

Appearance Motives to Tan and Not Tan: Evidence for Validity and Reliability of a New Scale

Guy Cafri, M.A. · J. Kevin Thompson, Ph.D. ·
Megan Roehrig, Ph.D. · Ariz Rojas, M.A. ·
Steffanie Sperry, M.A. · Paul B. Jacobsen, M.A. ·
Joel Hillhouse, Ph.D.

Published online: 21 March 2008
© The Society of Behavioral Medicine 2008

Abstract

Background Risk for skin cancer is increased by UV exposure and decreased by sun protection. Appearance reasons to tan and not tan have consistently been shown to be related to intentions and behaviors to UV exposure and protection.

Purpose This study was designed to determine the factor structure of appearance motives to tan and not tan, evaluate the extent to which this factor structure is gender invariant, test for mean differences in the identified factors, and evaluate internal consistency, temporal stability, and criterion-related validity.

Method Five-hundred eighty-nine females and 335 male college students were used to test confirmatory factor analysis models within and across gender groups, estimate latent mean differences, and use the correlation coefficient and Cronbach's alpha to further evaluate the reliability and validity of the identified factors.

Results A measurement invariant (i.e., factor-loading invariant) model was identified with three higher-order factors: sociocultural influences to tan (lower order factors: media, friends, family, significant others), appearance reasons to tan (general, acne, body shape), and appearance reasons not to tan (skin aging, immediate skin damage). Females had significantly higher means than males on all higher-order factors. All subscales had evidence of internal consistency, temporal stability, and criterion-related validity. **Conclusions** This study offers a framework and measurement instrument that has evidence of validity and reliability for evaluating appearance-based motives to tan and not tan.

Keywords UV exposure · Tanning · Body image · Scale development · Skin cancer

Introduction

The incidence of skin cancers has reached epidemic proportions in the United States; more than 1 million cases of basal and squamous cell carcinoma and 62,190 cases of malignant melanoma are expected to be diagnosed in 2006, resulting in a total of 10,710 deaths (7,910 from melanoma [1]). Research suggests that ultraviolet (UV) radiation through sun and sunbed/sunlamp exposure is a central risk factor for the development of skin cancers, but risk can be reduced through less UV exposure and sun protection [2]. A consistent finding in the literature is that a central reason for deliberate UV exposure is the positive effect exposure has on skin appearance [3–6]. Corresponding to this association, intervention research in the area has clearly shown that manipulation of appearance attitudes results in beneficial effects on these attitudes, intentions, and behaviors to tan less and sun protect more [7–10]. Despite the relevance of

G. Cafri · J. K. Thompson · M. Roehrig · A. Rojas · S. Sperry
Department of Psychology, University of South Florida,
Tampa, FL, USA

P. B. Jacobsen
Moffitt Cancer Center and Department of Psychology,
University of South Florida,
Tampa, FL, USA

J. Hillhouse
East Tennessee State University,
Johnson, TN, USA

G. Cafri (✉)
Department of Psychology, University of South Florida,
4202 East Fowler Ave. Tampa,
Tampa, FL 33620, USA
e-mail: cafri@mail.usf.edu

appearance for tanning and sun protection, little research has addressed multiple construct domains related to a tan appearance. Our previous studies in this area have demonstrated the multidimensional nature of appearance motives related to a tan appearance [11] and how these appearance constructs may be related to UV exposure and sun-protection intentions/behaviors in a series of mediational models [12]. In the current paper, we extend our work designed to uncover and quantify appearance-based motives to UV exposure and protection in several ways: test hypothesized models of factor structure, examine gender differences in these factor models, and evaluate the internal consistency, temporal stability, and criterion-related validity of the identified factors.

Appearance constructs are central to the development of theoretical models of UV exposure and protection behaviors. Previous research suggests that one of the best predictors of UV exposure behaviors and intentions is a positive attitude related to a tan appearance [3–6]. However, research in this area has been limited because of a focus on appearance constructs that are broad or confounded with indicators that are not appearance-based. In addition, for many of the measures, there is only selective evidence of reliability and validity (e.g., rarely evidence based on factor analysis). Cafri et al. [11] found, through factor analytic methods, that there are appearance reasons for pursuing a tan that are broad (e.g., general appearance enhancement) and specific (including reducing the appearance of acne and enhancing the appearance of body shape), appearance reason for not pursuing a tan (skin-aging), and sociocultural influences to tan (i.e., media, peer, and family), each of which are significantly related to UV risk and protective intentions. Given that a growing literature supports the efficacy of appearance-based interventions [7–10], identification of these more specific constructs and their interrelations affords researchers the opportunity to develop more refined models of risk that can be used to better identify intervention targets. This is particularly important because the above referenced interventions target only a single factor related to appearance, the negative appearance effects resulting from skin aging. Moreover, as new appearance-based intervention targets are tested, scales that have strong evidence of validity and reliability can be used to precisely investigate mechanisms of change.

Although previous research identified several relevant appearance attitudes based on exploratory and confirmatory factor analytic methods [11], it did not consult the target population when constructing items, resulting in possible exclusion of important factors and item content [13], and including only females in its sample. Although previous studies have found that females are more likely than males to use indoor tanning salons and engage in outdoor tanning behaviors [14–16], males, as a group, engage in tanning

behaviors at levels that place them at risk for skin cancer, a risk that is increased when considering that their sun-protective practices tend to be less than their female counterparts [17–18]. Therefore, from a public health perspective, there is ample reason for research to be inclusive of both genders. Moreover, recent research suggests that males increasingly value their appearance and, as a result, engage in risky health behaviors [19]. Given that a tan is generally viewed as attractive, it is likely that appearance motives to tan are salient among males. Consequently, including both genders in an examination of appearance motives related to UV exposure and protection would be important, as well as examining possible gender differences.

The first goal of this study is to identify a valid factor structure within each gender. To achieve this goal, we use items from a previous study in the area [11], develop additional items based on focus groups, and test a hypothesized factor structure separately for males and females. Although these methods are necessary to establish construct validity, it is possible that the constructs will not be comparable across genders. To determine if this form of construct validity exists, it is necessary to test for measurement invariance, the second goal of this study. If certain levels of measurement invariance can be established, it is also possible to test for latent mean differences between males and females on the factors, which is the third goal. The fourth goal is to provide evidence of internal consistency, temporal stability, and criterion-related validity of the identified factors.

