# FACTORIAL VALIDITY AND INVARIANCE OF THE PHYSICAL SELF-DESCRIPTION QUESTIONNAIRE AMONG BLACK AND WHITE ADOLESCENT GIRLS

Rod K. Dishman, PhD; Derek P. Hales, PhD; Maria Joao Almeida, PhD; Karin A. Pfeiffer, PhD; Marsha Dowda, DrPH; Russell R. Pate, PhD

Meaningful comparison of physical self-concept among racial or ethnic groups requires that the measurement instruments used have equivalent measurement properties. We tested the factorial validity and invariance of the Physical Self-Description Questionnaire (PSDQ) among Black ($n$=658) and White ($n$=479) adolescent girls in the 12th grade. Construct validity was examined by estimating correlations between PSDQ subscales and external criteria (physical activity, physical fitness, body mass index [BMI], and participation in sports). The hypothesized 11-factor model demonstrated adequate overall fit in both groups. Comparison of nested models supported the between-group invariance of the overall factor structure. Convergent and discriminant evidence for construct validity was supported by the pattern of correlations with the external criteria. The results indicate that a meaningful comparison of PSDQ scores can be made between Black and White girls in the 12th grade and that valid inferences from PSDQ scores can be made about specific aspects of physical self-concept. Despite lower levels of physical activity, sport participation, and fitness and higher BMI, Black girls had similar self-esteem and higher physical self-concept and perceived appearance compared to White girls. (*Ethn Dis*. 2006;16:551–558)

**Key Words:** African American, Confirmatory Factor Analysis, Measurement, Physical Self-Concept, Psychometrics, Self-Esteem

From the Department of Kinesiology, University of Georgia, Athens, Georgia (RKD, DPH); Department of Exercise Science, University of South Carolina, Columbia, South Carolina (MJA, KAP, MD, RRP).

Address correspondence and reprint requests to Rod K. Dishman, PhD; Department of Kinesiology; The University of Georgia; Ramsey Student Center, 300 River Road; Athens, Georgia 30602-6554; 706-542-9840; 706-542-3148 (fax); rdishman@uga.edu

## INTRODUCTION

As youths move through adolescence, their participation in physical activity declines markedly.[1,2] The declines are greater in girls, especially Black girls. From the ages of 9 to 19, the median level of leisure time physical activity for girls in the United States declines by ≈80%.[2] Among high school students in the United States, 45% of Black girls and 58% of White girls participate in sufficient vigorous physical activity, compared to 65% of Black boys and 72% of White boys.[3] Low levels of physical activity during adolescence are believed to contribute to increased incidence of obesity and type 2 diabetes[4,5] and increased risk of sedentary adult behavior.[6] This public health burden makes it important to determine what contributes to decreased physical activity among girls during late adolescence.

*From the ages of 9 to 19, the median level of leisure time physical activity for girls in the United States declines by ≈80%.[2]*

Positive physical self-concept is a putative mediator of participation in sport and physical activity[7,8] that has been understudied among adolescent girls,[9] despite evidence from studies of young athletes that females have lower physical self-concept than males.[10] Physical self-concept also mediates the positive relations of physical activity and sport participation with self-esteem and its inverse association with depression symptoms among Black and White adolescent girls.[11]

A widely used and validated[13–15] multidimensional measure of physical self-concept is the Physical Self-Description Questionnaire (PSDQ),[12] Measurement invariance of the PSDQ has been supported across sexes, over time, and between cultures,[14–17] but few studies have examined validity in US samples.[18,19] To our knowledge, no studies have tested the factorial validity and invariance of the PSDQ in a sample of Black and White adolescent girls in the United States. If the PSDQ is to be compared as an influence or outcome of physical activity between Black and White girls, we must first establish the factorial validity and invariance of the PSDQ for the two groups. Factorial validity is the extent a set of items, representing a construct, conforms to the theoretical definition of that construct.[20,21] Factorial invariance is the extent to which a model generalizes across points of time or between groups of people. Tests of factorial invariance explore the comparability of the form and values of parameters within a measurement model.[22]

The purpose of the current investigation was, therefore, to test the factorial validity and multigroup invariance of responses to the PSDQ in Black and White adolescent girls. In addition, convergent and discriminant evidence for construct validity of the PSDQ was examined by estimating factor correlations between PSDQ subscales and physical activity, physical fitness, body mass index (BMI), and sport participation.

