Latent Profile Analysis of the Physical Self-Description among Chinese Adolescents

Chee Keng John Wang[1,2]

Yanlin Sun[2]

Woon Chia Liu[1]

Jiaxin Yao[2]

&

Do Young Pyun[1]

[1]National Institute of Education, Nanyang Technological University, Singapore

[2]Tianjin University of Sport, China

Correspondence:

Professor Yanlin Sun

Tianjin University of Sport

51 Weijinnan Road, Hexi District, Tianjin, China

China (300381)

Tel: (86) 022-23012624

Fax: (86) 22-23012624

E-mail: ylsun1003@163.com

Latent Profile Analysis of the Physical Self-Description among Chinese Adolescents

Abstract

The purposes of this study were to validate the Physical Self-Description Questionnaire (PSDQ-S) and examine the physical self-description profiles using Latent Profile Analysis with a Chinese sample. A total of 744 secondary school students in China took part in the study. While the results provided support for internal reliability and discriminant validity of the PSDQ-S, they indicated that convergent validity required further testing. In addition, three distinct profiles were identified with unique physical self-concept and different levels of physical activity participation. The study showed that the PDSQ-S is useful in differentiating groups of adolescents with different levels of physical self-concept.

Keywords: PSDQ-S; latent profile analysis, confirmatory factor analysis, adolescents

Latent Profile Analysis of the Physical Self-Description among Chinese Adolescents

In education setting, a positive self-concept is an influential predictor of significant outcomes such as academic achievement, psychological well-being, and motivation (Marsh, 1988; Paradise & Kernis, 2002). In sport and physical education setting, physical self-concept is even more important as it has been found to be a significant indicator of health outcomes such as physical activity (Fox, 2000; Sonstroem, Harlow, & Josephs, 1994), physical fitness (Marsh, 1996; Marsh & Redmayne, 1994; van Vorst, Buckworth, & Mattern, 2002), eating behaviour (Monthuy-Blanc, Maïano, Morin, & Stehan, 2012), and well being (Paradise & Kernis, 2002). In a recent review of 113 studies with 128 effect sizes by Spence and his colleagues (Spence, McGannon, & Poon, 2005), a reciprocal effect of self-esteem was also found that participation in exercise had a positive effect on global self-esteem or self-concept ($d = .23$, $SE = .02$). Change in physical fitness and type of programme were two significant moderators of the effect of exercise on global self-esteem. This supports that exercise participation results in significant improvements in global self-esteem. The exercise and self-esteem model used by Spence et al. (2005) was developed on a multifaceted hierarchical concept of self-concept.

There is general consensus that self-concept is the conscious perceptions that individuals have of themselves and include both descriptive and evaluative content (Harter, 1996). Self-concept has been widely accepted as a multifaceted and hierarchically organised structure, based upon the model proposed by Shavelson, Hubner, and Stanton (1976). A multidimensional approach is preferred over traditional approaches (unidimensional) because it recognises that global self-concept arises from multiple sources and across different domains (Marsh, Hey, Roche, & Perry, 1997). These domains could be social, academic, physical and occupational. Within each domain, there could be numerous subdomains that represent the inferences about

oneself concerning his or her competencies in more specific contexts. The global self-concept is seen as relatively stable compared with domain and subdomain-level self-evaluations.

One advantage of the Shavelson et al.'s (1976) model is that it allows researchers to better understand self-concepts study self-concept in a single domain and yet maintain the relevance of the domain to global self-concept. Thus many domain specific measures of self-concepts were designed to differentiate self-concepts with regard to different domains. For example, Marsh and his colleagues (Marsh, Richards, Johnson, Roche, & Tremayne, 1994) developed the Physical Self-Description Questionnaire (PSDQ) to measure self-concept specific in the physical domain.

The original PSDQ contains 70 items with 11 self-concept factors: strength, body fat, activity, endurance and fitness, sport competence, coordination, health, appearance, flexibility, global physical self-worth, and global esteem. When used with other measures or inventories, the PSDQ is often perceived as too long and take too long to complete for applied research. To address this concern, Marsh and his colleagues (Marsh, Martin, & Jackson, 2010) recently constructed a short version of the Physical Self-Description Questionnaire (PSDQ-S). The 40-item PSDQ-S retains the same factor structure and reliability of the original 70-item PSDQ (Marsh, 1996; Marsh et al., 1994). They provided strong evidence for the reliability, invariance, convergent and discriminant validity of the subscales.

Since the development of the short version, not many studies have validated the applicability of the PSDQ-S to other countries and different cultures, particularly, non-English speaking countries with a collectivistic culture. There is a need to examine the psychometric properties of the PSDQ-S when use in a new population.