Method

Participants

Participants were 589 female and 335 male students from the University of South Florida. The inclusion criterion was being between the ages of 18 to 26. Participants with skin types V and VI [20] were excluded from the study because people with brown and black skin types are at a significantly reduced risk for developing skin cancer [21]. Participants were recruited from introductory psychology classes and given course credit for their participation. The study was completed online. All female data were collected October–November, 2005. The male data were collected during two time periods, October–November, 2005 ($n=164$) and February–April, 2007 ($n=171$), but using identical data collection methods. Given the differences between the two samples in terms of season and year in which the data were collected, it is necessary to test whether the groups are equal before aggregating the data. To test for group equality, the covariance matrices and means of all

items were constrained equal across the two samples using the estimation methods and tests described in the planned analyses. The result was $\chi^2(1,034)=1,511.96$, $p<0.05$, RMSEA=0.037 (90% CI 0.033, 0.041), CFI=0.96, NNFI=0.93, which suggests group equality. Consequently, data from the two time periods were aggregated. The female sample had a mean age of 19.88 (SD=1.84), and the male sample had a mean age of 20.06 (SD=1.89). The racial distribution of female participants was 80% White/Caucasian, 5% Black/African American, 3% Asian/Pacific Islander, >1% American Indian/Alaska Native, and 11% Other, and male participants 78% White/Caucasian, 4% Black/African American, 5% Asian/Pacific Islander, 1% American Indian/Alaska Native, and 12% Other. Frequencies of participant skin types [21] for females were: 12% type I, 26% type II, 36% type III, and 26% type IV, and for males: 9% type I, 27% type II, 36% type III, and 28% type IV.

Measures

Appearance Factors

The item content representing the appearance factors can be found in the [Appendix](#). Some of the items used were developed in a previous study [11]. The remaining items were developed based on focus groups. Participants of the focus groups were 18 students from the University of South Florida that closely resembled participants in the primary sample in demographic characteristics. Four focus groups and two individual interviews were conducted according to the methods prescribed by Vogt et al. [13]. The result of the focus groups and interviews was the generation of item content for some factors that had not been previously identified and adding item content to some already existing factors. Some items were also deleted based on redundancy with other items. The result was 44 total items.

UV Intention Outcome Variables

Intentions to behave were used to evaluate criterion-related validity because theory [22] and specific modeling efforts in this area [3] identify intentions as more proximal to attitudinal variables than behaviors.

Tanning Salon Intentions Participants were asked to provide a 1-year estimate of times they plan to go indoor tanning (“Please give me your best estimate of how many times you *PLAN TO* use an indoor tanning salon in the *NEXT 12 months*” [10]). Participants respond to the item by checking the box that best approximates the range of times they intend to indoor tan (0, 1–10, 11–20, 21–30, etc.). Test–retest reliability of this item over a 7- to 10-day period was adequate in a previous sample of females ($r=0.90$ [12]).

Future Sunbathing Intentions Participants were asked to provide a 1-year estimate of times they plan to go sunbathing (“Please give me your best estimate of how many times you *PLAN TO* sunbathe in the *NEXT 12 months*” [12]). The scoring method is identical to that used for future tanning salon intentions scale. Test–retest reliability of this item over a 7- to 10-day period was adequate in a previous sample of females ($r=0.88$ [12]).

Sun-Protection Intentions This is a measure modeled after the sun-protection behaviors scale [3], but with an adapted item wording that asks about intentions to protect by using sunscreen on the face, sunscreen on the body, hat use, sitting in the shade, and wearing additional clothing (e.g., item, “In the future, *how often do you intend to use sunscreen with sun protection factor (SPF) 15 or higher on your face when you were in the sun?*” [12]). The measure has shown adequate internal consistency (Cronbach’s $\alpha=0.70$) and test–retest reliability ($r=.87$) over a 7- to 10-day period in a previous study of females [12].

Planned Analyses

Study hypotheses were evaluated with AMOS 6.0 using structural equation modeling with maximum likelihood estimates of model parameters. Preliminary analyses included screening for outliers and examining the assumption of multivariate normality. There was evidence of multivariate non-normality in both male and female samples for the items used in the measurement models. Based on simulations studies [23, 24], the impact of non-normality on the χ^2 statistic, fit indexes, and estimation of SEs of parameters on the models tested in this study was deemed to be little (largest absolute value of univariate skewness for any individual item for either gender was 1.30 and kurtosis 1.31).

Confirmatory factor analysis (CFA) models were evaluated within each gender to determine the factor structure of the items. Given the relatively large number of items and hypothesized factors, estimating a single CFA model for each gender would be problematic because such models are unlikely to yield adequate fit, resulting in a search for misspecification that is extremely difficult [25]. Fitting a series of smaller models instead of a single model with all items/factors present can also cause problems, such as obscuring spurious or suppressor relations [25]. For these reasons, the jigsaw piecewise technique [25] was applied to examining the factor structure of the items. In this approach, models with a smaller number of items/factors are estimated separately, then brought together in a single model. In addition to examining model fit for the smaller models and the overall model, parameter shifts from the

smaller to larger model are examined for evidence of misspecification. In the current application, the smaller models were estimated based on grouping items/factors in the context of three higher-order factors: appearance reasons to tan, appearance reasons not to tan, and sociocultural influences to tan. The a priori decision to evaluate factor structures in this way was based on previous theory in the body image field distinguishing perceived sociocultural influences from individual body image attitudes [26], a previous study indicating small nonsignificant correlations between appearance reasons to tan and appearance reasons not to tan [11], and interventions in the area differentially targeting these appearance motives.

After establishing models that fit well within each gender, the goal was to test for measurement invariance [27]. Several kinds of invariance can be identified in CFA models. For instance, equality of factor loadings across groups offers evidence that the unit of measurement for the underlying factors is the same across groups, that is, the factors have the same meaning for males and females [27]. It should be noted that in a first-order factor model, there is only one set of factor loadings, whereas in a higher-order factor model (i.e., a second-order model) there are both first- and second-order factor loadings [28, 29]. In the present study, there is a higher-order factor structure (viz. there are three higher-order factors that are inter-correlated); therefore, it is necessary to show that both types of factor loadings are invariant. This form of invariance is often referred as “weak” invariance [30], which contrary to its title is of central interest from a construct validity perspective [31] and is a necessary precondition for establishing stricter forms of invariance. One stricter form of invariance is “strong” invariance [30], which is satisfied if the intercepts can be shown to be invariant. If this form of invariance exists, it suggests that the origins of the scales are equal across groups, a necessary condition for testing latent mean differences [27], one of the goals of this study. In a first-order factor model, it is only necessary that intercepts of measured variables are invariant, whereas to estimate mean differences of a higher-order factor, it is necessary that intercepts of the first-order factors be invariant as well [28, 29]. Finally, when tests of invariance suggest that the groups are not invariant on the parameters of interest, some argue that it is possible to identify a model that generalizes to both groups by isolating the source of the difference and allowing these parameters to be freely estimated in each group [32]. We apply this approach when appropriate, but consider its limitations in the discussion.

