

Perceived Physical Appearance: Assessing Measurement Equivalence in Black, Latino, and White Adolescents

Anna E. Epperson,¹ MA, Sarah Depaoli,¹ PhD, Anna V. Song,¹ PhD, Jan L. Wallander,¹ PhD, Marc N. Elliott,² PhD, Paula Cuccaro,³ PhD, Susan Tortolero Emery,³ PhD, and Mark Schuster,^{4,5} MD, PhD

¹University of California, Merced, ²RAND Corporation, Santa Monica, California, ³University of Texas School of Public Health, and ⁴Boston Children's Hospital and ⁵Department of Pediatrics, Harvard Medical School

All correspondence concerning this article should be addressed to Anna E. Epperson, MA, Psychological Sciences, SSHA, University of California, Merced, CA 95343, USA. E-mail: aepperson@ucmerced.edu

Received September 17, 2015; revisions received April 20, 2016; accepted April 29, 2016

Abstract

Objective This aim of this study was to examine whether the construct of physical appearance perception differed among the three largest racial/ethnic groups in the United States using an adolescent sample. **Methods** Black (46%), Latino (31%), and White (23%) adolescents in Grade 10 from the Healthy Passages study ($N=4,005$) completed the Harter's Self-Perception Profile for Adolescents–Physical Appearance Scale (SPPA–PA) as a measure of physical appearance perception. **Results** Overall, Black adolescents had a more positive self-perception of their physical appearance than Latino and White adolescents. However, further analysis using measurement invariance testing revealed that the construct of physical appearance perception, as measured by SPPA–PA, was not comparable across the three racial/ethnic groups in both males and females. **Conclusions** These results suggest that observed differences may not reflect true differences in perceptions of physical appearance. Measures that are equivalent across racial/ethnic groups should be developed to ensure more precise measurement and understanding.

Key words: adolescent; obesity; race/ethnicity; statistical applications.

Negative perceptions about one's body and/or physical appearance are common among today's children and adolescents (Neumark-Sztainer, Paxton, Hannan, Haines, & Story, 2006). Research indicates that more than half of high-school-aged adolescents have very negative perceptions about their physical appearance and low body satisfaction (Neumark-Sztainer et al., 2006). These negative perceptions about physical appearance can be harmful for youth in a variety of ways. For example, negative body perceptions are associated with risk for poorer overall health, depression, low self-esteem, obesity, and increased dieting, including unhealthy weight control behaviors such as purging (Eisenberg, Neumark-Sztainer, Paxton, 2006). Previous research has also found that for obese

youth, negative perceptions about weight are associated with a lower likelihood of achieving a healthy weight in the future (Schuster et al., 2014).

However, these negative physical appearance perceptions may be more prevalent within certain groups of youth. Studies have reported that significant differences exist among adolescents of different gender and racial/ethnic groups in perceptions or concerns about physical appearance (George & Franko, 2010; Mikolajczyk, Iannotti, Farhat, & Thomas, 2012; Thompson, Rafiroiu, & Sargent, 2003). Adolescent females are more likely to report having negative perceptions about their physical appearance compared with adolescent males (Eisenberg et al., 2006). African-American/Black (Black) adolescents are more

likely to report a positive perception of their physical appearance and body image compared with non-Hispanic White (White) and Hispanic/Latino (Latino) adolescents, and this pattern holds for both males and females (de Guzman & Nishina, 2014; George & Franko, 2010). Findings are less consistent about whether differences exist for Latinos compared with Whites, with some studies finding no differences in physical appearance perceptions (de Guzman & Nishina, 2014; Erickson, Hahn-Smith, & Smith, 2009) and others indicating that Latinos may have a more positive perception of their physical appearance compared with Whites (Mikolajczyk et al., 2012).

For those who have negative perceptions about their physical appearance, there are many potential factors that may influence these perceptions. Current theories, such as Bronfenbrenner's Ecological Theory (Bronfenbrenner, 1989) and the Tripartite Influence Model (Keery, van den Berg, & Thompson, 2004), propose that both individual and environmental factors influence the development of these negative perceptions of appearance (Paxton, Schutz, Wertheim, & Muir, 1999). Individual-level factors identified in previous research as having a strong relationship with negative perceptions of physical appearance include overweight or obesity, depression, and low self-esteem (Paxton et al., 1999). Environmental factors that may influence these perceptions include family appearance and weight-related attitudes and behaviors, especially of parents, peer norms, and sociocultural body ideals (e.g., thin body ideal) portrayed in the media (Haines et al., 2008; Mikolajczyk et al., 2012). Further, it appears that these negative physical appearance perceptions begin very early, even as early as six years of age, and increase as youth transition from childhood to adolescence and then to young adulthood (Calzo et al., 2012).

The influence of these factors on adolescent perceptions of physical appearance may vary by race/ethnicity. For Black adolescents, cultural attitudes of acceptance of a larger body size may influence perceptions about physical appearance (de Guzman & Nishina, 2014; George & Franko, 2010). Familial factors may be important for Latino adolescents, where research has identified social class and acculturation, typically measured by generational status, as inversely associated with negative perceptions about physical appearance (Erickson et al., 2009; George & Franko, 2010).

Most pediatric health research examining differences among racial/ethnic groups has primarily analyzed straightforward bivariate associations, such as comparing mean differences among racial/ethnic groups. However, there are drawbacks to these approaches because they are based on the assumption that the underlying construct, in this case perception

of physical appearance, is manifested identically across different racial/ethnic groups. In theory, current measures of physical appearance perceptions are assumed to measure this concept equally for all individuals such that results should reveal only actual differences among individuals (Meredith, 1993; Millsap, 2012). A physical appearance perception measure developed for children and/or adolescents should measure how an adolescent actually perceives his/her physical appearance, both on a personal level and in relation to others. The measure should not produce different results for individuals who are identical on this attribute but who might differ on another variable such as membership in a particular racial/ethnic group. For example, a measure developed to assess anxiety should measure an individual's level of anxiety and not produce different scores for a male or female if they have the same anxiety level. Methods exist to evaluate whether a measure behaves differently or can be considered invariant across groups. To our knowledge, no studies have examined whether the construct of physical appearance perception is equal across racial/ethnic adolescent groups using measurement invariance testing. Using this approach, we can assess whether observed differences are a reflection of true differences on this attribute or whether they are due to another variable such as race/ethnicity.

