from the Social Physique Anxiety Scale (SPAS: Hart, Leary, & Rejeski, 1989) were used to assess the degree to which participants become anxious about the real or perceived evaluation of their physique by others. Third, participants' physical self-image congruence was assessed with the Silhouette Matching Task (SMT: Marsh & Roche, 1996). The SMT asks participants to match themselves to a series of nine body silhouette images that vary from very thin to very fat (different series are presented for boys and girls) in relation to actual and ideal body image ratings. The discrepancies between the two ratings yield information regarding the participants' satisfaction with their physical self-image. Finally, participants from Sample 2 also completed the 10 items (rated on a four-point scale) from the Rosenberg Self-Esteem Inventory (RSEI: Rosenberg, 1965). The PSI-VS items and an extended presentation of the psychometric properties of the French versions of the convergent validity measures are reported in the online supplements.

Data Analysis

Analyses were conducted with Mplus 6.0 (Muthén & Muthén, 2010), using the robust full-information maximum likelihood (MLR) estimator, to correct for the small amounts of missing data (0.36–4.81%; $M_{missing}$ = 1.51%; $SD_{missing}$ = 1.18%). In the first stage, the a priori factor structure of the PSI-VS was tested with a confirmatory factor analytic (CFA) model hypothesizing that: (a) answers to the PSI-VS would be explained by six correlated factors (previously defined); (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) error terms would be uncorrelated. Relying on two indicators per construct creates locally unidentified factors even though the model remains identified with more than two factors (Bollen, 1989). A second model was thus estimated in which each factor was fully identified by using essentially tau-equivalent constraints (ETECs). Using ETECs involves placing equality constraints on both loadings to help locate the construct at the true centroid of the indicators (Little, Lindenberger, & Nesselroade, 1999). This procedure may results in a decrease in the fit of the models, which should not overly concern researchers if it is not dramatic (Little et al., 1999).

In the second stage, the measurement invariance of the PSI-VS was tested in the following sequence (e.g., Meredith, 1993): (a) configural invariance without ETECs (loadings, intercepts and uniquenesses are freely estimated, latent variances are constrained to 1 and latent means are constrained to 0); (b) loadings (weak) invariance without ETECs (constraining the loadings to invariance allows the variances to be freely estimated in all but one group); (c) loadings invariance with ETECs; (d) intercepts (strong) invariance (constraining the intercepts to invariance allows the means to be freely estimated in all but one group); (e) uniquenesses (strict) invariance; (f) variance–covariance invariance; and (g) latent means invariance. In each sequence of invariance, the preceding model served as reference. Models a and b were estimated without ETECs. If the ETECs were included directly in Model a, then Model b would have the same number of degrees of freedom as Model a (freeing up six variances while constraining six pairs of loadings to equality).

Assessment of model fit was based on multiple indicators (Hu & Bentler, 1999; Marsh, Hau, & Grayson, 2005): the chi-square (χ^2), the comparative fit index (CFI), the Tucker–Lewis index (TLI), the root mean square error of approximation (RMSEA), the 90% confidence interval of the RMSEA, and the standardized root mean square residual (SRMR). Values greater than .90 for the CFI and TLI indicate adequate model fit, although values greater than .95 are preferable. Values smaller than .08 or .06 for the RMSEA and than .10 and .08 for the SRMR support respectively acceptable and good model fit. Scale score reliability was computed from the CFA standardized parameter estimates, using McDonald's (1970) $\omega = \left(\sum|\lambda i|\right)^2 / \left(\left[\sum|\lambda i|\right]^2 + \sum \delta ii\right)$ where λi are the factor loadings and δii, the error variances. Compared with traditional scale score reliability estimates (e.g., alpha; see Sijtsma, 2009), ω has the advantage of taking into account the strength of association between items and constructs (λi) as well as item-specific measurement errors (δii). Measurement-invariance tests were evaluated by examining χ^2 difference tests (see the online supplements). However, studies suggest complementing this information with changes in CFIs and RMSEAs (Chen, 2007; Cheung & Rensvold, 2002). A ΔCFI of .01 or less and a ΔRMSEA of .015 or less between a more restricted model and the preceding one indicate that the invariance hypothesis should not be rejected. Since indices incorporating a penalty for parsimony (i.e., TLI and RMSEA) can also improve in more restricted models, ΔTLIs are also inspected (Marsh, Hau, & Grayson, 2005).