Typically, research in this area often uses a variable-centered approach that compares the structure of physical self-concept across elite athletes and physical education students (e.g., Marsh et al., 1997), age and sex effects (e.g., MaÔano, Ninot, & Bilard, 2004), and PE major and non PE major students (e.g., Chung, 2003). Recently, Marsh and his colleagues (Marsh, Lüdtke, Trautwein, & Morin, 2009) propose the use of latent profile analysis (LPA) to explore a person-centered approach to the study of multi-dimensional aspects of academic self-concept. Using this approach, homogenous grouping of students with similar profiles of academic self-concept can be identified and the profiles can then be validated across different correlates.

LPA focuses on grouping individuals with similar characteristics based on the values on the indicator variables. It is similar to the traditional cluster analysis methods. However, cluster analysis is an exploratory and LPA is a model-based approach whereby the researcher can hypothesize a statistical model for the population from which the sample is drawn. The results of the LPA provide the statistical likelihood of each variable and the probability of the group membership and there are several fit indexes that allow model comparisons and informed choices can be made (Marsh et al., 2009). Thus, LPA is a better approach for profiling individuals, compared to the traditional cluster analysis.

Purposes of the Current Study

The first purpose of this study was to validate the PSDQ-S with a Chinese sample. Specifically, we sought to examine reliability and validity (convergent and discriminant) of the subscales in PSDQ-S. The second purpose was to examine the physical self-description profiles among Chinese adolescents using a LPA approach. Finally, the study sought to validate the latent profiles with the level of physical activity participation among the sample.

Method

Participants

A total of 744 secondary school students from four schools in China were recruited for the study. The students were aged from 11 to 16 years (M = 13.29, SD = .94). There were 271 males and 473 females. They were representative of diverse socio-economic backgrounds.

Procedures

After securing permission from the head teachers, the heads of PE department of the schools were contacted, and arrangements for survey administration were made. Administration of the questionnaires took place in quiet classroom conditions under the supervision of one of the researcher. Students were told that their participation in the study was voluntary, and they were free to withdraw at any time and were assured that their responses would be kept confidential. All students gave informed consent and took about 10 minutes to complete the PSDQ-S administered at the beginning of their PE lessons. Ethical approval was granted from the host University's Institutional Review Board.

Measures

Physical self-description questionnaire - Short form. The PSDQ-S has 40 items that measure nine specific domains of components of physical fitness and competence (i.e., strength, body fat, activity, endurance/fitness, sports competence, coordination, health, appearance, and flexibility) and two global components (i.e., global physical competence and global self-esteem). In this study, the PDSQ-S was translated to Chinese using a forward and backward translation technique (Brislin, 1990). First, a local Chinese who is bilingual (Chinese and English) translated the questionnaire from English to Chinese. Thereafter, a Singaporean Chinese whom first language is English and well versed in Chinese did a back translation from Chinese to English.

The back-translate English version was compared with the original questionnaire. Both translators are experts in sport psychology. Participants' responses to each item were scored using a 5-point Likert scale ranged from "strongly disagree" (1) to "strongly agree" (5).

Physical activity participation. There were two items measuring the frequency of physical activity participation and total number of hours per week involved in physical activity. The frequency of participation ranged from 1 to 4 (1= occasionally, 2 = once or twice per week, 3 = three to four times per week, and 4 = five or more times a week). The duration of physical activity participation was made into a 4-point scale (1 = less than one hour, 2 = one to three hours, 3 = three to six hours, and 4 = more than six hours).

Data Analysis

In the initial analysis, the descriptive statistics and internal consistency coefficients of the main variables were computed. We examined the internal consistency for each subscale of the PSDQ-S by calculating the rho's coefficients and alpha coefficients. A composite reliability coefficient (rho) of greater than .60 is considered acceptable (Bagozzi & Yi, 1988), whereas alpha values of greater than .70 is considered satisfactory (Nunnally & Bernstein, 1995). Cronbach's (1951) alpha coefficients are based on the assumption that there are no measurement error covariances, probably resulting in a bias at the population level (Raykov, 1998). The use of the rho's coefficient corrects for this "bias". Raykov (1998) has demonstrated that Cronbach's alpha may over- or under-estimate scale reliability and for this reason, rho is now preferred and may lead to higher estimates of true reliability. For convergent validity of the measures, we computed the average variance extracted (AVE) values for each scale. The AVE index is a measure of the shared or common variance in a latent variable, that is, the amount of variance that is captured by the latent variable in relation to the amount of variance due to measurement error (Dillion & Goldstein, 1984). This provides a measure of convergent validity, and a value greater than .50 is considered

acceptable (Fornell & Larcker, 1981). For discriminant validity, the confidence intervals of the latent factor correlation between each pair of factors were examined ($\Box$-coefficients). If the correlations are significantly less than unity, discriminant validity of the measure is supported (Bagozzi, 1981).