Several indices were used to evaluate the structural equation models. When the aim was to evaluate an individual model for fit, the root mean square error of approximation (RMSEA), comparative fit index (CFI), and the non-normed fit index (NNFI) were used because of their

relatively good performance in simulation studies [33–35]. Several cut-off values were used to judge model-data fit: RMSEA < 0.05 suggested good fit, 0.05–0.08 suggested marginal fit, and > 0.10 suggested questionable fit, and a CFI > 0.95 indicated good fit [33–35]. When testing for invariance of parameters, a series of hierarchically nested models are compared, moving progressively from an unconstrained model to more constrained models. The likelihood ratio test (i.e., $\Delta\chi^2$; differences between chi-square values) can be used to test for invariance, with nonsignificant increases in χ^2 suggesting the presence of invariance. Because this test is sensitive to sample size and non-normality [25], some have argued for the use of goodness of fit indices [36]. Based on simulation results of CFA models, one fit index criteria proposed is that CFI values should not decrease by more than 0.01 [36]. Because $\Delta\chi^2$ and Δ CFI are the most widely researched [36, 37] and used indices for judging invariance in CFA models [28, 29], they are both applied to make determinations of measurement invariance. Specifically, both indices had to exceed their prespecified thresholds for a group of parameters to be judged as non-invariant. Based on the results of simulations [36, 37] and heeding the advice of others in the field [29], we regard Δ CFI as a possibly liberal test and $\Delta\chi^2$ as conservative, but also consider the magnitude of Δ CFI in making determinations about the likelihood of invariance.

Results

Within Gender Analyses of Factor Structure

Initially, three separate higher-order confirmatory factor analysis models were tested among females. For the purposes of identification, each of the three models had one arbitrarily selected lower-order loading per lower-order factor fixed to a value of one, and one higher-order loading fixed to a value of one. In the first model, the higher-order factor represents appearance reasons for tanning, with three lower-order factors of general appearance enhancement, improvement in the appearance of acne, and improvement in body shape. In the second model, the higher-order factor represents appearance reasons not to tan and two lower-order factors of skin aging and immediate skin damage.¹ The third model has sociocultural influences to tan as a higher-order factor and four lower-order factors of media, family, friends, and significant others. Pictorial representations of these models can be found by focusing on different

¹ To achieve identification for model 2, it was necessary to impose a constraint at the lower-order factor level. The constraint imposed was to equate the error terms of the lower-order factors. Subsequent models (e.g., model 4) were also estimated with this constraint.

portions of Fig. 1; the top and right-hand side represents the first model, the bottom represents the second model, and the left-hand side represents the third model. As can be seen from Table 1, the fit statistics suggest that the higher-order factor models adequately fit the data, whereas the unidimensional models (i.e., all items load on a single factor) did not. It should be noted that model 3 with females and males converged, but yielded a solution with a negative error variance for the error term of the friends factor. Based on the simulation and recommendations of Chen et al. [38], the source of this result was determined to be an error variance sufficiently close to 0 that sampling variability led to a negative estimate (constraining the error variance to 0

yielded: $\chi^2(1)=1.58, p>0.05$). This led to re-specification of the model by constraining the variance for the error term to 0. Subsequent models included this constraint.

When all three of the models were incorporated into a single model (see Fig. 1), the result suggests adequate fit, in contrast to the alternative unidimensional models that did not (Table 1). Moreover, comparing the smaller models with the larger model in terms of standardized lower- and higher-order factor loadings, suggests small shifts in the parameter estimates when the absolute value of the differences were aggregated across the parameters, $M=.008$ ($SD=0.017$). The largest outlying differences were the loadings of the general and body shape factors on the