# METHODS

## Participants

Participants were 658 Black girls and 479 White girls in the 12th grade from 22 public high schools in South Carolina. The two groups did not differ in age (17.7 years, standard deviation [SD] .61), but Black girls had higher estimated fatness determined by BMI expressed as $kg/m^2$ (26.30 vs 23.69, $t_{1125}=6.70$, $P<.001$), lower fitness determined by physical work capacity ($PWC_{170}$) expressed as $kg^{-m}/min/kg$ (10.64 vs 12.18, $t_{936}=6.47$, $P<.001$), and lower participation in physical activity expressed as multiples of resting metabolic rate (METs) accumulated across three days (61.03 vs 63.38, $t_{1036}=3.17$, $P<.002$) and number of sports (.90 vs 1.23, $t_{1132}=4.36$, $P<.001$). Parental education levels of Black girls compared with White girls were lower for mothers ($\chi^2_{(5)}=58.14$, $P<.001$) and fathers ($\chi^2_{(5)}=110.8$, $P<.001$). A high school diploma was the highest educational degree for 41% of the mothers and 44% of the fathers of Black girls compared to 28% each for the mothers and fathers of White girls. Fourteen percent of fathers and 11% of mothers of Black girls had not completed high school compared to 10% of fathers and 6% of mothers of White girls. Thirty-nine percent of the fathers and 35% of the mothers of White girls had a college degree or post-baccalaureate education compared to 11% of the fathers and 15% of the mothers of Black girls.

## Measures

The Physical Self-Description Questionnaire (PSDQ) consists of 70 statements rated on a six-point true-false scale. It measures nine scales specific to physical self-concept (ie, strength, body fat, physical activity, endurance/fitness, sports competence, coordination, health, appearance, and flexibility) and includes scales measuring global physical self-concept and self-esteem. Each scale includes six to eight items that are paired and averaged to form item parcels (three or four indicators per scale). The use of item parcels has been recommended based on psychometric and structural modeling advantages over the uses of single-item indicators.[15,23]

Four external criteria (physical activity, physical fitness, BMI, and sport participation) were selected for comparison based on theoretical similarities to specific PSDQ subscales. Physical activity was measured by using the 3-Day Physical Activity Recall (3DPAR) described elsewhere.[24] The 3DPAR provides a reliable estimate of usual physical activity by assessing multiple days (two weekdays and one weekend day) of physical activity in a single reporting session. Responses are converted into 30-minute blocks of total METs (ie, physical activity level expressed as multiples of basal metabolic rate) for each day. Factorial invariance of the 3DPAR has been established between Black and White adolescent girls across 12 months.[25] Criterion-related validity of the 3DPAR for use with adolescent girls has been established based on the relation between the 3DPAR and an objective measure of physical activity derived by accelerometry.[24] The 3DPAR was hypothesized to be most closely related to the physical activity subscale of the PSDQ, with significant but lower correlations with strength, endurance, and sport competence. Physical fitness was measured with the physical work capacity 170 ($PWC_{170}$). This test requires participants to cycle at three submaximal rates of power output ($kg^{-m}/min$). Heart rate measured at each stage is used to estimate power output at a heart rate of 170 beats per minute. The $PWC_{170}$ was expressed per kg ($kg^{-m}/min/kg$) and was expected to correlate highest with the endurance scale, with significant but lower correlations with the physical activity and sport subscales of the PSDQ. Height and weight were measured and used to calculate BMI ($kg/m^2$), a common measure of excess weight for height. Body mass index (BMI) was used as an estimate of fatness and was therefore thought to relate most closely with the body fat and appearance subscales. The final criterion was sport participation, which was calculated as the sum of two items adapted from the Youth Risk Behavior Survey[26]: "During the past 12 months, how many sports teams run by your school did you play on? (Do not include PE classes.)" and "During the past 12 months, how many sports teams run by organizations outside your school did you play on?" Scores for each item range from none to three. Sport participation was expected to correlate most strongly with the sports competence subscale and to a lesser extent with the physical activity, endurance, and strength subscales. Each criterion (3DPAR, $PWC_{170}$, BMI, and sport participation) was modeled as a latent variable. The 3DPAR was indicated by the sum of three single-day estimates of total METs from the 3DPAR. Each of the other three variables was modeled as a single-item indicator of a latent construct with a loading of 1.0 and error variance set based on reliabilities for each measure.

## Procedures

Data for the 3DPAR were collected on a Wednesday, and $PWC_{170}$, BMI, sport participation, and PSDQ were usually collected within one week of the 3DPAR. The University of South Carolina Institutional Review Board approved all procedures. Participants and parents or legal guardians of girls <18 years of age provided written informed consent. The PSDQ was administered by trained data collectors to groups of 6 to 30 girls.

## Data Analysis

All models were tested with full-information maximum likelihood (FIML) estimation in AMOS 5.0 (SmallWaters Corp., Chicago, Ill).[27] Results were also obtained from LISREL 8.51

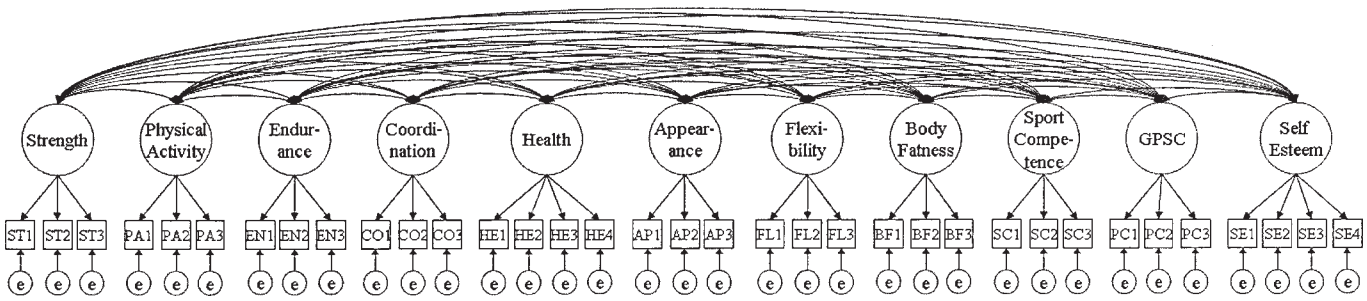


**Fig 1. Measurement model for tests of the factorial validity and multigroup invariance of the Physical Self-Description Scale in Black and White girls. Notes: GPSC=global physical self-concept; e=item error or uniqueness**