We conducted a confirmatory factor analyses (CFA) on the 11-factor measurement model of PSDQ-S, proposed by Marsh et al. (2009). The CFA was conducted using EQS for Windows 6.2 (Bentler & Wu, 1998) with maximum likelihood estimates derived from covariance matrices. Multiple indices of fit provided by EQS were examined to evaluate the adequacy of the models: Satorra-Bentler scaled chi-square statistic, the robust non-normed fit index (NNFI), the robust comparative fit index (CFI), the robust Bollen's incremental fit index (IFI), and the root mean square error of approximation (RMSEA) and its confidence intervals. Typically, for these fit indices, values greater than .95 were considered satisfactory (Hu & Bentler, 1999). The RMSEA was based on the analysis of residuals and compensates for the effects of model complexity. For this, Hu and Bentler (1999) recommended the cut-off of .06. Additionally, a LPA was conducted using Mplus 7.0 (Muthén, 2001; Muthén & Muthén, 2004). All the mean scores of the 11 subscales of the PDSQ-S were used as the grouping variables. The analyses were conducted from one to eight groups in order to select the best solution based on their fit statistics. The algorithm used was integration, and the number of initial stage randomly started at 2000 with 200 stages of final stage optimizations. The maximum number of iterations was 500, and the estimator used was Multivariate Linear Regression (MLR). The indices used to evaluate the fit of the model were entropy statistics (Muthén & Muthén, 2004), Akaike's information criterion (AIC; Akaike, 1987), Bayesian information criterion (BIC, Schwartz, 1978), sample-size adjusted BIC, and Lo-Mendell-and Rubin likelihood ratio test (LMR; Lo, Mendell, & Rubin, 2001). The entropy statistics higher than .60 indicates high classification utility (see Pastor, Barron, Miller, & Davis, 2007).

Lower values of both AIC and BIC are indicators of better model fit. The LMR test was used to compare an estimated model with a model whose class was one less than the estimated model. A small p-value associated with the LMR test supports the retention of a more complex solution with k clusters (Pastor et al., 2007). In addition, we also examined the models with number of groups that had less than 1% and less than 5% of the cases. It is suggested that solutions with small number of cases may not be feasible (Marsh et al., 2009). After the cluster profiles have been determined, a MANOVA was conducted using the cluster as grouping variable and the frequency and duration of physical activity participation as the outcome variables. This helped to verify the concurrent validity of the cluster solution.

Results

The mean scores, standard deviations, reliability coefficients, skewness and kurtosis for the PDSQ-S scales are shown in Table 1. All the PDSQ-S subscales had adequate internal consistency of alpha coefficients equal or higher than .70, and rho coefficients higher than .60. This indicates that all the subscales had satisfactory reliability. Univariate skewness and kurtosis statistics indicated that the observed variables in the main sample were approximately normal (within $\pm$ 1.00), except for sports competence. In terms of convergent validity, eight out of the eleven scales showed unsatisfactory AVE values. The results showed that several PDSQ-S subscales were highly correlated with each other. The latent factor correlations and its confidence intervals between each pair of scale are shown in Table 2. All the correlation coefficients and confidence intervals were significantly lower than 1, except for the confidence intervals between sports competence and endurance (with 90 % CI of SE [.78, 1.02]). In general, the scales of the PDSQ-S had adequate discriminant validity (Bagozzi, 1981). The results of the CFA on the measurement model of the PDSQ-S showed that the model fit the data adequately (Scaled χ^2 = 1188.43, df = 685; robust NNFI = .934, robust CFI = .942; robust IFI = .943, RMSEA = .031, 90% CI of RMSEA [.028, .034]).

The results of the latent profile analysis are presented in Table 3. From the values of AIC, BIC, and SSA-BIC, it can be seen that when the clusters increased from three to four clusters onwards, there were only marginal decrease (less than 0.5% decrease in SSA-BIC). The entropy values for all the models were greater than .70. The p values of the LMR for K versus K-1 classes were only significant for two- to three-group solutions. That is, the four-group solution was not significantly better than the three-group solution. Based on the pLMR and values of AIC, BIC and SSA-BIC, it was clear that a three-group solution was the best.

Figure 1 shows the profiles of the three clusters. Z scores of +/-.50 or greater were used as criteria to describe whether a group scored relatively 'high' or 'low' in comparison to their peers. In Cluster 1, there were 35% of the participants (N = 260), with 53.8% males and 46.2% females. The unique characteristics of this group was having all PDSQ-S scores higher than Z scores of .50, except for body fat (Z = .42) and health (Z = .37). Cluster 2 consisted of 17.5% of the sample (N = 130). This female-dominated group (84.6%) had the characteristics of low PDSQ-S scores of lower than Z = -.50. Cluster 3 (47.5% of the sample, N = 354) appeared to have a moderate PDSQ-S scores with 31.4% males and 68.6% females.