The Harter Self-Perception Profile for Adolescents-Physical Appearance subscale is commonly used to assess adolescents' perceptions about their physical appearance both individually and in relation to others (Harter, 2012). The Self-Perception Profile (SPP) for Adolescents has been widely used to measure different aspects of self-perceptions in adolescents, with specific subscales for athletic competence, physical appearance, scholastic competence, social competence, and behavioral and global self-worth (Harter, 1998; Michaels, Barr, Roosa, & Knight, 2007; Schumann et al., 1999). The physical appearance subscale more specifically measures how good-looking adolescents perceive themselves to be and how happy they are with their appearance and body (Harter, 2012). As one of the most widely used measures of self-concept during adolescence, the SPP is reported to have high internal consistency and short-term test-retest reliability, and numerous studies have produced results that support its construct validity (Harter, 1998; Michaels et al., 2007; Schumann et al., 1999).

A few studies have examined whether the overall SPP or its subscales measure the same construct in a similar or comparable way across groups. Results from these studies indicated that overall self-perception measured by the SPP and the scholastic competence subscale was equivalent for clinical and nonclinical adolescent populations, females and males, and those of varying school grade level (fourth to sixth grade)

(Van den Bergh & Van Ranst, 1998; Veerman, ten Brink, Straathof, & Treffers, 1996). Only one study could be located that examined measurement invariance among racial/ethnic groups, but this study only focused on elementary and middle school children (Michaels et al., 2007). Results indicated that the overall SPP and its subscales were invariant across American Indians/Alaska Natives, Blacks, Latinos, and Whites. Given marked developmental changes that occur during puberty, with accompanying cognitive, emotional, and social changes, these findings may not necessarily generalize to older, high-school-aged adolescents. For example, youth who are late in pubertal development are more likely to have a more negative perception of their physical appearance (de Guzman & Nishina, 2014; Eisenberg et al., 2006) compared with those who initiate puberty earlier. Moreover, the previous study sample was predominately low-income, further restricting its generalizability (Michaels et al., 2007). Finally, this study used a modified version of the SPP for Children (SPPC), which comprises different question formatting and wording compared with the SPP for Adolescents (SPPA).

Given the high numbers of youth reporting negative perceptions about physical appearance and indications from previous research that there may be racial/ethnic differences, it is important to assess whether items from physical appearance measures function the same way across racial/ethnic groups. If they do not, results cannot be compared and conclusions may be invalid. One of the first steps to obtaining a clearer understanding of how perceptions of physical appearance influence health behaviors among adolescents therefore is to ensure that measurement outcomes are comparable across youth from varying racial/ethnic groups. Without this information, findings from research on perceived physical appearance and related interventions in diverse populations would be challenging to interpret, such as when findings are used to inform about other important health outcomes, for example, disordered eating (Kimber, Couturier, Georgiades, Wahoush, & Jack).

The current study applies measurement invariance testing to assess whether the construct of physical appearance perception, as measured by the subscale of the SPPA, is comparable across Black, Latino, and White adolescents, ages 15–16, and whether there are true racial/ethnic group differences in such perceptions. Because a large body of research indicates that physical appearance perceptions vary significantly among female and male adolescents, all analyses were conducted separately by gender (Neumark-Sztainer et al., 2006; Thompson et al., 2003).

Methods

Data for this project are from the Healthy Passages (HP) study, a longitudinal, multisite study of health

and health behaviors in youth (Windle et al., 2004). The data used here were from the final wave when participants were generally in Grade 10.

Participants

The sample included youth initially recruited and enrolled during the first wave of data collection when participants were in Grade 5 (M age = 11.12 years old). Participants were recruited from public schools with ≥ 25 students in regular academic classrooms in and around the metropolitan areas of Birmingham, Alabama, Houston, Texas, and Los Angeles, California. Schools and students were selected by using a two-stage probability sampling procedure where stratified sampling was used to ensure adequate sample sizes of non-Latino Black, Latino, and non-Latino White students (further details of the sampling procedure are provided elsewhere; Windle et al., 2004). Information about the study was distributed to each student in all Grade 5 classrooms in the selected schools for them to take home to their parents or primary caregivers (parent). Of the 11,532 children eligible for the study, 58% of the parents agreed to be contacted and receive further information about the study. Of these families, 77% completed the assessment during the first wave ($N = 5,147$). The sample closely resembled the sampled population and all eligible students on basic demographic characteristics, and sampling weights adjusted for any selection bias due to differential nonresponse. The first wave of data collection occurred from 2004 to 2006. After five (2009–2011) years, 4,521 families (88%) completed the final wave. Overall exclusion criteria included not attending a regular academic classroom or a parent who could not complete interviews in English or Spanish. Of the 4,371 youth who completed the SPPA–Physical Appearance Scale (SPPA–PA) at the final wave, only those identified as being members of one of the three major racial/ethnic groups (Black, Latino, and White) were included in the current analysis. This resulted in an analysis sample of 4,005 students (51% female), with a weighted distribution of 31% Black, 46% Latino, and 23% White (M age = 16.12 years). Highest level of household education was used as an indicator of socioeconomic status (SES) because previous research has indicated that it is a more stable predictor of SES for members of racial/ethnic minority groups (Kaufman, Cooper, & McGee, 1997; Williams & Collins, 1995). In the current sample, 44% of the sample was of low SES, as indicated by parent attainment of a high school degree or less. Each adolescent and parent who participated in the study received a gift certificate for their time. Further details on demographic information about this sample are provided elsewhere (Schuster et al., 2012).