(Scientific Software International, Inc., Chicago, Ill).[28] Preliminary analyses, creation of item parcels, and comparison of observed scores were conducted by using SPSS 13.0 (SPSS, Inc., Chicago, Ill) and PRELIS.[29]

Multigroup factorial invariance was established by testing and comparing a series of models with standard procedures.[22,30,31] The measurement model depicted in Figure 1 was first tested separately for Black and White girls. The metric was established by setting the loading for the first indicator of each latent factor to 1.0. Then we tested the invariance of the overall factor structure (model 1), factor loadings (model 2), factor variances (model 3), factor covariances (model 4), item uniquenesses (model 5), and item intercepts (model 6).

The equivalence of factor structure (model 1) and loadings (model 2) was selected as the minimal standard for concluding measurement invariance across groups. The test of item intercepts (model 6) is not essential, but it examines the similarity in item response pattern between races. Support for model 6 would allow for stronger conclusions from latent means analysis of the PSDQ factors in which factor loadings and item intercepts were held invariant between groups, while the latent means for each PSDQ factor were estimated for White girls and fixed to zero for Black girls. The estimated latent mean difference for each factor was then standardized by dividing the mean difference by the pooled standard deviation for that latent factor.

In addition to the multigroup invariance analysis, two models (model 7 and model 8) that included each of the external criteria were tested. Based on similarities in fit for models 1 and 2, factor loadings were held invariant in model 7, while covariance between all PSDQ scales, 3DPAR, $PWC_{170}$, BMI, and sport participation were estimated freely. In model 8, factor loadings, variances, and covariance among PSDQ subscales and all external criteria were held invariant between Black and White girls. Comparison of models 7 and 8 allowed for an invariance test of the covariance among PSDQ subscales and the external criteria between White and Black girls.

The root mean square error of approximation (RMSEA), non-normed fit index (NNFI), comparative fit index (CFI), and the chi-square statistic were used to evaluate and compare model fit. The chi-square statistic assesses absolute fit of the model to the data, but it is sensitive to sample size.[20,28,32] The RMSEA is a standardized estimation that represents closeness of fit of population data to the model. It is widely used and is considered one of the most informative fit criteria.[33] Values approximating .06 to .08 reflect close and reasonable fit of the model.[34,35] The 90% confidence interval (CI) around the RMSEA point estimate should contain .06 or zero to indicate the possibility of close or exact fit. The RMSEA point estimate and 90% CIs in the multigroup analyses were multiplied by the $\sqrt{2}$ as suggested by Steiger.[36] This correction adjusts for the downward bias in RMSEA values estimated in multigroup analyses. The CFI and NNFI test the proportionate improvement in fit by comparing the target model to some baseline model.[35,37] The NNFI, but not the CFI, is affected by model parsimony (more complex models are penalized). Values for the CFI and NNFI approximating .95 indicate good fit.[35] Acceptable model fit was based on CFI and NNFI values of .90.[35,37] Nested models were compared based on $\chi^2$ difference tests and changes in the values of the RMSEA, CFI, and NNFI. Parameter estimates, standard errors, *t* values, and squared multiple correlations were inspected for sign and magnitude.

## RESULTS

### General Descriptives

Scale means, standard deviations, and alpha reliabilities are shown in Table 1. In addition, univariate ANOVA *P* values and effect sizes comparing scale scores for Black and White girls are presented. Scale comparisons between groups were adjusted for potential confounding effects of parental education by including mother and father education as covariates in the

**Table 1. Descriptive statistics (M, SD) and internal consistency reliability for Physical Self-Description Questionnaire (PSDQ) scores among Black and White girls in the 12th grade**