Results

Factor Validity

The goodness-of-fit statistics and standardized parameter estimates of the a priori CFA measurement model are respectively displayed in Tables 1 and 2. First, the results showed that the a priori six-factor CFA model provided an adequate degree of fit to the data with: (a) significant χ^2 values; (b) CFI and TLI > .95; (c) RMSEA and SRMR < .05. Using ETECs resulted in a slight decrease in fit according to the χ^2 difference test, but in an equivalent fit according to changes in fit indices (remaining under the .01 criterion). These results suggest the appropriateness of the ETECs. The standardized results from this model, reported in Table 2, reveal that all loadings are substantial and significant and that the scale score reliability coefficients are modest to acceptable. Modification indices confirmed the adequacy of this model. However, the latent correlations are quite elevated (r = .44–.87). These results are consistent with those from previous studies of the PSI (Maïano et al., 2008) and of other instruments (e.g., Marsh, Asçl, & Marco, 2002; Marsh & Cheng, in press).

In contrast with the results from some previous studies (e.g., Fox & Corbin, 1989; Maïano et al., 2008), the current results revealed that the GSW scale exhibited a stronger relation with PA than with PSW. This is consistent with the observation made by previous scholars that physical appearance represents a major component of global self-esteem in modern Western societies (e.g., Harter, 1999). However, the relation between GSW and PSW was stronger than between GSW and the remaining subscales, supporting previous results. In addition, and in conformity with previous results, all subscales were significantly and positively correlated with PSW and exhibited stronger relations with PSW than with GSW (with the exception of PA). These comparisons are all significant according to Steiger's (1980) criteria (see the online supplements). To assess the discriminant validity of factors, Bagozzi and Kimmel (1995) suggest ensuring that the 95% confidence intervals for the correlations exclude 1. This verification confirms the discriminant validity of the PSI-VS factors. In addition, post hoc models in which highly correlated factors were combined were also estimated (see the online supplements) and failed to provide a better fit to the data than the a priori model.

Measurement Invariance

To ensure meaningful between-group comparisons, it needs to be shown that the measurement scales are psychometrically equivalent across subsamples. Measurement-invariance tests were thus conducted across genders (Models 2-1 to 2-7), age categories (Models 3-1 to 3-7), parental origins (Models 4-1 to 4-7) and weight categories (Models 5-1 to 5-7). The results from these tests were highly similar and showed that (a) all of the χ^2 were significant but few of the χ^2 difference tests were significant (i.e., when ETECs were added, when latent means were constrained to equality across genders and age groups, and when the variance–covariance matrix was constrained to equality across age groups); (b) the CFI, TLI, RMSEA and SRMR indicated adequate model fit at all steps; (c) the ΔCFI never exceeded .01; (d) the ΔRMSEA remained under .015; and (e) the fit indices that control for model parsimony were similar or slightly better at the end of the sequence than at the beginning. These results confirm the full measurement invariance of the PSI-VS across genders, age groups, parental origins and weight categories. The only systematic decrease in fit was associated to ETECs, and this observation parallels the results from the main model. Since this slight decrease was not dramatic (Little et al., 1999), ETECs were kept to retain locally identified factors.

Given that gender- and age-related differences in latent means were of substantive interest and that the χ^2 differences were significant at this step, latent mean differences across genders and age categories were probed. For gender, the results showed that when boys' latent means are fixed to zero, girls' latent means were significantly ($p \leq .01$), but slightly (by about a quarter of a *SD*), lower on all subscales: (a) −.243 on GSW; (b) −.305 on PSW; (c) −.188 on PC; (d) −.308 on SC; (e) −.176 on PA; and (f) −.329 on PS. For age, the results showed that when early adolescents' latent means are fixed to zero, late adolescents' latent means were slightly but significantly ($p \leq .01$) higher on PS (latent mean = .221). Early and late adolescents did not differ significantly on other subscales. These

Table 1
Goodness-of-fit statistics of confirmatory factor analytic (CFA) models.