Procedure

Following standard procedures approved by the institutional review boards at each of the data collection sites and the Centers for Disease Control and Prevention (CDC), two trained interviewers met the participants and their parent (87% were the biological mother) at their home or another agreed-upon location. Informed consent was provided by the parent and the adolescent provided assent. Computer-assisted self-interviews and interviewer-administered questions were then conducted in private spaces with the parent and adolescent separately to complete the full study protocol. Both the adolescent and his/her parent were given a choice of completing the interviews in English or Spanish. All adolescents and 83% of the parents chose English.

Measures

Race/Ethnicity

Race/ethnicity was classified based on the adolescent's response when asked which of seven racial/ethnic categories described him/her. The adolescent was classified as Latino, if so indicated, regardless of whether other categories had also been chosen. Adolescents not categorized as Latino were classified as (non-Latino) Black, (non-Latino) White, or other (which was excluded from the current analysis).

Self-Perception Profile for Adolescents–Physical Appearance Subscale

The SPPA-PA is a subscale used to measure adolescents' perceptions about their physical appearance (Harter, 2012), consisting of five items (Table 1) in which participants are asked to identify which contrasting description fit them best (e.g., "Some teenagers are not happy with the way they look, other teenagers are happy with the way they look"). Items were categorized as "way they look," "body different," "physical appearance," "good looking," and "like their looks." Participants were then asked to select how true the selected statement was for them ("sort of true" or "really true"). Each item is scored from 1 to 4 and certain item scores are reversed, such that the total score for the SPPA-PA ranges from 5 to 20, with higher scores indicating more positive physical appearance satisfaction. The SPPA was validated for use for ages 13-18 years, with a sample ($N=1,099$) that was approximately 90% White. Results indicated high internal consistency scores across all subscales, ranging from .87 to .96 (Harter, 2012). For the current sample, the internal consistency of the scale was $\alpha=0.66$. Previous research also indicates high construct validity, for example, by finding expected differences among normal weight, overweight, and obese youth (Schuster et al., 2012).

Table 1. Comparison of Racial/Ethnic Groups on Self-Perception Profile–Physical Appearance Subscale Score Means (SD)

SPPA-PA Item	Females					Males				
	Overall ($n=2,046$)	Latino ($n=759$)	Black ($n=796$)	White ($n=491$)	F-test η^2	Overall ($n=1,959$)	Latino ($n=725$)	Black ($n=697$)	White ($n=537$)	F-test η^2
SPPA-PA Scale Score	15.01 (3.92)	13.93 ^c (3.82)	16.37 ^a (3.64)	14.48 ^b (3.85)	88.04	16.11 (3.34)	15.30 ^c (3.36)	17.02 ^a (3.19)	16.04 ^b (3.21)	49.64
S1: Some teens are not happy with the way they look, other teens are happy with the way they look	3.07 (1.01)	2.89 ^b (1.02)	3.29 ^a (0.98)	3.02 ^b (0.97)	33.32	3.36 (0.87)	3.24 ^c (0.90)	3.49 ^a (0.87)	3.36 ^b (0.82)	14.00
S2: Some teens wish their body was different, other teens like their body the way it is	2.73 (1.14)	2.54 ^b (1.11)	2.97 ^a (1.15)	2.65 ^b (1.10)	29.41	3.10 (1.05)	2.98 ^b (1.06)	3.21 ^a (1.09)	3.13 ^a (0.98)	9.10
S3: Some teens wish their physical appearance was different, other teens like their physical appearance the way it is	2.99 (1.07)	2.83 ^b (2.99)	3.17 ^a (1.08)	2.94 ^b (1.03)	20.71	3.19 (0.99)	3.07 ^b (0.99)	3.32 ^a (1.02)	3.18 ^b (0.95)	11.48
S4: Some teens think that they are good looking, other teens think that they are not very good looking	3.13 (0.92)	2.83 ^b (0.92)	3.54 ^a (0.83)	2.93 ^b (0.81)	149.99	3.23 (0.86)	2.97 ^c (0.92)	3.57 ^a (0.79)	3.16 ^b (0.73)	95.10
S5: Some teens really like their looks, other teens wish they looked different	3.08 (0.99)	2.84 ^b (0.99)	3.40 ^a (0.95)	2.94 ^b (0.89)	73.37	3.22 (0.93)	3.04 ^c (0.95)	3.43 ^a (0.97)	3.21 ^b (0.81)	33.45

Note. SPPA-PA = Self-Perception Profile for Adolescents–Physical Appearance; teens = teenagers.

a,b,c Different superscripts within race/ethnicity subgroups for row variable indicate statistically significant difference as per ANOVA ($p < .05$).

Boldface values indicate significant results.