| | Black girls ($n$=658) | | | | White Girls ($n$=479) | | | | Total ($N$=1137) | | Group Comparison | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Scale | M | SD | α* | 95%CI | M | SD | α* | 95%CI | M | SD | *P* value† | g‡ |
| Strength | 4.17 | 1.03 | .77 | .74,.80 | 4.04 | 1.07 | .89 | .87,.90 | 4.11 | 1.05 | .038 | .12 |
| Physical activity | 3.44 | 1.43 | .87 | .85,.88 | 4.02 | 1.47 | .93 | .91,.94 | 3.68 | 1.47 | <.001 | −.40 |
| Endurance | 3.10 | 1.23 | .85 | .82,.87 | 3.38 | 1.33 | .91 | .90,.92 | 3.22 | 1.28 | .013 | −.22 |
| Coordination | 4.31 | 1.03 | .84 | .82,.86 | 4.39 | 1.02 | .90 | .89,.92 | 4.34 | 1.03 | .968 | −.08 |
| Health | 4.47 | .98 | .79 | .76,.81 | 4.44 | .96 | .87 | .85,.89 | 4.46 | .97 | .197 | .03 |
| Appearance | 5.07 | .90 | .83 | .80,.85 | 4.33 | .97 | .90 | .88,.92 | 4.76 | 1.00 | <.001 | .80 |
| Flexibility | 4.23 | 1.15 | .83 | .80,.85 | 4.45 | 1.09 | .89 | .88,.91 | 4.33 | 1.13 | .019 | −.20 |
| GPSC | 4.45 | 1.25 | .89 | .88,.91 | 4.06 | 1.16 | .92 | .90,.93 | 4.29 | 1.23 | <.001 | .33 |
| Body fat | 4.24 | 1.61 | .92 | .91,.93 | 4.02 | 1.54 | .95 | .95,.96 | 4.15 | 1.58 | .002 | .14 |
| Sport competence | 3.69 | 1.35 | .92 | .91,.93 | 3.83 | 1.24 | .95 | .94,.95 | 3.75 | 1.31 | .378 | −.11 |
| Self-esteem | 5.10 | .83 | .77 | .74,.80 | 5.06 | .74 | .82 | .80,.85 | 5.08 | .79 | .093 | .04 |

* α=Cronbach α reliability coefficient for item parcels.
† *P* values are for independent samples *t* tests corrected by Bonferroni adjustment.
‡ g=effect size calculated as (Black-White)/pooled SD.
GPSC=Global Physical Self-Concept; M=median; SD=standard deviation.

ANOVA. Effect sizes were calculated as [g=($Mean_{Black}$−$Mean_{White}$)/Pooled SD];[38] values of .2, .5, and .8 were considered small, medium, and large effects.[39] Mardia test of multivariate kurtosis was significant ($P<.001$), but the relative kurtosis index (1.297) and univariate kurtosis values indicate that violations of multivariate normality should have minimal effect on model estimates[29,40] Only three item parcels had kurtosis values >2.0, AP1 (2.11) and SE4 (3.22) for Black girls and SE4 (4.63) for White girls.

## Multi-Group Factorial Invariance

The results of the multigroup invariance analysis are presented in Table 2. The RMSEA (90% CI), CFI, and NNFI indicate adequate fit of the 11-factor model to the data in both groups (Black girls CFI=.92, NNFI=.91, RMSEA=.06; White girls CFI=.93, NNFI=.92, RMSEA=.07). The test of equal factor structures (model M1) was not rejected based on acceptable (CFI=.93, NNFI=.91) and close (RMSEA=.062) model-data fit. The invariance of factor loadings is

**Table 2. Results of the confirmatory factor analyses testing the multigroup factorial invariance of the PSDQ measurement model between Black and White girls in the 12th grade**

| Model | *df* | $\chi^2$ | RMSEA (90% CI) | CFI | NNFI |
|---|---|---|---|---|---|
| Black girls | 505 | 1654.608 | .059 (.056–.062) | .923 | .909 |
| White girls | 505 | 1567.016 | .066 (.063–.070) | .930 | .918 |
| Invariance | | | | | |
| Factor structure (M1) | 1010 | 3221.744 | .062 (.059–.065) | .927 | .909 |
| Factor loadings (M2) | 1034 | 3295.996 | .062 (.059–.065) | .925 | .909 |
| Variance (M3) | 1045 | 3328.948 | .062 (.059–.065) | .924 | .909 |
| Covariance (M4) | 1100 | 3495.353 | .062 (.059–.064) | .921 | .909 |
| Uniquenesses (M5) | 1135 | 4480.576 | .072 (.069–.074) | .889 | .877 |
| Intercepts (M6) | 1170 | 4988.543 | .076 (.074–.078) | .873 | .863 |
| External criteria | | | | | |
| External (M7) | 1380 | 3776.708 | .055 (.054–.058) | .927 | .909 |
| Covariance (M8) | 1485 | 4064.013 | .055 (.054–.058) | .922 | .909 |

| Model Comparisons | $df_{diff}$ | $\chi^2_{diff}$ | *P* value |
|---|---|---|---|
| M1 vs M2 | 24 | 74.252 | $P<.001$ |
| M2 vs M3 | 11 | 32.952 | $P<.001$ |
| M3 vs M4 | 55 | 166.405 | $P<.001$ |
| M4 vs M5 | 35 | 985.223 | $P<.001$ |
| M5 vs M6 | 35 | 507.967 | $P<.001$ |
| M7 vs M8 | 105 | 287.305 | $P<.001$ |

PSDQ=Physical Self-Descriptive Questionnaire; RMSEA=root mean square error of approximation; CFI=comparative fit index; NNFI=non-normed fit index.