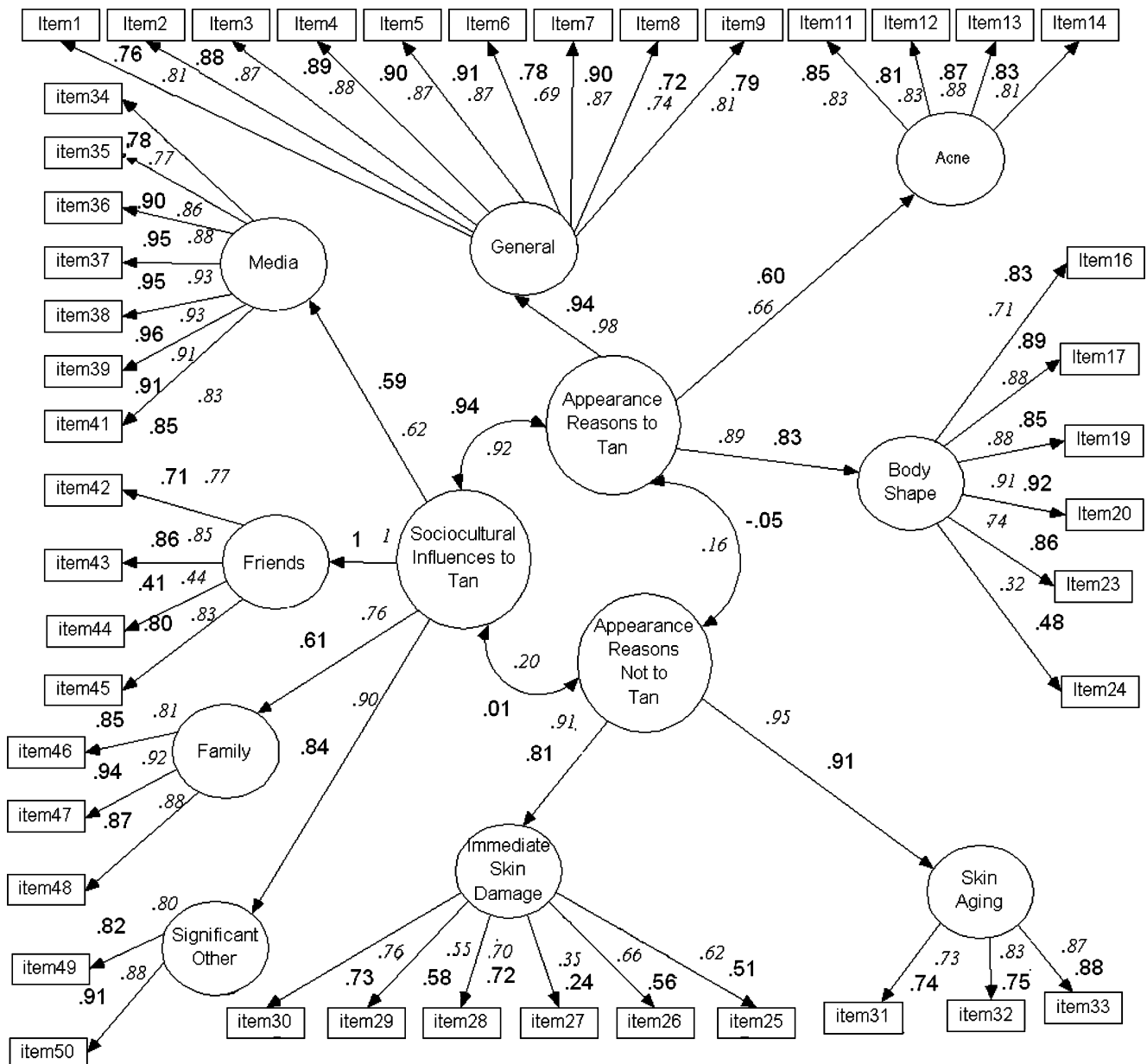


Fig. 1 Confirmatory factor analysis model. Note: The standardized estimates are presented above for females (*bold*) and males (*italics*). All factor loadings are statistically significant at the 0.05 level

Table 1 Within gender analyses of CFA model fit

Gender	Model	Hypothesized model					Alternative model (unidimensional)				
		χ^2	<i>df</i>	RMSEA (90% CI)	CFI	NNFI	χ^2	<i>df</i>	RMSEA (90% CI)	CFI	NNFI
Female	Appearance reasons to tan	540.72 *	149	0.067 (0.061,0.073)	0.96	0.96	2,05.54 *	152	0.169 (0.163, 0.174)	0.75	0.72
	Appearance reasons not to tan	44.30 *	26	0.035 (0.016, 0.052)	0.99	0.98	200.55 *	27	0.104 (0.091, 0.118)	0.89	0.86
	Sociocultural influences	425.95 *	101	0.074 (0.067, 0.081)	0.96	0.96	2,598.43 *	104	0.202 (0.195, 0.209)	0.73	0.68
	Models 1–3 estimated together ^a	2,292.61*	892	0.052 (0.049, 0.054)	0.94	0.93	7,247.82*	899	0.109 (0.107, 0.112)	0.71	0.70
Male	Appearance reasons to tan	449.81*	149	0.077 (0.070, 0.086)	0.94	0.94	1,197.3*	152	0.144 (0.136, 0.151)	0.81	0.78
	Appearance reasons not to tan	36.38	26	0.035 (0.00, 0.059)	0.99	0.99	81.84*	27	0.078 (0.059, 0.098)	0.95	0.93
	Sociocultural influences	334.56*	101	0.083 (0.073, 0.093)	0.95	0.94	1,377.21*	104	0.192 (0.182, 0.201)	0.73	0.69
	Models 1–3 estimated together ^a	2,010.66*	892	0.061 (0.058, 0.065)	0.91	0.90	4,088.50*	899	0.103 (0.100, 0.106)	0.74	0.73

χ^2 Chi-square, *RMSEA* root mean square error of approximation, *CFI* comparative fit index, *NNFI* non-normed fit index, * $p < 0.05$.
^aThe unidimensional model reported here is one in which there are three correlated first-order factors, corresponding to: appearance reasons to tan, appearance reasons not to tan, and sociocultural influences to tan. Strictly speaking, this is not a unidimensional model, but seems to be a “stronger” alternative model to make comparisons against than a true unidimensional model. Nevertheless, a pure unidimensional model was tested, which yielded inadequate fit for females, $\chi^2(902) = 11,494.39$, $p < 0.05$, *RMSEA* = 0.141 (90% CI 0.139, 0.144), *CFI* = 0.52, *NNFI* = 0.50, and males, $\chi^2(902) = 6,156.14$, $p < 0.05$, *RMSEA* = 0.132 (90% CI 0.129, 0.135), *CFI* = 0.57, *NNFI* = 0.55.

higher-order factor of appearance reasons to tan, with an increase of 0.07 units and a decrease of 0.08 units (from the smaller to the larger model), respectively. The relatively small shift in parameter estimates suggest no major misspecification present. Furthermore, a test of statistical independence was conducted [39] to evaluate the distinctiveness of the sociocultural influences to tan and appearance reasons to tan factors. Applying single-factor constraints resulted in: $\chi^2(894) = 2,330.77$, which when compared to the model without the constraints (fourth entry in Table 1), was significant, $\Delta\chi^2(\Delta df = 2) = 38.61$, $p < .05$. This suggests that the single-factor constraints are unreasonable, and two higher-order factors fit the data better than one. Collectively, the results suggest that the aggregated model fits adequately among females.

Identical analyses were applied to studying the factor structure among males. As can be seen from Table 1, the fit statistics suggest that the hypothesized higher-order factor models adequately fit the data, whereas the unidimensional models did not. The exception to this generalization is model 2, for which the unidimensional did have adequate fit, but the hypothesized higher-order model fit better. When all three of the models were incorporated into a single model, the result suggests adequate fit, in contrast to the alternative unidimensional models that did not (Table 1). As with females, model 3 yielded a solution with a negative error variance for the error term of the friends factor, which was determined to be an error variance close to 0 (constraining the error variance to 0 yielded: $\chi^2(1) = 0.85$,

$p > 0.05$). This led to re-specification of the model by constraining the variance for the error term to 0. Subsequent models included this constraint.

As with females, small shifts in the parameter estimates were observed when comparing parameter estimates from the smaller models to the aggregated model, with the absolute value of the differences yielding little change, $M = .006$ ($SD = .009$). The largest difference was the loading of the friend factor on the higher-order factor of appearance reasons to tan, with a decrease of 0.03 units (from the smaller to the larger model). The relatively small shift in parameter estimates suggest no major misspecification present. Collectively, the results suggest that the aggregated model fits adequately among males. Employing the test of statistical independence described above yielded $\chi^2(894) = 2,063.92$, $p < .05$ for a model with single-factor constraints. When compared to the model without the constraints (eighth entry in Table 1), the result was significant, $\Delta\chi^2(\Delta df = 2) = 53.26$, $p < 0.05$, suggesting that the two higher-order factors fit the data better than one. Collectively, the results suggest that the aggregated model fits adequately among males.