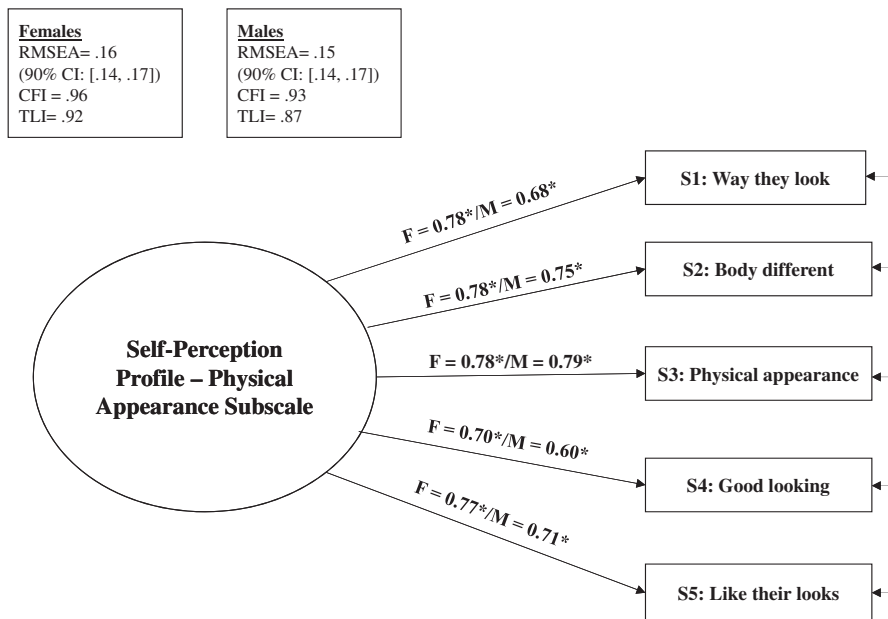


Figure 1. Confirmatory factor analysis model for 10th grade female and male high-school-aged adolescents (standardized estimates). RMSEA = root mean squared error of approximation; CFI = comparative fit index; TLI = Tucker–Lewis index; F = female; M = male. * $p < .05$.

Data Analysis

Descriptive analyses and tests for group differences were conducted using IBM SPSS Statistics™ Complex Sampling module with weighted data to adjust for the complex survey design and account for the effects of design and nonresponse weights, clustering of youth within schools, and stratification by site, as detailed elsewhere (Banis et al., 1988). Racial/ethnic groups are compared separately for females and males.

To conduct the measurement invariance test of the latent factor physical appearance perception using the SPPA–PA, all further analyses were conducted with *Mplus*, version 7.3 (Muthén & Muthén, 2012), while accounting for the complex survey design. Because all the items were categorical, models were estimated with weighted least-squares means and variance-adjusted (WLSMV) estimation and theta parameterization. We first verified that all five items from the SPPA–PA constituted one latent factor (physical appearance perception) by using confirmatory factor analysis (CFA) to test an overall model for the entire sample (Figure 1). Model fit was assessed using the comparative fit index (CFI), Tucker-Lewis index (TLI), and root mean square error of approximation (RMSEA) indexes. The literature recommends CFI and TLI values of $>.90$ to indicate good model fit, $.80$ – $.90$ to indicate acceptable fit, and $<.80$ to indicate poor fit; for RMSEA, values of $<.05$ indicate close fit, $.05$ – $.08$ indicate fair fit, and $>.10$ indicate poor fit (Browne & Cudeck, 1992; Hu & Bentler, 1999; Yu, 2002). It is recommended to use more than one measure of fit, especially when categorical data are present

(Hutchinson & Olmos, 1998). We expand on the interpretation of these measures in the Results section.

Following the overall CFA, as a preliminary step prior to measurement invariance testing, separate CFAs were conducted on SPPA–PA for Black, Latino, and White racial/ethnic groups to assess whether the measurement model differed. Next, invariance of the measurement model was tested among Black, Latino, and White adolescents using recommended steps (Millsap & Olivera-Aguilar, 2012). Invariance testing steps, where each subsequent step was completed pending confirmation of each previous step, were as follows: (1) configural or baseline invariance model where a single factor (physical appearance perception) was specified for the measured variables in each group; (2) metric invariance model where factor loadings are held equal across groups, also referred to as weak factorial invariance; (3) scalar invariance model where intercepts or thresholds are constrained to be equal across group, referred to as strong factorial invariance; (4) residual variances are constrained to be equal across groups, referred to as strict factorial invariance; and (5) factor variance is constrained to be equal across groups.

At each step in the process, model fit was tested using the chi-square difference test ($\Delta\chi^2$) and change in CFI (Δ CFI), as recommended by Cheung and Rensvold (2002), comparing the current step model with the previous step model. For Δ CFI, values that are smaller than or equal to -0.01 indicate invariance of the current model (Cheung & Rensvold, 2002). For the $\Delta\chi^2$, models were estimated with WLSMV using

the “difftest” option in *Mplus*. If invariance is rejected (i.e., if the χ^2 test is significant), then an attempt is made to locate the violation of invariance and the source of model misfit (i.e., by examining loading, intercepts/thresholds, etc.) using Lagrangian multipliers (modification indexes; Asparouhov & Muthén, 2009). If located, this parameter can then be “freed” across the groups and the model retested for potential achievement of partial invariance (Millsap & Olivera-Aguilar, 2012). This process will ultimately uncover whether a measurement model is invariant across groups. If differences are found in the measurement model (i.e., invariance is rejected), this may point to the differences in interpretation of the measure items and/or how physical appearance is conceptualized by individuals in particular racial/ethnic groups.

Results

Descriptive information appears in Table I. Overall, males had a significantly higher, approximately one-third standard deviation (*SD*) higher, SPPA-PA score compared with females ($t(4003) = -9.56, p < .01$). For both females and males, Blacks had a more positive perception of their physical appearance, especially Black females, compared with Whites and Latinos. Moreover, for both females and males, Whites reported more positive perceptions of their physical appearance than Latinos.

Measurement Model Testing

Figure 1 presents the results for the CFAs for the latent factor perception of physical appearance separately by females and males. For both sexes, all of the observed variables significantly loaded onto the latent factor. Model fit indexes for both the female and male models indicated a mixed pattern of fit. For the female model, CFI and TLI indicated good model fit, but RMSEA indicated poor or unacceptable fit (CFI = .96; TLI = .92; RMSEA = .16). For the male model, CFI indicated good fit and TLI indicated acceptable fit, but RMSEA again indicated poor fit (CFI = .93; TLI = .87; RMSEA = .15). Previous research has reported that CFI may be a more reliable indicator compared with RMSEA (Meade, Johnson, & Braddy, 2008; Yu, 2002). Further, previous work has indicated that the interpretation of RMSEA should be avoided when categorical indicators are used because this fit measure is unreliable under cases of model misspecification, which is unavoidable during the measurement invariance testing process (Sass, Schmitt, & Marsh, 2014). Thus, the interpretation of CFI, Δ CFI (Cheung & Rensvold, 2002), and $\Delta\chi^2$ (Sass, Schmitt, & Marsh, 2014) are recognized as reliable assessments of each stage of the measurement invariance testing process. Although patterns of model fit were mixed

(especially regarding RMSEA), the decision was made to continue with the steps for testing measurement invariance because no study to date has examined invariance across race/ethnicity for adolescent physical appearance measures, despite the call for this research (Kimber et al., 2015).