**Table 3. Factor correlations of PSDQ subscales and external criteria (from model 7) among Black and White girls in the 12th grade**

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Strength | | .582 | .589 | .523 | .693 | .739 | .145 | .518 | .391 | .627 | .281 | .317 | .271 | *.019* | .422 |
| 2. Self-Esteem | .526 | | .755 | .590 | .513 | .493 | .463 | .412 | .729 | .665 | .376 | .143 | .259 | −.188 | .292 |
| 3. GPSC | .690 | .621 | | .481 | .647 | .555 | .611 | .517 | .639 | .609 | .270 | .204 | .349 | −.339 | .398 |
| 4. Flexibility | .731 | .529 | .606 | | .564 | .432 | .307 | .390 | .393 | .751 | .260 | .189 | .217 | −.192 | .276 |
| 5. Endurance | .647 | .265 | .520 | .675 | | .724 | .325 | .739 | .328 | .682 | .266 | .392 | .457 | −.250 | .554 |
| 6. Sports Competence | .729 | .346 | .573 | .621 | .760 | | .158 | .598 | .251 | .694 | .281 | .333 | .335 | −.155 | .602 |
| 7. Body Fat | .204 | .441 | .579 | .399 | .320 | .289 | | .127 | .543 | .287 | .155 | *−.042* | .298 | −.665 | .145 |
| 8. Physical Activity | .642 | .202 | .472 | .518 | .740 | .678 | .151 | | .162 | .602 | .190 | .503 | .378 | −.135 | .490 |
| 9. Appearance | .493 | .764 | .630 | .429 | .182 | .242 | .423 | .109 | | .371 | .146 | *.028* | .152 | −.199 | *.091* |
| 10. Coordination | .708 | .579 | .641 | .840 | .676 | .694 | .315 | .642 | .418 | | .318 | .290 | .316 | −.224 | .393 |
| 11. Health | .215 | .410 | .139 | .199 | .115 | .110 | .193 | *.027* | .240 | .196 | | *.058* | .230 | *−.045* | .229 |
| 12. 3DPAR | .225 | *.032* | .128 | .130 | .257 | .235 | *.069* | .370 | *−.007* | .157 | *.037* | | .155 | *.004* | .247 |
| 13. $PWC_{170}$ | .171 | .167 | .273 | .228 | .260 | .287 | .447 | .263 | .170 | .263 | *.079* | .231 | | −.405 | .274 |
| 14. BMI | *−.013* | −.239 | −.330 | −.268 | −.269 | −.206 | −.747 | −.088 | −.256 | −.274 | *−.083* | *−.028* | −.552 | | −.150 |
| 15. Sport Participation | .326 | .128 | .258 | .262 | .420 | .457 | .160 | .454 | .116 | .327 | *.070* | .291 | .261 | −.146 | |

Correlations above the diagonal (White girls), below the diagonal (Black girls).

Italicized correlations considered nonsignificant (critical *t* value <2).

PSDQ=Physical Self-Descriptive Questionnaire; GPSC=Global Physical Self-Concept; 3DPAR=3-Day Physical Activity Recall; $PWC_{170}$=physical work capacity 170; BMI=body mass index.

often considered the minimal condition for establishing factorial invariance across groups. The fit of model M2 is similar to model M1, indicating that when factor loadings are constrained, model fit does not suffer. Similarly, factor structure and loadings were similarly invariant for Black and White girls when items were used instead of parcels as indicators (fit for M2 using items as indicators: $\chi^2(4639)=12209.2$, CFI=.86, NNFI=.85, RMSEA [90% CI]=.052 [.051–.053]). This finding suggests that the conclusions concerning invariance between Black and White girls were not affected by the use of item parcels in the invariance analysis.

Subsequent analyses tested the invariance of factor variances, factor covariance, and item uniquenesses. Although the $\chi^2_{diff}$ tests are significant, the similarity in values for the RMSEA, CFI, and NNFI indicate that factor loadings (M2), variance (M3), and covariance (M4) were similar across groups. Item uniquenesses and intercepts were not invariant. Completely standardized invariant loadings for item parcels ranged from .55 to .93. All were significant (critical value >2.0). Squared multiple correlations ranged from .26 to .85 and from .39 to .89 for Black and White girls, respectively.

### External Criteria

The RMSEA (90% CI), CFI, and NNFI indicated adequate fit of the model including external criteria (model 7) to the data (CFI=.93, NNFI=.91, RMSEA=.055 [.054–.058]). Although the $\chi^2_{diff}$ test was significant, the similarity in values for the RMSEA, CFI, and NNFI for models 7 and 8 indicate that the covariance among PSDQ subscales and the external criteria were similar between groups. For comparison, factor correlations among PSDQ subscales and the external criteria are presented separately for Black and White girls in Table 3. As expected, factor correlations between the PSDQ subscales and their closest matching criteria (eg, BMI and the body fat subscale) were each significant. The largest correlations with sport participation, for both White and Black girls, were with the sports competence, endurance, and physical activity scales. Body mass index (BMI) was strongly related to the body fat scale for Black (−.75) and White (−.67) girls, with the next largest correlation (≈−.33) being with the global physical self-concept scale in both groups. The 3DPAR had the largest correlation with the physical activity scale and did not significantly correlate with the body fat or appearance scales in either group. The $PWC_{170}$ had the largest correlation with the endurance (.46) and body fat scales (.45) in White and Black girls, respectively; its correlation with the endurance scale (.26) was smaller but significant in Black girls.