Model	Description	$\chi^2(df)$	CFI	TLI	RMSEA	90% CI	SRMR	$\Delta R\chi^2(df)$	ΔCFI	ΔTLI	ΔRMSEA
CFA	1-1. A priori 6-factor model	139.142(39)*	.982	.970	.048	.040–.057	.018	–	–	–	–
	1-2. A priori 6-factor model with ETEC	186.720(45)*	.975	.963	.053	.046–.061	.038	53.174(6)*	–.007	–.007	+.005
CFA: genders	2-1. Configural invariance without ETEC	172.198(78)*	.983	.971	.047	.037–.056	.021	–	–	–	–
	2-2. λ invariant without ETEC	183.131(84)*	.982	.972	.046	.037–.055	.025	10.535(6)	–.001	+.001	–.001
	2-3. λ invariant with ETEC	226.968(90)*	.975	.964	.053	.044–.061	.041	50.536(6)*	–.007	–.008	+.006
	2-4. λ, τs invariant	232.894(96)*	.975	.966	.051	.043–.059	.042	4.441(6)	.000	+.002	–.002
	2-5. λ, τs, δs invariant	238.448(108)*	.975	.971	.047	.039–.055	.044	9.187(12)	.000	+.005	–.004
	2-6. λ, τs, δs, ξ/φ invariant	269.634(129)*	.975	.974	.044	.037–.052	.052	30.089(21)	.000	+.003	–.003
	2-7. λ, τs, δs, ξ/φ, ηs invariant	294.769(135)*	.971	.972	.046	.039–.054	.066	27.266(6)*	–.004	–.002	+.002
CFA: age	3-1. Configural invariance without ETEC	183.510(78)*	.982	.969	.050	.040–.059	.023	–	–	–	–
	3-2. λ invariant without ETEC	191.807(84)*	.981	.970	.048	.039–.057	.025	7.037(6)	–.001	+.001	–.002
	3-3. λ invariant with ETEC	238.094(90)*	.974	.962	.055	.046–.063	.042	52.029(6)*	–.007	–.008	+.007
	3-4. λ, τs invariant	246.040(96)*	.974	.964	.053	.045–.061	.041	7.069(6)	.000	+.002	–.002
	3-5. λ, τs, δs invariant	261.711(108)*	.973	.967	.051	.043–.059	.045	17.745(12)	–.001	+.003	–.002
	3-6. λ, τs, δs, ξ/φ invariant	305.826(129)*	.969	.968	.050	.043–.057	.057	43.707(21)*	–.004	+.001	–.001
	3-7. λ, τs, δs, ξ/φ, ηs invariant	331.801(135)*	.966	.966	.051	.044–.058	.060	26.970(6)*	–.003	–.002	+.001
CFA: origins	4-1. Configural invariance without ETEC	190.454(78)*	.980	.967	.051	.042–.060	.022	–	–	–	–
	4-2. λ invariant without ETEC	196.496(84)*	.980	.969	.049	.040–.058	.024	4.694(6)	.000	+.002	–.002
	4-3. λ invariant with ETEC	243.085(90)*	.973	.960	.056	.047–.064	.041	52.097(6)*	–.007	–.009	+.007
	4-4. λ, τs invariant	247.126(96)*	.973	.963	.053	.045–.062	.041	2.505(6)	.000	+.003	–.003
	4-5. λ, τs, δs invariant	271.007(108)*	.971	.965	.052	.045–.060	.046	24.952(12)	–.002	+.002	–.001
	4-6. λ, τs, δs, ξ/φ invariant	284.801(129)*	.971	.972	.047	.039–.054	.052	11.779(21)	.000	+.007	–.005
	4-7. λ, τs, δs, ξ/φ, ηs invariant	289.949(135)*	.971	.973	.046	.038–.053	.057	4.452(6)	.000	+.001	–.001
CFA: weight	5-1. Configural invariance without ETEC	232.715(117)*	.981	.967	.052	.042–.062	.024	–	–	–	–
	5-2. λ invariant without ETEC	250.577(129)*	.980	.969	.051	.041–.060	.032	16.979(12)	–.001	+.002	–.001
	5-3. λ invariant with ETEC	300.239(135)*	.972	.960	.058	.049–.066	.048	53.408(6)*	–.008	–.009	+.007
	5-4. λ, τs invariant	316.966(147)*	.971	.962	.056	.048–.065	.048	15.687(12)	–.001	+.002	–.002
	5-5. λ, τs, δs invariant	349.766(171)*	.970	.965	.053	.045–.061	.050	34.181(24)	–.001	+.003	–.003
	5-6. λ, τs, δs, ξ/φ invariant	385.575(213)*	.970	.973	.047	.039–.054	.056	36.011(42)	.000	+.008	–.006
	5-7. λ, τs, δs, ξ/φ, ηs invariant	402.712(225)*	.970	.974	.046	.039–.054	.059	16.608(12)	.000	+.001	–.001