Invariance Testing: Females

All invariance testing results are reported in Table II, indicating that fit for the configural model was mixed but adequate (Step 1). For full metric invariance, where all factor loadings are constrained to be equal across the three racial/ethnic groups (Step 2a), the $\Delta\chi^2$ test indicated that invariance was rejected, but the Δ CFI test indicated that invariance would not be rejected. Following the $\Delta\chi^2$ results, modification indexes (MI) for the model pointed to strong noninvariance for Item 1 (*Way they look*) for White females and Item 5 (*Like their looks*) for Latina females, MI = 24.836 and MI = 10.438, respectively. For the noninvariant loading of Item 1, this indicates that content for the item *Way they look* was less ambiguous for White compared with the Black and Latina females. For Item 5, results indicated that the content for the item *Like their looks* was more ambiguous for Latina than for Black and White adolescents. To achieve partial metric invariance and proceed with the invariance testing (Step 2b), loadings for Item 1 (*Way they look*) for White and Item 5 (*Like their looks*) for Latina females were allowed to freely be estimated across the groups ($p > .05$).

Both the $\Delta\chi^2$ and Δ CFI tests indicated that scalar invariance (Step 3a), where the goal is to constrain all thresholds to be equal across groups, was rejected. MIs for the model pointed to strong noninvariance of thresholds for Item 3 (*Physical appearance*) for Black and Latina females, Item 4 (*Good looking*) for Latina and White females, and Item 5 (*Like their looks*) for all three groups, with MIs ranging from 6 to 273. This can be taken to mean that the endorsement of these items was different for some of the racial/ethnic groups compared with the others. Partial scalar invariance (Step 3b) was achieved when almost all of the item thresholds for Latina females and about half of the thresholds for the White females were allowed to be freely estimated ($p > .05$). For full invariance of residual variances (Step 4a), or testing for strict factorial invariance, the $\Delta\chi^2$ test indicated that the model should be rejected; however, the Δ CFI test indicated that it should not be rejected. MIs pointed to the noninvariance of one item residual. Partial invariance (Step 4b) was achieved according to both the $\Delta\chi^2$ and Δ CFI tests when this item's residual variance was allowed to be freely estimated. Finally, for the overall factor variance (Step 5), the Δ CFI test indicated that invariance of factor variance would not be rejected,

Table II. Goodness of Fit Indexes for Measurement Invariance Testing of Self-Perception Profile–Physical Appearance Subscale Across White, Black, and Latino High-School-Aged Adolescents

Models	Fit indices for alternative models			Difference in fit between current and previous models				
	CFI	TLI	RMSEA (90% CI)	Δ CFI	χ^2_{diff}	df	fp	p
Females								
1. Configural (baseline model)	0.962	0.923	0.146 (0.130–0.163)	–	–	–	60	–
2a. Metric (all loadings constrained)	0.970	0.961	0.104 (0.090–0.118)	.008	32.301	8	52	.001
2b. Partial Metric (some loadings constrained)	0.971	0.959	0.108 (0.094–0.122)	.001	9.335	6	54	.156
3a. Scalar (all thresholds constrained)	0.916	0.948	0.120 (0.111–0.130)	–.055	383.408	28	26	.001
3b. Partial scalar (some thresholds constrained)	0.973	0.973	0.088 (0.076–0.100)	.057	14.446	9	45	.107
4a. Residual variances (all variances constrained)	0.973	0.973	0.088 (0.076–0.100)	0	20.197	10	45	.027
4b. Partial residual variances (some constrained)	0.972	0.971	0.090 (0.078–0.102)	–.001	15.488	9	46	.078
5. Factor variance	0.976	0.977	0.080 (0.068–0.092)	.004	16.581	2	44	.001
Males								
1. Configural (baseline model)	0.944	0.888	0.151 (0.135–0.168)	–	–	–	60	–
2. Metric (all loadings constrained)	0.964	0.953	0.098 (0.084–0.112)	.020	10.746	8	52	.217
3a. Scalar (all thresholds constrained)	0.897	0.939	0.111 (0.102–0.121)	–.067	347.282	28	24	.001
3b. Partial scalar (some thresholds constrained)	0.961	0.967	0.082 (0.071–0.093)	.064	18.999	13	39	.123
4a. Residual variances (all variances constrained)	0.961	0.967	0.082 (0.071–0.093)	0	23.174	10	39	.010
4b. Partial residual variances (some constrained)	0.963	0.968	0.080 (0.069–0.092)	.002	10.959	9	40	.279
5. Factor variance	0.961	0.968	0.080 (0.070–0.092)	–.002	18.708	2	38	.001

Note. CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean squared error of approximation; χ^2_{diff} = chi-square difference test; df = degrees of freedom; fp = free parameters; Δ CFI = change in comparative fit index.

Boldface values indicate significant results.

but the $\Delta\chi^2$ test indicated that the model should be rejected. The result based on the $\Delta\chi^2$ test indicated that Black and Latina females had significantly less variance in the latent factor of physical appearance perceptions compared with White females.

Invariance Testing: Males

Results are reported in Table II, indicating that again fit for the male configural model was mixed but adequate (Step 1). For full metric invariance (Step 2), where all factor loadings are constrained to be equal across the three racial/ethnic groups, the $\Delta\chi^2$ test indicated that invariance was achieved, but results from the Δ CFI test indicated that invariance was not achieved. Because the results from the $\Delta\chi^2$ indicated invariance, MI tests for noninvariant loadings to achieve partial metric invariance were not conducted.