Measurement Invariance

The next aim is to test for measurement invariance using the aggregated models for males and females. All previously applied constraints are included in tests of measurement invariance. The threshold for measurement invariance

Table 2 Fit statistics related to tests of measurement invariance and latent mean differences

Model	Model comparison	χ^2	<i>df</i>	CFI	$\Delta\chi^2$	Δdf	ΔCFI	NNFI	RMSEA (90% CI)
<i>Model 1</i> configural model (i.e., baseline model with no constraints)	–	4,307.59*	1,785	0.927	–	–	–	0.923	0.039 (0.038, 0.041)
<i>Model 2</i> first-order factor loadings invariant	2 vs. 1	4,388.51*	1,820	0.926	80.92*	35	0.001	0.923	0.039 (0.038, 0.041)
<i>Model 3</i> Model 2 + second-order factor loadings invariant	3 vs. 2	4,395.31*	1,826	0.926	6.80	6	0.001	0.923	0.039 (0.038, 0.041)
<i>Model 4</i> Model 3 + intercepts of measured variables invariant	4 vs. 3	4,833.17*	1,861	0.914	437.86*	35	0.012	0.913	0.042 (0.040, 0.043)
Tests based on partial measurement invariance									
<i>Model 5</i> Model 4 w/ no constraints on intercepts of six items	5 vs. 3	4,499.68*	1,855	0.923	104.37*	29	0.003	0.922	0.039 (0.038, 0.041)
<i>Model 6</i> Model 6 + intercepts of first-order factors invariant	6 vs. 5	4,572.80*	1,861	0.923	73.12*	6	<0.001	0.920	0.040 (0.038, 0.041)

* $p < 0.01$

is $\Delta CFI = 0.01$ and $\Delta\chi^2$ with $\alpha = 0.01$ (Bonferonni adjustment for primary model comparisons; 0.05/5). It is necessary to first evaluate the extent to which the model fits across groups without imposing any cross-group constraints (model 1, Table 2).² This model fit adequately, therefore the next model (model 2) tested whether constraining the first-order factor loading equal across groups was feasible. Although the $\Delta\chi^2$ criterion suggested that this was not a reasonable constraint, ΔCFI was not greater than 0.01 and was approximately ten times less than this value; therefore, it was concluded that the first-order factor loadings were invariant. When the second-order loadings, in addition to the first-order loadings, were constrained equal across groups, both $\Delta\chi^2$ and ΔCFI suggested the presence of invariance (model 3). Collectively, these results imply that the unit of measurement for the underlying factors is the same across groups. Next, the intercepts of the measured variables were constrained equal, in addition to the aforementioned constraints (model 4). The result suggests that this constraint did not hold; $\Delta\chi^2$ and ΔCFI both exceeded their prespecified thresholds. Consequently, we sought to identify a model that is partially invariant [28]. One approach to identifying such a model is examining modification indices for improvement in fit resulting from releasing the group equality constraints on specific items [28]. Six intercept equality constraints were relaxed based on modification indices (i.e., a univariate Lagrange Multiplier type test), which indicated that χ^2 would decrease significantly by releasing each constraint.

² Parameters of this configural or baseline model are estimated simultaneously for males and females, which by using the default values in AMOS yields χ^2 and *df* values that are approximately equal to the addition of the χ^2 and *df* of models estimated individually within each gender [40].

The intercept parameters corresponded to items 1, 6, 9, 16, 23, and 24. Retrospective examination of the content of these items suggests that they target the appearance concerns of females more than males; therefore, the modifications seemed reasonable from a theory perspective. Comparing this partially invariant intercept model (model 5) to the model with only the first- and second-order factor loadings constrained equal (model 3) suggested the presence of invariance based on ΔCFI criterion but not $\Delta\chi^2$, which again is interpreted as being invariant by the criteria set forth in this paper. Constraining the intercepts of the lower-order factors equal across groups (model 6) suggested the presence of invariance. Consequently, a model that can be compared across groups with respect to its latent means was identified.

Internal Consistency, Temporal Stability, and Relationship with Outcome Variables

Using unit weights for individual items, the internal consistency of each of the lower-order factors was calculated (Table 3). The internal consistency ranged from 0.73 to 0.96. Using an independent sample of 14 males and 80 females, the temporal stability of the lower- and higher-order factors was evaluated by measuring individuals twice over the course of 1 week. The resulting reliability coefficients ranged from 0.74 to 0.92. Examining the relationship between the factors and the three intention variables (Table 4) suggests the presence of criterion-related validity.

Latent Mean Differences

The specifications for tests of latent mean differences of the higher-order factors were the same as those implemented in

Table 3 Descriptive statistics and reliability estimates by gender

Factor	Female			Male			Mixed gender
	Alpha	Mean	SD	Alpha	Mean	SD	Test–retest reliability (<i>r</i>)
General attractiveness (9 items)	0.95	3.60	1.06	0.95	3.14	1.09	0.92
Acne (4 items)	0.91	2.77	1.18	0.90	2.57	1.14	0.79
Body shape (6 items)	0.92	3.23	1.11	0.89	2.91	0.99	0.87
Immediate skin damage (6 items)	0.83	3.19	0.83	0.78	2.74	0.84	0.86
Skin aging (3 items)	0.73	3.05	1.06	0.85	2.51	1.04	0.74
Media (7 items)	0.97	2.31	1.14	0.96	2.02	0.99	0.89
Family (3 items)	0.77	2.39	1.15	0.91	2.36	1.10	0.78
Friends (4 items)	0.91	2.96	0.95	0.81	2.71	0.96	0.87
Significant Other (2 items)	0.86	3.10	1.19	0.83	3.04	1.14	0.80
Appearance reasons to tan	0.78	3.20	0.94	0.80	2.87	0.94	0.90
Appearance reasons not to tan	0.69	3.12	0.84	0.77	2.62	0.87	0.84
Sociocultural influences to tan	0.73	2.69	0.90	0.79	2.53	0.88	0.89
Indoor frequency intentions (1 item)	–	13.12	23.18	–	3.57	11.13	0.91
Sunbathe frequency intentions (1 item)	–	22.00	23.23	–	12.97	17.04	0.92
Sun-protection intentions (5 items)	0.79	3.48	1.21	0.75	3.48	1.23	0.87