The results of the MANOVA using cluster as grouping variables and physical activity frequency and duration as outcome variables showed that the three cluster differed significantly in physical activity participation (Pillai's Trace = .88, F [22, 1464] = 52.47, p < .001, $\eta p2$ = .44). The results of the first ANOVA showed that the three clusters differed significantly in their duration of physical activity participation, F (2, 741) = 34.49, p < .001, $\eta p2$ = .08. Follow-up tests using Tukey's HSD showed that Cluster 1 reported highest duration of weekly physical activity duration, followed by cluster 3. Cluster 2 had the lowest score in the duration of physical activity participation. The pairwise comparisons indicated all the differences among the three clusters were significantly (ps < .05). Similarly, the results on the frequency of physical activity participation differed among the three clusters (F [2, 741] =

65.39, p < .001, η_p^2 = .15). Again, Cluster 1 had the highest score in the frequency of physical activity participation, and Cluster 2 reported lowest physical activity frequency (see Table 4). The pairwise comparisons indicated all the differences among the three clusters were significantly (ps < .05).

Discussion

The present study was designed to validate the PSDQ-S among Chinese adolescents and to examine the LPA of the physical self-concept, measured by the PSDQ-S. In this study, we examined the internal consistency, convergent validity, and discriminate validity of the subscales in PSDQ-S. In terms of reliability, all the scales of the PSDQ-S were found to have adequate internal consistency. In addition to Cronbach's (1951) coefficient alpha, we also computed the "unbiased" rho's coefficients which also supported the reliability of the subscales.

The test of discriminant validity suggested that all the confidence intervals between the scales of the PSDQ-S, except between sports competence and endurance, were less than 1.00. Therefore, it was deeded discriminant validity of the PSDQ-S was generally supported. However, the results showed convergent validity was not supported as AVEs of most of the PDSQ-S scales did not meet the cut-off, indicating more than a half of its total variance was derived from measurement errors (Fornell & Larcker, 1981).

Marsh et al. (2010) has shown that PDSQ-S has strong construct validity in terms of convergent validity, discriminant validity, and reliability. They also support the invariance measurement models across different samples. This study uses a different approach to the test of construct validity by calculating the CI of the latent factor correlations, AVE, and rho, and provides support to discriminant validity and reliability of the PDSQ-S. However, there seems to be some issues with convergent validity using the AVE approach. As this is the first study to examine the AVE of the PDSQ-S, it suggests that replication studies are needed to

further examine convergent validity of the PDSQ-S. It should be noted that the multitrait-multimethod used by Marsh et al. essentially uses the correlation coefficients of different traits and a large sample size. The judgement nature of the MTMM means that different researchers could arrive at different conclusions (Cole, 1987).

The findings of the present study show that there are at least three groups of students with unique physical self-description profiles. Two of the three clusters consist of more females than males (Groups 2 and 3). Group 2 has about 85% of females with a generally low self-description profile; they also had the lowest physical activity participation, compared to the other two clusters. Group 3 is also a female-dominated group (69% females) with moderate levels of physical self-concept. On the other hand, Group 1 is has more males than females. The typical characteristic of this group is high in sports competence, strength, flexibility, endurance, coordination, appearance, physical activity, general self-concept and self-esteem. Physical activity participation is the highest among this group. The findings of the current study show that the PDSQ-S provides a clear analysis of the unique characteristics of the participants. The LPA shows a clear distinction in all the 11 factors of the PDSQ-S.

The current study provides evidence that PDSQ-S has adequate internal reliability and discriminant validity. However, convergent validity of the PDSQ-S is questionable and there is a need for more studies to investigate the psychometric properties of the PDSQ-S, particularly with non-English speaking nations.

What is clear is that the 11 factors of the PDSQ-S provide three clear distinctive profiles that possess unique characteristics and that the profiles could be differentiated by the levels of physical activity level. Future studies could use the method LPA to design intervention programmes suitable for each profile group (Donovan & Owen, 1994; Killoran, Cavill, & Walker, 1994). There are a few limitations of the current study. First, this study is cross- sectional, it is not possible to provide information on trends. A longitudinal design

would be required. Second, the participants may respond to the questionnaire in a socially desirable way. Finally, this study lacks objective measures of sedentary and physical activities. Future studies need to consider other more objective measures such as the use of direct observations or pedometers. Another possibility is to consider garnering information from other sources, such as from parents, peers, and teachers to triangulate the findings.

In conclusion, the use of LPA in this study shows that it is useful in differentiate groups of students with different physical self-concept. This will help in studying the defining characteristics of important subgroups and consequently intervention programmes can be designed to better target such groups. Given the multidimensional nature of physical self-concept, studying such factors in combination may be fruitful.