Full scalar invariance (Step 3a), where thresholds are held equal, was not achieved according to results from both the $\Delta\chi^2$ and Δ CFI tests. MIs pointed to nonequivalence of Item 2 (*Body different*) for Black and Latino adolescents, Item 4 (*Good looking*) for all groups, and Item 5 (*Like their looks*) for Latino and White adolescents, with MIs ranging from 5 to 107. As in the female sample, this means that the endorsement of Items 2, 4, and 5 was different for each of the three racial/ethnic groups when compared with each other. To continue with the testing, partial scalar invariance (Step 3b) was achieved after 10 thresholds for the Latino male adolescents and 3 for both Black and White male adolescents were allowed to be freely estimated. For invariance testing of the model where residual variances were constrained (Step 4a), the $\Delta\chi^2$

test indicated that invariance should be rejected, but the Δ CFI test indicated that invariance was not rejected. MIs pointed to the noninvariance of one item residual. Partial invariance (Step 4b) was achieved according to both $\Delta\chi^2$ and Δ CFI tests when this item's residual variance was allowed to be freely estimated. Finally, for the overall factor variance (Step 5), the $\Delta\chi^2$ indicated that invariance of structural variance was rejected, but the Δ CFI test indicated that invariance should not be rejected. The results based on the $\Delta\chi^2$ indicated that Black and Latino males have significantly less variance in the latent factor of physical appearance perceptions compared with White males.

Discussion

Previous research has indicated that having a poor perception of one's physical appearance is common among high-school-aged adolescents. Using the physical appearance subscale from the SPPA, our results initially indicated that there were significant racial/ethnic differences in observed mean scores. Both Black males and females reported the most positive perception about their physical appearance and Latino males and females the most negative perceptions, with White males and females reporting in between. However, further examination using measurement invariance testing indicated that these group differences may not have reflected true differences in physical appearance perceptions. We found evidence against measurement invariance, indicating that the construct of physical appearance perception, as measured by the SPPA–PA, may not be conceptualized in the same way across

racial/ethnic groups. Therefore, results from this measure should not be compared across 15–16-year-old adolescents of different race/ethnicity.

More specifically for males, invariance of thresholds was not achieved, indicating that, in particular, Latino males at a given level of physical appearance perception endorse items from the SPPA–PA at a different rate compared with Black and White males. For example, results indicated that Latino males reported more negative perceptions for items *Body different*, *Good looking*, and *Like their looks* compared with Black and White males. Previous research has indicated that for Latino youth, acculturation as indicated by generational status and primary language use may play a role in perceptions of body and physical appearance (Nieri, Kulis, Keith, & Hurdle, 2005). Those who are less acculturated, especially if male, report higher levels of physical appearance dissatisfaction. The current study also found that for Black adolescent males, items *Body different* and *Good looking* were endorsed at a different rate compared with Latino and White males, with Blacks reporting more positive perceptions. This is in line with previous findings that Black adolescents have the highest levels of body satisfaction stemming from a larger ideal body size compared with Latino and White adolescents (George & Franko, 2010).

For females, the evidence against racial/ethnic invariance was even stronger, with lack of equivalence found for both factor loadings and thresholds. Loadings were invariant for item *Way they look* for White females and item *Like their looks* for Latina females compared with the other racial/ethnic groups. This indicates that the content of these items may be interpreted differently for some of the racial/ethnic groups. Further, similar to the findings with Latino adolescent males, Latina females endorsed items *Physical appearance*, *Good looking*, and *Like their looks* at different rates compared with Black and White adolescent females. One cultural factor that may account for these is the social norm of higher body satisfaction at all sizes for Black adolescent females compared with White females (Martin, May, & Frisco, 2010). Again, acculturation may play a role for Latina females, where previous findings indicate that lower levels of acculturation are associated with higher physical appearance satisfaction (Nieri et al., 2005).

The evidence against measurement invariance indicates that there are significant differences in how high-school-aged adolescents from different racial/ethnic groups respond to the items even when or if they might ascribe to similar perceptions about their physical appearance. Further, this rejection of invariance at several levels (e.g., loadings and thresholds) negatively affects the ability to make substantive comparisons

across race/ethnicity for both males and females. Lack of scalar, or strong factorial, invariance, where the thresholds were constrained to be equal across groups, was found for both males and females, indicating that any substantive findings with regard to factor and observed means and factor variances among the three racial/ethnic groups may be unreliable. Our findings that differences exist among the three racial/ethnic groups in perceptions of physical appearance are consistent in a general sense with some previous research, but the direction of these differences varies (Neumark-Sztainer et al., 2006; Thompson et al., 2003). Previous studies found that Blacks have the most positive physical appearance perceptions and Whites the most negative perceptions among the racial/ethnic groups. Our descriptive results found that Latinos reported the worst perceptions about their physical appearance compared with White and Black adolescents. However, these different findings from comparing mean scores across the three racial/ethnic groups are partially explained by our measurement invariance analysis. Moreover, our findings differ from the findings from the one published study that has assessed racial/ethnic measurement equivalence in physical appearance perceptions in youth (Michaels et al., 2007). In that study, complete measurement invariance was achieved across Black, Latino, and White children for the overall SPPC and partial invariance for the physical appearance subscale. However, comparison with the current study is hampered by different versions of the SPP being used in the two studies (SPPC vs. SPPA) because of the age differences in the samples. Further, the current study examines adolescents in high school, at ages 15–16, where the majority of youth are well into or have completed pubertal development. Indeed, pubertal development has been linked to variation in physical appearance and body perceptions among youth (Eisenberg et al., 2006), making comparisons of findings with the previous younger study sample difficult. To our knowledge, our study is the first to examine invariance on a physical appearance measure among Black, Latino, and White high school youth.