Note: χ^2: chi-square; *df*: degrees of freedom; CFI: comparative fit index; TLI: Tucker–Lewis index; RMSEA: root mean square error of approximation; 90% CI: 90% confidence interval of the RMSEA; SRMR: standardized root mean square error of approximation; ETEC: essentially tau-equivalent constraints; λ: loading; τ: intercept; δ: uniqueness; ξ: variance; φ: covariance; η: factor means; $\Delta R\chi^2$: Robust chi-square difference tests (calculated from loglikelihoods for greater precision); Δ: change from previous model.

* $p < .01$.

Table 2
Standardized parameters estimates from the confirmatory factor analytic (CFA) model and convergent validity analyses.

Items	GSW (λ)	PSW (λ)	PC (λ)	SC (λ)	PA (λ)	PS (λ)	δ
Standardized factor loadings and uniquenesses from the CFA model							
GSW1	.762*						.419*
GSW2	.609*						.629*
PSW1		.855*					.269*
PSW2		.861*					.259*
PC1			.927*				.141*
PC2			.883*				.221*
SC1				.823*			.322*
SC2				.862*			.256*
PA1					.784*		.385*
PA2					.756*		.429*
PS1						.727*	.472*
PS2						.703*	.506*
ω	.64	.85	.90	.83	.74	.68	

Factor	PSW	PC	SC	PA	PS
Latent factor correlations from the CFA model (95% confidence intervals)					
GSW	.78* (.72/.83)	.44* (.37/.51)	.57* (.50/.63)	.84* (.78/.90)	.64* (.57/.71)
PSW		.70* (.66/.75)	.87* (.84/.91)	.66* (.61/.72)	.80* (.75/.85)
PC			.68* (.63/.73)	.47* (.40/.53)	.67* (.62/.73)
SC				.60* (.54/.66)	.86* (.91/.82)
PA					.72* (.78/.65)

	GSW	PSW	PC	SC	PA	PS
Latent factor correlations from the convergent validity analyses (95% confidence intervals)						
EAT-D	−.48* (−.56/−.40)	−.29* (−.36/−.22)	−.22* (−.28/−.16)	−.20* (−.27/−.13)	−.30* (−.38/−.23)	−.16* (−.24/−.08)
EAT-BFP	−.30* (−.40/−.19)	−.17* (−.25/−.08)	−.10* (−.18/−.02)	−.04 (−.01/.04)	−.08 (−.17/.02)	.05 (−.05/.15)
EAT-OC	−.14* (−.23/−.05)	−.09* (−.17/−.02)	−.01 (−.08/.07)	−.08 (−.15/.00)	.05 (−.04/.13)	−.07 (−.16/.02)
SPAS	−.50* (−.59/−.42)	−.32* (−.40/−.25)	−.18* (−.25/−.11)	−.22* (−.29/−.14)	−.31* (−.39/−.23)	−.18* (−.27/−.10)
SMT	.33* (.26/.40)	.28* (.21/.36)	.29* (.22/.35)	.25* (.18/.32)	.34* (.27/.42)	.17* (.09/.24)
RSEI (S2)	.70* (.47/.93)	.54* (.35/.72)	.35* (.15/.54)	.30* (.13/.47)	.48* (.29/.67)	.41* (.24/.58)

Note: λ: standardized factor loading; δ: standardized item's uniquenesses; ω: McDonald (1970) scale score reliability coefficient; GSW: Global Self-Worth; PSW: Physical Self-Worth; PC: Physical Condition; SC: Sport Competence; PA: Physical Attractiveness; PS: Physical Strength; EAT: Eating Attitude Test; EAT-D: Dieting subscale of the EAT; EAT-BFP: Bulimia and Food Preoccupation subscale of the EAT; EAT-OC: Oral Control subscale of the EAT; SPAS: Social Physique Anxiety Scale; SMT: Silhouette Matching Task (physical self-image congruence); RSEI: Rosenberg Self-Esteem Inventory (global self-esteem); S2: computed only in the second subsample.
* $p \leq .05$.

results generally replicate those from previous studies (e.g., Maïano et al., 2008; Marsh et al., 2007).