Means and SD values for multi-item Likert scales expressed as item averages. Measures of frequency expressed as mean frequency of the behavior over the course of 12 months. Composites of subscales (i.e., entries 10–12) are expressed with unit weights for individual subscales.

other studies [28, 29]. In this model, males were chosen as the reference group, therefore their higher-order latent means were set to zero, while the latent means for females were freely estimated. The result is the difference between the factor means of the two groups. Tests of latent mean differences of the second-order factors are based on $\alpha=0.0167$ (0.05/3 tests of mean difference). Using the partially invariant measurement model previously identified, the results suggest that females score significantly higher than males on sociocultural influences to tan, mean difference=0.16, $z=3.52$, $p<0.001$, appearance reasons to tan, mean

difference=0.30, $z=4.41$, $p<0.001$, and appearance reasons not to tan, mean difference=0.57, $z=8.34$, $p<0.001$.

Discussion

Confirmatory factor analysis supported the validity of the factor structure within each gender, and tests of measurement invariance suggested that this factor structure is gender invariant. Estimates of internal consistency, temporal stability, and correlations with relevant intention out-

Table 4 Factor and criterion-related correlations for females and males

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
General attractiveness	–	0.60	0.79	0.24	0.11	0.51	0.61	0.81	0.75	0.91	0.18	0.80	0.26	0.48	–0.19
Acne	0.49	–	0.55	0.22	0.24	0.48	0.49	0.56	0.53	0.83	0.25	0.61	0.13	0.30	–0.13
Body shape	0.74	0.52	–	0.28	0.17	0.51	0.53	0.72	0.66	0.88	0.24	0.72	0.25	0.37	–0.15
Immediate skin damage	0.11	0.09	0.18	–	0.70	0.29	0.18	0.34	0.23	0.28	0.91	0.30	0.09	0.01	0.21
Skin aging	–0.12	–0.02	0.02	0.58	–	0.27	0.11	0.18	0.06	0.20	0.94	0.18	0.06	–0.08	0.25
Media	0.51	0.39	0.48	0.10	0.01	–	0.62	0.62	0.42	0.57	0.30	0.78	0.13	0.20	–0.11
Family	0.46	0.40	0.41	0.06	–0.07	0.46	–	0.68	0.57	0.62	0.16	0.86	0.04	0.30	–0.09
Friends	0.78	0.52	0.68	0.21	–0.02	0.60	0.59	–	0.75	0.80	0.27	0.91	0.23	0.40	–0.18
Significant other	0.66	0.42	0.56	0.15	–0.06	0.44	0.48	0.73	–	0.74	0.15	0.82	0.14	0.38	–0.15
Appearance reasons to tan	0.86	0.80	0.88	0.15	–0.04	0.54	0.50	0.77	0.64	–	0.25	0.81	0.24	0.44	–0.18
Appearance reasons not to tan	–0.02	0.03	0.11	0.86	0.92	0.06	–0.02	0.09	0.04	0.05	–	0.26	0.08	–0.04	0.25
Sociocultural influences to tan	0.73	0.53	0.65	0.15	–0.04	0.77	0.78	0.89	0.82	0.75	0.05	–	0.16	0.38	–0.16
Indoor intentions	0.40	0.21	0.32	–0.05	–0.23	0.25	0.21	0.32	0.27	0.37	–0.17	0.32	–	0.16	–0.09
Sunbathe intentions	0.45	0.24	0.29	–0.08	–0.25	0.15	0.19	0.32	0.26	0.38	–0.20	0.28	0.32	–	–0.19
Sun-protection Intentions	–0.32	–0.11	–0.14	0.29	0.35	–0.18	–0.13	–0.22	–0.16	–0.22	0.37	–0.21	–0.24	–0.28	–

Female correlations are below the diagonal, and male correlations are above the diagonal. For females and males $r>0.07$ has a $p<0.05$.

comes further supported the reliability and validity of the measures of the factors. Tests of latent mean differences indicated that females had significantly higher means than males on all higher-order factors. In the sections that follow, the implications of the findings are considered in terms of how the statistical methods used in this paper can improve behavioral medicine research generally and how the specific findings contribute substantively to research on skin cancer prevention.

Factor Structure

The importance of factor analysis in evaluating the construct validity of scales assessing psychological constructs is widely accepted [41]. Less well known however, are the merits of using factor analysis to test the assumption that two groups (e.g., males vs. females, adolescents vs. adults, etc.) are equivalent on a psychological construct. Indeed, whenever two groups are being compared, or if a single sample consists of a mixture of two groups, it is assumed that the dependent variable is measuring the same construct in both groups, otherwise the results are not meaningful [40]. The merits of evaluating group equivalence of psychological constructs through tests of measurement invariance are clear; however, few studies in behavioral medicine employ such tests. Despite the lack of past studies evaluating the assumption of group equivalence, the increased use of structural equation models combined with the growth in user friendly software, will likely serve to increase number of studies employing tests of measurement invariance.

In the current study, evidence of construct validity from conventional single group factor analytic methods and tests of measurement invariance suggest motives to tan and not tan are quite multidimensional and comparable across gender. It would therefore be prudent to use scales based on the higher- and lower-order factors identified in this study with future mixed-gender samples or studies designed to make comparisons across gender. Particularly lacking in skin cancer prevention research are studies designed to make inferences about risk factors (e.g., longitudinal studies) for increased UV exposure and decreased protection. Given the significant associations between the appearance constructs and these outcomes in the current study, and past research identifying the importance of appearance in tanning and sun protection [3–6], use of the constructs identified in this study to evaluate risk factors would be well worthwhile. With respect to intervention research specifically, the lower-order factors identified in this study can provide direction for the design of novel intervention targets. The necessity for new intervention targets is clear because appearance-based interventions focus almost exclusively on appearance reasons not to tan

[7–10]. However, it should be mentioned that one recent multicomponent intervention has targeted a sociocultural reason to tan, media influence, with evidence of program efficacy [42]. In addition, scales based on the lower-order factors may be used to test mechanism of change of existing appearance-based interventions, and future appearance-based interventions.