Several limitations in the present study need to be noted. The first is that data were drawn from high-school-aged adolescents in three specific communities, limiting generalizability to the national high-school-aged adolescent population. In addition, Latino adolescents were mainly recruited in Houston and Los Angeles, therefore representing a heritage primarily from Mexico and Central America, which may not reflect Latinos with different origins. Only one subscale of the SPPA was analyzed, thus limiting the ability to assess the invariance of the entire instrument. Given the way measurement invariance testing is conducted, other important covariates, such as SES, were not

included in the analysis. Further, the use of only level of education for either parent in the household may not provide a complete account of the adolescent's level of SES. Previous research has indicated that differences exist in self-perceptions about physical appearance among youth of varying SES (George & Franko, 2010).

The present findings point to several directions for future research, including assessment of measurement equivalence of the most commonly used physical appearance measures and development of new measures tailored specifically to each of the largest racial/ethnic groups in the United States. Further examination of how physical appearance is conceptualized among diverse adolescent populations through the use of focus groups may further clarify cultural factors associated with these differences. Identifying measures that are invariant and developing new, more precise measures will enable future researchers to more precisely assess racial/ethnic disparities in pediatric health. As these measures are often used to establish whether relationships exist with other important health factors, such as disordered eating patterns or weight loss-related behaviors, it is important to measure true differences across an increasingly diverse adolescent population in the United States.

Conclusions

The SPPA-PA is a commonly used method to assess negative or positive perceptions about one's physical appearance. This study is one of the first studies to examine differences in measurement of physical appearance perceptions among racial/ethnic groups in high-school-aged adolescents. Whereas negative perceptions about physical appearance appear to be very common among youth at this age and differences may exist among high-school-aged adolescents of different racial/ethnic membership, it remains unclear whether these differences reflect true differences or measurement artifact. Different responses elicited on this scale appear to be partially due to racial/ethnic group membership rather than real or true differences on the construct of perceived physical appearance. Researchers may need to develop racially/ethnically relevant measures for perceived physical appearance. This may also indicate that interventions to address associated body image behaviors, such as unsafe weight loss methods, may need to be racially/ethnically relevant to ensure healthy and realistic perceptions about one's body and physical appearance.

Funding

The Healthy Passages study was funded by the Centers for Disease Control and Prevention (Cooperative Agreements CCU409679, CCU609653, CCU915773, U48DP000046,

U48DP000057, U48DP000056, U19DP002663, U19DP002664, and U19DP002665). The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the CDC. No author received any form of payment to produce this paper beyond grant funding.

Conflicts of interest: None declared.

Acknowledgments

The contributions made to this research by study participants in the Birmingham, Houston, and Los Angeles areas, other Healthy Passages investigators, field teams at each site, and the CDC Division of Adolescent and School Health are gratefully acknowledged.

References

- Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling, 16*, 397–438. doi:10.1080/10705510903008204
- Banis, H. T., Varni, J. W., Wallander, J. L., Korsch, B. M., Jays, S. M., Adler, R., ... Negrete, V. (1988). Psychological and social adjustment of obese children and their families. *Child: Care, Health and Development, 14*, 157–173. doi:10.1111/j.1365-2214.1988.tb00572.x
- Bronfenbrenner, U. (1989). Ecological systems theory. *Annals of Child Development, 6*, 187–249.
- Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods and Research, 21*, 230–258. doi:10.1177/0049124192021002005
- Calzo, J. P., Sonnevile, K. R., Haines, J., Blood, E. A., Field, A. E., & Austin, S. B. (2012). The development of associations among body mass index, body dissatisfaction, and weight and shape concern in adolescent boys and girls. *Journal of Adolescent Health, 51*, 517–523. doi:10.1016/j.jadohealth.2012.02.021
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255. doi:10.1207/S15328007SEM0902_5
- de Guzman, N. S., & Nishina, A. (2014). A longitudinal study of body dissatisfaction and pubertal timing in an ethnically diverse adolescent sample. *Body Image, 11*, 68–71. doi:10.1016/j.bodyim.2013.11.001
- Eisenberg, M. E., Neumark-Sztainer, D., & Paxton, S. J. (2006). Five-year change in body satisfaction among adolescents. *Journal of Psychosomatic Research, 61*, 521–527. doi:10.1016/j.jpsychores.2006.05.007
- Erickson, S. J., Hahn-Smith, A., & Smith, J. E. (2009). One step closer: Understanding the complex relationship between weight and self-esteem in ethnically diverse preadolescent girls. *Journal of Applied Developmental Psychology, 30*, 129–139. doi:10.1016/j.appdev.2008.11.004
- George, J. B., & Franko, D. L. (2010). Cultural issues in eating pathology and body image among children and adolescents. *Journal of Pediatric Psychology, 35*, 231–242. doi:10.1093/jpepsy/jsp064
- Haines, J., Neumark-Sztainer, D., Hannan, P., & Robinson-O'Brien, R. (2008). Child versus parent report of parental