Convergent Validity

Finally, we assessed the convergent validity of the PSI-VS latent factors in relation to measures of disturbed eating attitudes and behaviours (EAT-D, EAT-BFP, EAT-OC), social physique anxiety (SPAS), physical self-image congruence (SMT) and global self-esteem (RSEI). The latent variable correlations are reported at the bottom of Table 2. These latent correlations show that GSW and more specific physical self-concept perceptions are negatively and modestly to moderately related to (a) disturbed eating attitudes and behaviours: EAT-D, EAT-BFP (except for the SC, PA and PS subscales) and EAT-OC (except for the subdomains); and (b) social physique anxiety. Additionally, all of the PSI-VS factors are positively and moderately correlated with (a) physical self-image congruence and (b) global self-esteem. These results are consistent with previously reported results for other physical self-concept instruments (e.g., Crocker et al., 2003; Hagger et al., 2010; Marsh et al., 2007), which reinforces the convergent and construct validity of the PSI-VS.

Discussion

The first objective of this study was to evaluate the robustness of the psychometric properties of the PSI-VS. The results confirmed that the a priori six-factor model of the PSI-VS provides a satisfactory fit to the data, and presented acceptable levels of scale score reliability. Although some of the observed reliability coefficients appear lower than conventional rules of thumb (i.e., .70 or .80), one of the potential explanations for this result involves the limited number of items used in the PSI-VS. Reliability coefficients are positively affected by the number of items in a scale and, consequently, acceptability levels must be adjusted for shorter measures (Streiner, 2003). However, this observation reinforces the need to rely on latent variable methodologies in studies based on the PSI-VS in order to take into account these moderate levels of reliability (e.g., Fan, 2003).

Since the convergent validity of the PSI-VS had never been evaluated, a second objective of the present study was to verify the convergent validity of the PSI-VS with measures of theoretically related constructs, such as global self-esteem, physical self-image congruence, social physique anxiety, and disturbed eating attitudes and behaviours. These analyses clearly confirmed the convergent validity of the PSI-VS and thus further support its overall construct validity.

Additional results indicate that the PSI-VS latent factor correlations are very elevated. Although these latent correlations are similar to those found in previous studies of instruments based on the PSPP (Maïano et al., 2008; Marsh et al., 2002; Marsh & Cheng, in press), they still "*bring into question the real independence of some of the models' sub-dimensions*" (Maïano et al., 2008, p. 844). Fortunately, in this study, the discriminant validity of the PSI-VS factors was confirmed, suggesting that these elevated latent correlations may be part of the PSI-VS. Marsh et al. (2002) suggested that inflated correlations could sometimes emerge when short scales attempt to cover a broad range of dimensions with few items. More recently, Marsh et al. (2010) suggested that the independent cluster model inherent in CFA—in which each item loads on a single factor—could be too restrictive for many multidimensional constructs and will result in inflated latent correlations unless all non-target loadings are close to zero. They proposed exploratory structural equation modeling (ESEM) as a way to relax this assumption, while retaining most advantages of CFAs, such as tests of measurement invariance. Unfortunately, ESEM models based on two indicators per construct are non-identified and could not be estimated, but should be investigated in future studies based on the 18-item PSI.

To further probe the robustness of these results, measurement-invariance tests were performed across subgroups of students, based on the assumptions that physical self-concept may be organized differently in these subgroups. First, the results confirm the complete measurement and latent mean invariance across parental origins and weight categories. Although previous studies suggest that average levels of physical self-concept may differ according to adolescents' ethnicity (e.g., Morin et al., 2011; Siegel et al., 1999; Twenge & Crocker, 2002), these studies also show that these effects differ according to specific ethnic groups (Twenge & Crocker, 2002) or may emerge only in the context of a three-way interaction involving gender and pubertal development (Morin et al., 2011; Siegel et al., 1999). Unfortunately, these distinctions could not be made in the present study due to the low number of participants forming the non-European subgroup. Similarly, although previous studies suggested that average levels of physical self-concept may be lower in overweight and obese adolescents than in normal-weight youths (Hau et al., 2005; Marsh et al., 2007), this observation applies to obese or combined obese–overweight samples, whereas the present sample included very few obese youths. Weight categories were also created from adolescents' self-reports rather than from objective measures of height/weight or body fat. It has been previously noted that self-reports tend to be biased as adolescents generally underestimate their weight and overestimate their height, a trend that is more pronounced in youths at risk of being overweight or obese (Sherry, Jefferds, & Grummer-Strawn, 2007). Moreover, at least some youths may have been classified as overweight due to strong muscular or bone structure, rather than body fat—thus inducing another bias. Alternatively, this lack of differences could also be related to the fact that the PSI-VS was developed specifically for in-depth longitudinal studies (e.g., Stone & Shiffman, 1994) and thus to be sensitive to intra-individual differences rather than to group differences. The absence of mean-level differences could thus be related to a loss of precision in group-based comparisons due to the selection of so few items. However, although this could suggest that the PSI-VS may not be optimal for group-based comparisons, the results confirm that its measurement model is fully invariant across ethnic and weight-related subgroups and could thus provide unbiased estimates of group-based differences.