## GROUP COMPARISON

The physical activity, endurance, and flexibility scale means were significantly lower for Black girls, while the strength, appearance, global physical self-concept, and body fat scale means were significantly lower ($P<.05$) for White girls. No group differences were found for the sports competence, coordination, health, or self-esteem subscales. Most effect sizes were small (median g=.14), but the higher scores among Black girls on the global physical self-concept (g=.32) and appearance (g=.80) subscales and lower score on the physical activity subscale (g=.40)

*Black girls in the present study had substantially higher physical appearance subscale scores than White girls. . .*

were moderate to large in size. The standardized latent mean differences were nearly identical to the effect sizes calculated from the observed scores. The pattern of significant and nonsignificant mean differences was the same for observed and latent mean analyses. For this reason, only the significance and effect sizes (g) from the observed score analyses are presented. As shown in Table 1, Cronbach alpha coefficients for PSDQ subscales ranged from .77 to .95; reliabilities were slightly higher for White (median $\alpha$=.91) compared to Black (median $\alpha$=.84) girls.

## DISCUSSION

Our results support the hypothesized measurement model, multigroup invariance, and construct validity of the PSDQ in a sample of Black and White girls in the 12th grade. To our knowledge, this is the first reported evidence indicating that the measurement structure of the PSDQ is equivalent between samples of Black and White adolescent girls. Model fit and factor correlations were similar to those reported for a sample of undergraduate, American females[18] but not quite as good compared to results for Australian adolescents.[15,41] Similar to past studies,[13–15] the internal consistency reliability of the 11 subscales approximated or exceeded .80, but scale reliabilities were slightly lower for responses by the Black girls. Means for the appearance and self-esteem scales were substantially larger in our sample compared to those previously reported.[15,41] Why differences exist in model fit or scale means is not apparent, but the similarities between results from samples of adolescent American females indicate that age, culture, and/or sex could explain the discrepancy.

Higher and more stable ratings of self-perceived physical appearance and global self-worth in Black, compared to White, girls in the National Heart Lung and Blood Institute (NHLBI) growth and health study[42] were interpreted as the result of cultural differences in acceptable body shape and size.[43] Similar to those findings, Black girls in the present study had substantially higher physical appearance subscale scores than White girls. Factor correlations between self-esteem, global physical self-concept, and physical appearance were, however, similar for Black and White girls (see Table 3).

The relations between PSDQ subscales and external criteria of BMI, 3DPAR, $PWC_{170}$, and sport participation indicate acceptable construct validity of the PSDQ. A number of the subscales correlated significantly with the external measures, but the largest correlations were noted between the criterion measure and its conceptually matched subscale in all but one case. In this instance, $PWC_{170}$ for Black girls was correlated with body fat (.45) more so than endurance (.26). Why this measure of physical fitness was more strongly related to body fat than endurance physical self-concept in Black but not White girls is unclear. As expected, sport participation was correlated most highly with the sport competence subscale in both groups, but it was similarly correlated with the endurance and physical activity subscales. In contrast, the health subscale was not significantly related to any of the external criteria.

Despite the fact that Black girls had higher BMI, their ratings of physical appearance and body fatness were more positive than White girls' ratings. This enhanced "body satisfaction" is consistent with previous findings[42] and suggests that cultural values influence perception and acceptance of body size and shape. Post hoc analysis revealed that Black girls had more favorable scores on the appearance and body fatness scales across the range of BMI scores. The importance, or effect, of these differences for the design of interventions to increase physical activity is less clear, however, given the lower levels of fitness and physical activity in the Black girls compared to the White girls. Further evidence from longitudinal cohort and intervention studies is needed to determine the natural course and effect of racial differences in physical self-concept and its relations with body fatness, fitness, physical activity participation, and self-esteem.

Although factor variances and covariances were found to be invariant between the groups, some evidence (albeit subjective) indicated that certain factor covariances differed between the groups. For example, the covariances of self-esteem with endurance and with physical activity were smaller for Black girls (.235 [95% CI .15–.32]; .180 [95% CI .10–.26]) compared to White girls (.484 [95% CI .38–.59]; .384 [95% CI .28–.49]). Because of the number of parameters estimated in the model, freeing or constraining these two factor correlations did not change model fit. Similar conclusions are supported for the invariance of the relationships among PSDQ subscales and the four external criteria. While comparison of models 7 and 8 suggests that constraining covariances does not substantially change model fit, the magnitude of several of the factor correlations between the PSDQ subscales and the external criteria seem to differ between Black and White girls.