Latent Mean Differences

The conventional approach to testing mean differences between two groups on a scale representing a psychological construct is to assign unit weights to the items, calculate a sum for each individual, calculate a mean and variance for each group, then compute a *t* test. A somewhat better approach is to test for differences based on latent mean. The benefit of testing for latent mean difference is that it offers a more valid test of mean differences, by using an invariant measurement model with item weights based on factor loadings, and estimate means in the absence of measurement error [40]. As with measurement invariance, tests of latent mean differences offer an improvement to existing methods but are rarely implemented. For the same reasons stated previously however, such tests should see increased use in the future.

The tests of latent mean differences conducted in this study offers substantive information about how males and females differ on the appearance constructs. Women were found to have higher scores than males on all higher-order latent factors. The finding that females have higher ratings of sociocultural influences to tan and appearance reasons to tan is consistent with their use of indoor tanning salons and sunbathing at levels greater than males [14–16], while the finding that females had higher ratings on appearance reasons not to tan is consistent with females engaging in more sun-protective practices than males [17–18]. Therefore, it seems that the gender differences related to UV exposure and protection observed in past studies is consistent with, and may be explained in part by, appearance attitudes.

Limitations and Directions for Future Research

Despite the merits of testing for measurement invariance and latent mean differences, some caveats should be heeded. For instance, with respect to test of measurement invariance, there are only a few simulations that can be used to inform statistical decisions. Although this limits confidence in the application of the statistical procedures to behavioral medicine research, such limitations are characteristic of many advanced statistical methods. Moreover, it is always possible to re-evaluate the results of previous studies in light of the results of newer simulations.

Consequently, the benefits of conducting the kind of statistical procedures used in this paper would seem to outweigh their cost. With respect to the current study specifically, the lack of simulations with higher-order factor models, mean structures, and non-normal data [29] leads to concern about how effective the cut-off criteria were in detecting non-invariant parameters. Another concern is the validity of the latent mean difference test because equality of the origin points of the scales was not found on all items across groups. Although a partial scalar invariant model was identified, partial measurement invariance is somewhat controversial, and the invariant model was obtained through the use of modification indices, which requires cross-validation [28, 29].

Another important limitation of the present study and an area for future research is the need to evaluate the identified factors in more demographically diverse samples of adults and adolescents. For instance, it will be important to test the models developed here with a younger sample, preferably high school students, who have been found to engage in high levels of UV risk behaviors [14–16]. Improved identification and evaluation of factors involved in UV exposure and protection will ultimately lead to more success in the prevention of skin cancer.

Acknowledgment Funding for this study was supported by a grant from the National Cancer Institute, R21 CA102205-01A2 to the second author.

Appendix

Appearance Reasons to Tan

General

1. Having a tan gives me more sex appeal.
2. I tan because it makes me more attractive.
3. I tan because it makes me look better.
4. I tan because it makes me more confident in my appearance.
5. I feel more confident in my appearance when I am tan.
6. I tan before a big social event because it makes me feel more attractive.
7. The tanner I am, the more attractive I feel.
8. I tan to avoid looking pale.*
9. I tan because it adds a nice glow to my appearance.*

Acne

11. When I am tan, I feel less concerned about the appearance of acne.
12. The less tan I am the more I'm worried about my acne showing.

13. I tan because it helps reduce the amount of acne on my face and body.
14. Tan skin helps me cover up acne-related scars.
15. I tan before a big social event because it helps reduce the appearance of acne. **DELETED**

Body shape

16. I look like I have less fat on my body when I am tan.
17. The more tan I am the more physically fit I look.
18. I tan because it helps me look in shape. **DELETED**
19. A tan gives my body the appearance of having more muscle tone.
20. A tan helps me look like I'm in good physical shape.
21. My muscles look more defined when I'm tan. ***DELETED**
22. I tan to reduce the appearance of cellulite. ***DELETED**
23. I look slimmer with a tan.*
24. Being tan conceals my appearance of stretch marks.

Appearance Reasons Not to Tan

Immediate Skin Damage

25. I'm concerned about getting blemished skin as a result of tanning.*
26. I'm concerned about freckling from tanning.*
27. The appearance of a sunburn makes me look unattractive.*
28. Getting sunspots worries me.*
29. I'm concerned about my skin peeling after too much tanning.*
30. I'm concerned about the appearance of rough or leathery skin from tanning.*

Aging

31. I don't tan as much as I would like because I'm worried about premature skin aging.
32. I don't tan because it will age my skin quicker.
33. I'm hesitant to tan because it will wrinkle my skin.

Sociocultural Influences

Media

34. I try to have a tan like famous people I see in magazines.
35. I wish I was as tan as celebrities in the media.
36. I want to be as tan as TV stars.
37. I wish I had a tan like people on TV.
38. I want to be as tan as people in magazines.
39. I try to be as tan as people in movies.

40. I would like my skin tone to be darker like people in TV and movies.

Friends

41. I like to be as tan as my friends.*
 42. Positive appearance comments from my friends make me want to tan more.*
 43. I receive negative appearance comments from my friends when I am not tan.*
 44. My friends say I look good when I am tan.*

Family

45. I want a tan because people in my family think it makes my skin look nice.
 46. I try to get a tan because my family members say it is attractive.
 47. I want to be tan because my family members think it makes me look healthier.

Significant Other

48. My boyfriend/girlfriend likes the way I look when I am tan.*
 49. Comments about my appearance from my boyfriend/girlfriend encourage me to tan.*

* Indicates item developed based on information provided from the focus group.

DELETED indicates removal of item because of redundancy with other item(s).