Second, the results confirm the complete measurement invariance across age groups and genders and also show that: (a) girls tend to present lower average levels for most dimensions of the PSI-VS and that (b) late adolescents show some improvements in physical self-concepts. These results are consistent with those from previous studies (e.g., Maïano et al., 2008; Marsh et al., 2007), and suggest that the absence of mean-level differences across ethnic groups and weight categories may not be related to a lack of sensitivity after all.

Additional limitations must be taken into account. First, the present study is based on a convenience sample of normally achieving adolescents, which cannot be considered representative of the French adolescent population. This indicates that use of the PSI-VS should be limited to normally achieving French adolescents. Clearly, the next step in the evaluation of the PSI-VS would be to verify its applicability and validity with non-French populations. Second, although we found evidence of the convergent and construct validity of the PSI-VS, additional tests remain to be conducted in relation to (a) other physical self-concept instruments and (b) external criteria (i.e., fitness exams, percentage of body fat, etc.). Third, the scale score reliability and construct validity (i.e., in terms of latent mean comparisons) of the PSI-VS are modest, potentially due to the brevity of the scales. For these reasons, the use of this instrument should be limited for contexts where there is a clear need for a very short scale. This is encouraging, as the PSI-VS was developed specifically for these contexts. However, even though Maïano et al. (2008) relied on a rigorous state-of-the-art process to select the PSI-VS items (e.g., Marsh, Ellis, Parada, Richards, & Heubeck, 2005) in order to maximize the breadth of construct coverage (as confirmed by the convergent validity results), longer instruments do provide broader construct coverage and should be preferred whenever possible. Finally, although the PSI-VS was specifically designed for in-depth idiographic studies, it was never tested in this context. For the moment, we feel that this study accomplished an important step in this direction by confirming the robustness of this instrument across various subgroups and its convergent (in relation to other constructs) and construct (in relation to grouping variables) validity. But clearly, one of the next steps should be to verify the PSI-VS sensitivity to intra-individual changes in the context of in-depth idiographic assessment.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bodyim.2011.06.005.

References

Bagozzi, R. P., & Kimmel, S. K. (1995). A comparison of leading theories for the prediction of goal directed behaviours. *British Journal of Social Psychology*, *34*, 437–461.

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

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement. *Structural Equation Modeling*, *14*, 464–504.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255.

Cole, T. J., Bellizzi, M., Flegal, K., & Dietz, W. (2000). Overweight and obesity worldwide: International establishing a standard definition for child survey. *British Medical Journal*, *320*, 1240–1243.

Cole, T. J., Flegal, K., Nicholls, D., & Jackson, A. (2007). Body mass index cut offs to define thinness in children and adolescents: International survey. *British Medical Journal*, *335*, 194–197.

Crocker, P., Sabiston, C., Forrestor, S., Kowalski, N., Kowalski, K., & McDonough, M. (2003). Predicting change in physical activity, dietary restraint, and physique anxiety in adolescent girls: Examining covariance in physical self-perceptions. *Canadian Journal of Public Health*, *94*, 322–337.

Fan, X. (2003). Two approaches for correcting correlation attenuation caused by measurement error: Implications for research practice. *Educational & Psychological Measurement*, *63*, 915–930.

Fox, K. R., & Corbin, C. B. (1989). The physical self-perception profile: Development and preliminary validation. *Journal of Sport & Exercise Psychology*, *11*, 408–430.

Garner, D. M., Olmstead, M. P., Bohr, Y., & Garfinkel, P. E. (1982). The eating attitude test: Psychometric features and clinical correlates. *Psychological Medicine*, *12*, 871–878.