Some researchers have recently suggested that certain items from the PSDQ be revised because of their limited measurement precision.[44] Negatively worded items were specifically cited as providing low information to the overall scale. In our sample, post hoc

examination revealed a large modification index for the error covariance between two item parcels from the self-esteem scale. These parcels, SE2 and SE4, are composed of negatively worded items. Removal of these parcels from model 2 used in this study improved fit (CFI=.937, NNFI=.922, RMSEA [90% CI]=.059 [.057–.062]).

Latent means analysis allows researchers to control error variance between items and constrain the underlying measurement structure of an instrument between groups. Strong conclusions about the latent mean comparisons between the Black and White girls were limited by the lack of invariant item intercepts. This finding indicates that the response pattern across item parcels was not the same for Black and White girls, consistent with their different internal consistency coefficients. Nonetheless, from a practical standpoint, the similarity of latent and observed mean score comparisons and measurement invariance of the PSDQ suggest that analysis of observed scores in Black and White adolescent girls will result in similar conclusions as latent covariance modeling.

Physical self-concept is thought to develop and change based on a series of personal experiences and judgments. Although cause and effect cannot be established from previous and current cross-sectional evidence, support exists for distinct but related categories of physical self-concept that correlate with behavior and/or physiologic ability.[14,15,19,45] The strength and consistency of the cross-sectional evidence lay the ground work for future studies designed to examine variation in physical self-concept as it relates to lifestyle or physiological change (ie, increased physical activity, weight loss, increased fitness).

The results of this cross-sectional investigation support the hypothesized factor structure and between-group invariance of responses to the PSDQ, indicating that the scales have an equivalent measurement structure in samples of Black and White girls in the 12th grade. In addition, our finding of coherent relations of the scales with external measures of fitness, fatness, physical activity and sport participation adds to the evidence that supports the construct validity of scores on the PSDQ for the assessment of physical self-concept among Black and White American girls during late adolescence.

## ACKNOWLEDGMENT

Supported by NIH HL 57775 National Heart, Lung, and Blood Institute.

## REFERENCES

1. Caspersen CJ, Pereira MA, Curran KM. Changes in physical activity patterns in the United States, by sex and cross-sectional age. *Med Sci Sports Exerc*. 2000;32:1601–1609.
2. Kimm SYS, Glynn NW, Kriska AM, et al. Decline in physical activity in Black girls and White girls during adolescence. *N Engl J Med*. 2002;347:709–715.
3. Grunbaum JA, Kann L, Kinchen SA, et al. Youth Risk Behavior Surveillance—United States, 2003. *MMWR Surveill Summ*. 2004;53(SS-2):21–24.
4. Ogden CL, Flegal KM, Carroll MD, Johnson CL. Prevalence and trends in overweight among US children and adolescents, 1999–2000. *JAMA*. 2002;288:1728–1732.
5. Wang G, Dietz WH. Economic burden of obesity in youths aged 6 to 17 years: 1979–1999. *Pediatrics*. 2002;109(5):E81, 1–6.
6. Campbell PT, Katzmarzyk PT, Malina RM, Rao DC, Perusse L, Bouchard C. Prediction of physical activity and physical work capacity (PWC150) in young adulthood from childhood and adolescence with consideration of parental measures. *Am J Hum Biol*. 2001; 13(2):190–196.
7. Sonstroem RJ. Physical self-concept: assessment and external validity. *Exerc Sport Sci Rev*. 1998;26:133–164.
8. Sonstroem RJ, Morgan WP. Exercise and self-esteem: rationale and model. *Med Sci Sports Exerc*. 1989;21(3):329–337.
9. Strong WB, Malina RM, Blimkie CJR, et al. Evidence based physical activity for school-age youth. *J Pediatr*. 2005;146:732–737.
10. Lirgg CD. Gender differences in self-confidence in physical activity: a meta-analysis of recent studies. *J Sport Exerc Psychol*. 1991; 13:294–310.
11. Dishman RK, Hales DA, Pfeiffer KA, et al. Physical self-concept and self-esteem mediate cross sectional associations of physical activity and sport participation with depression symptoms among adolescent girls. *Health Psychol*. In press.
12. Marsh HW, Richards GE, Johnson S, Roche L, Treymayne P. Physical self-description questionnaire: psychometric properties and a multimethod analysis of relations to existing instruments. *J Sport Exerc Psychol*. 1994b;16: 270–305.
13. Marsh HW, Asci FH, Tomás IM. Multitrait-multimethod analyses of two physical self-concept instruments: a cross-cultural perspective. *J Sport Exerc Psychol*. 2002a;24:99–119.
14. Marsh HW. Physical self-description questionnaire: stability and discriminant validity. *Res Q Exerc Sport*. 1996;67:249–264.
15. Marsh HW, Redmayne RS. A multidimensional physical self-concept and its relations to multiple components of physical-fitness. *J Sport Exerc Psychol*. 1994;16:43–55.
16. Marsh HW, Marco IT, Asçi FH. Cross-cultural validity of the physical self-description questionnaire: comparison of factor structures in Australia, Spain, and Turkey. *Res Q Exerc Sport*. 2002b;73:257–270.
17. Marsh HW. Age and gender effects in physical self-concept for adolescent elite athletes and nonathletes: a multicohort-multioccasion design. *J Sport Exerc Psychol*. 1998;20:237–259.
18. Nigg CR, Norman GJ, Rossi JS, Benisovich SV. Examining the structure of physical self-description using an American University sample. *Res Q Exerc Sport*. 2001;72:78–83.
19. Dunton GF, Jamner MS, Cooper DM. Physical self-concept in adolescent girls: behavioral and physiological correlates. *Res Q Exerc Sport*. 2003;74:360–365.
20. Bollen KA. *Structural Equations With Latent Variables*. New York, NY: John Wiley & Sons, Inc; 1989.
21. Messick S. Validity of psychological assessment – validation of inferences from persons' responses and performances as scientific meaning into score meaning. *Am Psychol*. 1995; 50:741–749.
22. Vandenberg R, Lance CE. A review and synthesis of the management invariance literature: suggestions, practices, and recommendations for organizational research. *Organ Res Methods*. 2000;3:4–69.
23. Little TD, Cunningham WA, Shahar G, Widaman KF. To parcel or not to parcel: exploring the question, weighing the merits. *Struct Equat Model*. 2002;9:151–173.
24. Pate RR, Ross R, Dowda M, Trost SG, Sirard JR. Validation of a 3-day physical activity recall instrument in female youth. *Pediatr Exerc Sci*. 2003;15:257–265.
25. Motl RW, Dishman RK, Dowda M, Pate RR. Factorial validity and invariance of a self-report measure of physical activity among adolescent girls. *Res Q Exerc Sport*. 2004;75: 259–271.