- influences on children's weight-related attitudes and behaviors. *Journal of Pediatric Psychology*, *33*, 783–788. doi:10.1093/jpepsy/jsn016
- Harter, S. (1998). The development of self-representation. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 553–617). New York, NY: Wiley.
- Harter, S. (2012). Self-perception profile for adolescents: Manual and questionnaires. Retrieved from <https://portfolio.du.edu/downloadItem/221931>.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55. doi:10.1080/10705519909540118
- Hutchinson, S. R., & Olmos, A. (1998). Behavior of descriptive fit indexes in confirmatory factor analysis using ordered categorical data. *Structural Equation Modeling*, *5*, 344–464. doi:10.1080/10705519809540111
- Kaufman, J. S., Cooper, R. S., & McGee, D. L. (1997). Socioeconomic status and health in Blacks and Whites: The problem of residual confounding and the resiliency of race. *Epidemiology*, *8*, 621–628.
- Keery, H., van den Berg, P., & Thompson, J. K. (2004). An evaluation of the Tripartite Influence Model of body dissatisfaction and eating disturbance with adolescent girls. *Body Image: An International Journal of Research*, *1*, 237–251. doi:10.1016/j.bodyim.2004.03.001
- Kimber, M., Couturier, J., Georgiades, K., Wahoush, O., & Jack, S. M. (2015). Ethnic minority status and body image dissatisfaction: A scoping review of the child and adolescent literature. *Journal of Immigrant and Minority Health*, *17*, 1567–1579. doi:10.1007/s10903-014-0082-z
- Martin, M. A., May, A. L., & Frisco, M. L. (2010). Equal weights but different weight perceptions among US adolescents. *Journal of Health Psychology*, *15*, 493–504. doi:10.1177/1359105309355334
- Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology*, *93*, 568–592. doi:10.1037/0021-9010.93.3.568
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*, 525–543. doi:10.1007/BF02294825
- Michaels, M. L., Barr, A., Roosa, M. W., & Knight, G. P. (2007). Self-esteem: Assessing measurement equivalence in a multiethnic sample of youth. *The Journal of Early Adolescence*, *27*, 269–295. doi:10.1177/0272431607302009
- Mikolajczyk, R. T., Iannotti, R. J., Farhat, T., & Thomas, V. (2012). Ethnic differences in perceptions of body satisfaction and body appearance among U.S. schoolchildren: A cross-sectional study. *BMC Public Health*, *12*, 425–458. doi:10.1186/1471-2458-12-425
- Millsap, R. E. (2012). *Statistical approaches to measurement invariance*. New York, NY: Routledge.
- Millsap, R. E., & Olivera-Aguilar, M. (2012). Investigating measurement invariance using confirmatory factor analysis. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 380–392). New York, NY: Guilford Press.
- Muthén, L., & Muthén, B. (1998–2012). *Version 7 Mplus User's Guide*. Los Angeles, CA: Muthén & Muthén.
- Neumark-Sztainer, D., Paxton, S. J., Hannan, P. J., Haines, J., & Story, M. (2006). Does body satisfaction matter? Five-year longitudinal associations between body satisfaction and health behaviors in adolescent females and males. *Journal of Adolescent Health*, *39*, 244–251. doi:10.1016/j.jadohealth.2005.12.001
- Nieri, T., Kulis, S., Keith, V. M., & Hurdle, D. (2005). Body image, acculturation and substance abuse among boys and girls in the southwest. *The American Journal of Drug and Alcohol Abuse*, *31*, 617–639. doi:10.1081/ADA-200068418
- Paxton, S. J., Schutz, H. K., Wertheim, E. H., & Muir, S. L. (1999). Friendship clique and peer influences on body image concerns, dietary restraint, extreme weight-loss behaviors, and binge eating in adolescent girls. *Journal of Abnormal Psychology*, *108*, 255–266.
- Sass, D. A., Schmitt, T. A., & Marsh, H. W. (2014). Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators. *Structural Equation Modeling*, *21*, 167–180. doi:10.1080/10705511.2014.882658
- Schumann, B. C., Striegel-Moore, R. H., McMahon, R. P., Waclawiw, M. A., Morrison, J. A., & Schreiber, G. B. (1999). Psychometric properties of the self-perception profile for children in a biracial cohort of adolescent girls: The NHLBI Growth and Health Study. *Journal of Personality Assessment*, *73*, 260–275. doi:10.1207/S15327752JPA7302_5
- Schuster, M. A., Elliott, M. N., Bogart, L. M., Klein, D. J., Feng, J. Y., Wallander, J. L., ... Tortolero, S.R. (2014). Changes in obesity between fifth and tenth grades: A longitudinal study in three metropolitan areas. *Pediatrics*, *134*, 1051–1058. doi:10.1542/peds.2014-2195
- Schuster, M. A., Elliott, M. N., Kanouse, D. E., Wallander, J. L., Tortolero, S. R., Ratner, J. A., ... Banspach, S. W. (2012). Racial and ethnic health disparities among fifth-graders in three cities. *New England Journal of Medicine*, *367*, 735–745. doi:10.1056/NEJMsa1114353
- Thompson, S. H., Rafiroiu, A. C., & Sargent, R.G. (2003). Examining gender, racial, and age differences in weight concern among third, fifth, eighth, and eleventh graders. *Eating Behaviors*, *3*, 307–323. doi:10.1016/S1471-0153(02)00093-4
- Van den Berg, B. R., & Van Ranst, N. (1998). Self-concept in children: Equivalence of measurement and structure across gender and grade of Harter's self-perception profile for children. *Journal of Personality Assessment*, *70*, 564–582. doi:10.1207/s15327752jpa7003_13
- Veerman, J. W., ten Brink, L. T., Straathof, M. A., & Treffers, P. D. (1996). Measuring children's self-concept with a Dutch version of the self-perception profile for children: Factorial validity and invariance across a nonclinic

- and a clinic group. *Journal of Personality Assessment*, 67, 142–154. doi:10.1207/s15327752jpa6701_11
- Williams, D. R., & Collins, C. (1995). U.S. *socioeconomic and racial differences in health: Patterns and explanations*. *Annual Review of Sociology*, 21, 349–386.
- Windle, M., Grunbaum, J. A., Elliott, M., Tortolero, S. R., Berry, S., Gilliland, J., & Schuster, M. (2004). Healthy passages: A multilevel, multimethod longitudinal study of adolescent health. *American Journal of Preventive Medicine*, 27, 164–172. doi:10.1016/j.amepre.2004.04.007
- Yu, C. Y. (2002). Evaluation of model fit indices for latent variable models with categorical and continuous outcomes (*Unpublished dissertation*). Los Angeles, CA: University of California.