References

- American Cancer Society: *Cancer facts and figures 2006*. Retrieved from http://www.cancer.org/docroot/STT/stt_0_2006.asp?sitearea=STT&level=1.
- U.S. Department of Health and Human Services: *Ultraviolet radiation-related exposures: broad spectrum ultraviolet (UV) radiation, UVA, UVB, UVC, solar radiation, and exposure to sunlamps and sunbeds. In the Tenth Annual Report on Carcinogens (2002)*. Retrieved August 27th, 2004 from <http://ehis.niehs.nih.gov/roc/tenth/profiles/s183uvrr.pdf>.
- Jackson KM, Aiken LS. A psychosocial model of sun protection and sunbathing in young women: The impact of health beliefs, attitudes, norms, and self-efficacy for sun protection. *Health Psychol.* 2000; 19: 469–478.
- Hillhouse JJ, Turrisi R, Kastner M. Modeling tanning salon behavioral tendencies using appearance motivation, self-monitoring, and the theory of planned behavior. *Health Education Review.* 2000; 15: 405–414.
- Wichstrom L. Predictors of Norwegian adolescents' sunbathing and use of sunscreen. *Health Psychol.* 1994; 13: 412–420.
- Leary MR, Jones JL. The social psychology of tanning and sunscreen use: self-presentational motives as a predictor of health risk. *J Appl Soc Psychol.* 1993; 23: 1390–1406.
- Mahler HIM, Kulik JA, Gibbons FX, et al. Effects of appearance-based interventions on sun protection intentions and self-reported behaviors. *Health Psychol.* 2003; 22: 199–209.
- Gibbons FX, Gerrard M, Lane DJ, et al. Using UV photography to reduce use of tanning booths: a test of cognitive mediation. *Health Psychol.* 2005; 24: 358–363.
- Mahler HIM, Kulik JA, Harrell J, et al. Effects of UV photographs, photoaging information, and use of sunless tanning lotion on sun protection behaviors. *Arch Dermatol.* 2005; 141: 373–380.
- Hillhouse JJ, Turrisi R. An examination of the efficacy of an appearance-focused intervention to reduce UV exposure. *J Behav Med.* 2002; 25: 395–409.
- Cafri G, Thompson JK, Roehrig M, et al. An investigation of appearance motives for tanning: the development and evaluation of the Physical Appearance Reasons for Tanning Scale (PARTS) and its relation to sunbathing and indoor tanning intentions. *Body Image.* 2006; 3: 199–209.
- Cafri G, Thompson JK, Jacobsen P. Appearance reasons for tanning mediate the relationship between media influence and UV exposure and protection. *Arch Dermatol.* 2006; 142: 1067–1069.
- Vogt DS, King DW, King LA. Focus groups in psychological assessment: enhancing content validity by consulting members of the target population. *Psychological Assessment.* 2004; 16: 231–243.
- Lazovich D, Forster J, Sorensen G, et al. Characteristics associated with use or intention to use indoor tanning among adolescents. *Arch Pediatr Adolesc Med.* 2004; 158: 918–924.
- Demko CA, Borawski EA, Debanne SM, et al. Use of indoor tanning facilities by white adolescents in the United States. *Arch Pediatr Adolesc Med.* 2003; 157: 854–860.
- Davis KJ, Cokkinides VE, Weinstock MA, et al. Summer sunburn and sun exposure among U.S. youths ages 11 to 18: National prevalence and associated factors. *Pediatrics.* 2002; 110: 27–35.
- Cardinez CJ, Cokkinides VE, Weinstock MA, et al. Sun protective behaviors and sunburn experiences in parents of youth ages 11 to 18. *Preventive Medicine.* 2005; 41: 108–117.
- Cokkinides VE, Davis KJ, Weinstock MA, et al. Sun exposure and sun protection behaviors and attitudes among U.S. youth, 11 to 18 years. *Preventive Medicine.* 2001; 33: 141–151.
- Thompson JK, Cafri G (eds): *Pursuit of the Muscular Ideal: Social, Psychological and Medical Perspectives*. Washington: American Psychological Association, 2007.
- Fitzpatrick TB. The validity and practicality of sun-reactive skin types I through VI. *Arch Dermatol.* 1988; 124: 869–871.
- Goldsmith MF. Paler is better, say skin cancer fighters. *JAMA.* 1987; 257: 893–894.
- Ajzen I, Fishbein M. *Understanding Attitudes and Predicting Social Behavior*. Englewood Cliffs, NJ: Prentice-Hall; 1980.
- Curran PJ, West SG, Finch JF. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychol Methods.* 1996; 1: 16–29.
- West SG, Finch JF, Curran PJ. Structural equation modeling with nonnormal variables: problems and remedies. In: Hoyle RH, ed. *Structural Equation Modeling: concepts Issues and Applications*. Thousand Oaks CA: Sage; 1995: 56–75.
- Bollen K. Modeling strategies: in search of the holy grail. *Struct Equ Modeling.* 2000; 7: 4–81.
- Thompson JK, Heinberg L, Altabe M, et al. *Exacting Beauty*. Washington, DC: American Psychological Association; ; 1999.
- Widaman KF, Reise SP. Exploring the measurement invariance of psychological instruments: applications in the substance use domain. In: Bryant KJ, Windle M, West SG, eds. *The science of prevention: methodological advances from alcohol and substance abuse research*. Washington, DC: American Psychological Association; 1997: 281–324.
- Byrne BM, Stewart SM. The MACS approach to testing for multigroup invariance of a second-order structure: a walk through the process. *Struct Equ Modeling.* 2006; 13: 287–321.

29. Chen FF, Sousa KH, West SG. Testing measurement invariance of second-order factor models. *Struct Equ Modeling*. 2005; 12: 471–492.
30. Meredith W. Measurement invariance, factor analysis and factorial invariance. *Psychometrika*. 1993; 58: 525–543.
31. Marsh HW. Confirmatory factor analysis models of factorial invariance: a multifaceted approach. *Struct Equ Modeling*. 1994; 1: 5–34.
32. Bryne M, Shavelson RJ, Muthen B. Testing for the equivalence of factor covariance and mean structures: the issue of partial measurement invariance. *Psychol Bull*. 1989; 105: 456–466.
33. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling*. 1999; 6: 1–55.
34. Hu L, Bentler PM. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychol Methods*. 1998; 3: 424–453.
35. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structural modeling. *Psychol Methods*. 1996; 1: 130–149.
36. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equ Modeling*. 2002; 9: 233–255.
37. French BF, Finch WH. Confirmatory factor analytic procedures for the determination of measurement invariance. *Struct Equ Modeling*. 2006; 13: 378–402.
38. Chen F, Bollen KA, Paxton P, et al. Improper solutions in structural equation models. *Sociol Methods Res*. 2001; 29: 468–508.
39. van der Sluis S, Dolan CV, Stoel RD. A note on testing perfect correlations in SEM. *Struct Equ Modeling*. 2005; 12: 551–577.
40. Bryne BM. *Struct Equ Modeling with AMOS: Basic Concepts, Applications, and Programming*. Mahwah, NJ: Lawrence Erlbaum; 2001.
41. Allen MJ, Yen WM. *Introduction to Measurement Theory*. McClendon: Waveland Press; ; 2001.
42. Jackson KM, Aiken LS. Evaluation of a multicomponent appearance-based sun-protective intervention for young women: uncovering the mechanisms of program efficacy. *Health Psychol*. 2006; 25: 34–46.