26. Centers for Disease Control and Prevention. Methodology of the Youth Risk Behavior Surveillance System. *Morb Mortal Wkly Rep*. 2004;53(No. RR-12):4–6.
27. Arbuckle JL, Wothke W. *Amos 4.0 User's Guide*. Chicago, Ill: Small Waters Corporation; 1999.
28. Jöreskog KG, Sörbom D. *LISREL® 8: User's Reference Guide*. Chicago, Ill: Scientific Software International, Inc; 1996a.
29. Joreskog K, Sorbom D. *PRELIS 2: Users Reference Guide*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1996b.
30. Horn JL, McArdle JJ. A practical and theoretical guide to measurement invariance in aging research. *Exp Aging Res*. 1992;18: 117–144.
31. Marsh HW. Confirmatory factor analysis models of factorial invariance: a multifaceted approach. *Struct Equat Mod*. 1994a;1:5–34.
32. Jöreskog KG. Testing structural equation models. In: Bollen KA, Long JS, eds. *Testing Structural Equation Models*. Newbury Park, Calif: Sage Publications, Inc; 1993;294–316.
33. Byrne BM. *Structural Equation Modeling with LISREL, PRELIS, and SIMPLIS: Basic Concepts, Applications, and Programming*. Mahwah, NJ: Lawrence Erlbaum Associates; 1998.
34. Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS, eds. *Testing Structure Equation Models*. Newbury Park, Calif: Sage Publications, Inc; 1993; 136–162.
35. Hu L, Bentler PM. Cutoff criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equat Model*. 1999;6:1–55.
36. Steiger JH. A note on multiple sample extensions of the RMSEA fit index. *Struct Equat Model*. 1998;5:411–419.
37. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. *Psychol Bull*. 1980;88: 588–606.
38. Hedges LV, Olkin I. *Statistical Methods for Meta-Analysis*. San Diego, Calif: Academic Press; 1985.
39. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum; 1988.
40. Kline RB. *Principles and Practices of Structural Equation Modeling*. New York, NY: The Guilford Press; 1998.
41. Marsh HW, Hey J, Roche LA, Perry C. Structure of physical self-concept: elite athletes and physical education students. *J Educ Psychol*. 1997;89:369–380.
42. Brown KM, McMahon RP, Biro FM, et al. Changes in self-esteem in Black and White girls between the ages of 9 and 14 years. The NHLBI Growth and Health Study. *J Adolesc Health*. 1998;23(1):7–19.
43. Kumanyika S, Wilson JF, Guilford-Davenport M. Weight-related attitudes and behaviors of Black women. *J Am Diet Assoc*. 1993;93: 416–422.
44. Fletcher RB, Hattie JA. An examination of the psychometric properties of the physical self-description questionnaire using a polytomous item response model. *Psychol Sport Exerc*. 2004;5:423–447.
45. Crocker PRE, Eklund RC, Kowalski KC. Children's physical activity and physical self-perceptions. *J Sports Sci*. 2000;18:383–394.

## AUTHOR CONTRIBUTIONS

*Design concept of study:* Dishman, Almeida, Pfeiffer, Dowda, Pate
*Acquisition of data:* Almeida, Pfeiffer, Pate
*Data analysis interpretation:* Dishman, Hales, Pate
*Manuscript draft:* Dishman, Hales, Pfeiffer, Dowda
*Statistical expertise:* Hales
*Acquisition of funding:* Pate
*Administrative, technical, or material assistance:* Almeida, Pfeiffer, Dowda, Pate
*Supervision:* Dishman, Pate