Hagger, M. S., Stevenson, A., Chatzisarantis, N. L. D., Pereira Gaspar, P. M., Leitão Ferreira, P. M., & González Ravé, J. M. (2010). Physical self-concept and social physique anxiety: Invariance across culture, gender and age. *Stress & Health*, *26*, 304–329.

Hart, E. H., Leary, M. R., & Rejeski, W. J. (1989). The measurement of social physique anxiety. *Journal of Sport & Exercise Psychology*, *11*, 94–104.

Harter, S. (1999). *The construction of the self: A developmental perspective*. New York, NY: Guilford.

Hau, K. T., Sung, R. Y. T., Yu, C. W., Marsh, H. W., & Lau, P. W. C. (2005). Factorial structure and comparison between obese and nonobese: Chinese children's physical self-concept. In H. Marsh, R. Craven & D. McInerney (Eds.), *The new frontiers of self research* (pp. 259–272). Charlotte, NC: Information Age.

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

Little, T. D., Lindenberger, U., & Nesselroade, J. R. (1999). On selecting indicators for multivariate measurement and modeling with latent variables: When "good" indicators are bad and "bad" indicators are good. *Psychological Methods*, *4*, 192–211.

Maïano, C., Morin, A. J. S., Ninot, G., Monthuy-Blanc, J., Stephan, Y., Florent, J.-. F., et al. (2008). A short and very short form of the physical self-inventory for adolescents: Development and factor validity. *Psychology of Sport & Exercise*, *9*, 830–847.

Marsh, H. W., Asçl, F. H., & Marco, I. T. (2002). Multi-trait multi-method analyses of two physical self-concept instruments: A cross-cultural perspective. *Journal of Sport & Exercise Psychology*, *24*, 99–119.

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

Marsh, H. W., Ellis, L., Parada, R., Richards, G. E., & Heubeck, B. (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.

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.

Marsh, H. W., Lüdtke, O., Muthén, B. O., Asparouhov, T., Morin, A. J. S., Trautwein, U., et al. (2010). A new look at the big-five factor structure through exploratory structural equation modeling. *Psychological Assessment*, *22*, 471–491.

Marsh, H. W., & Roche, L. A. (1996). Predicting self-esteem from perceptions of actual and ideal ratings of body fatness: Is there only one ideal supermodel? *Research Quarterly for Exercise & Sport*, *6*, 13–23.

McDonald, R. P. (1970). Theoretical foundations of principal factor analysis, canonical factor analysis, and alpha factor analysis. *British Journal of Mathematical & Statistical Psychology*, *23*, 1–21.

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

Morin, A. J. S., Maïano, C., Marsh, H. W., Janosz, M., & Nagengast, B. (2011). The longitudinal interplay of adolescents' self-esteem and body image: A conditional autoregressive latent trajectory analysis. *Multivariate Behavioral Research*, *46*, 157–201.

Muthén, L. K., & Muthén, B. (2010). *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén.

Ninot, G., Delignières, D., & Fortes, M. (2000). L'évaluation de l'estime de soi dans le domaine corporel. *Sciences et Techniques des Activités Physiques et Sportives*, *53*, 35–48.

Puhl, R. M., & Latner, J. D. (2007). Stigma, obesity, and the health of the nation's children. *Psychological Bulletin*, *133*, 557–580.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University.

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

Sherry, B., Jefferds, M.-E., & Grummer-Strawn, L. M. (2007). Accuracy of adolescent self-report of height and weight in assessing overweight status. *Archives of Pediatrics & Adolescent Medicine*, *161*, 1154–1161.

Siegel, J. M., Yancey, A. K., Aneshensel, C. S., & Schuler, R. (1999). Body image, perceived pubertal timing, and adolescent mental health. *Journal of Adolescent Health*, *25*, 155–165.

Sijtsma, K. (2009). On the use, misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*, 107–120 (Introduction to a special issue)

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*, 245–251.

Steinberg, L., & Morris, A. (2001). Adolescent development. *Annual Review Psychology*, *52*, 83–110.

Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment in behavioral medicine. *Annals of Behavioral Medicine*, *16*, 199–202.

Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, *80*, 99–103.

Twenge, J. M., & Crocker, J. (2002). Race and self-esteem: Meta-analyses comparing Whites, Blacks, Hispanics, Asians, and American Indians and comment on Gray-Little and Hafdahl (2000). *Psychological Bulletin*, *128*, 371–408.