Psychological Assessment
2011, Vol. 23, No. 4, 958–969

1040-3590/11/$12.00 DOI: 10.1037/a0024082

An Abbreviated Tool for Assessing Feminine Norm Conformity: Psychometric Properties of the Conformity to Feminine Norms Inventory–45

Mike C. Parent and Bonnie Moradi
University of Florida

The Conformity to Feminine Norms Inventory–45 (CFNI–45; Parent & Moradi, 2010) is an important tool for assessing level of conformity to feminine gender norms and for investigating the implications of such norms for women's functioning. The authors of the present study assessed the factor structure, measurement invariance, reliability, and validity of the CFNI–45 with data from 520 college women (55% White). Confirmatory factor analyses with data from this sample suggested acceptable fit for the posited 9-factor structure. Furthermore, analyses of measurement invariance indicated similar structural properties with members of socioculturally dominant (i.e., White) and nondominant (i.e., women of color) racial/ethnic status groups. Also, subscales of the CFNI-45 demonstrated acceptable internal consistency reliability coefficients, and correlations with convergent and discriminant validity indicators supported the validity of subscales scores. Overall, results offered support for the CFNI–45 as a multidimensional measure of women's conformity to feminine norms. The CFNI–45 can be used in research to facilitate evaluation of the theorized roles of conformity to feminine norms in women's mental health, vocational behavior, interpersonal relationships, and other domains. The CFNI–45 can be used in clinical practice to assess and attend to clients' conformity to feminine norms as is called for in the American Psychological Association's (2007) *Guidelines for Psychological Practice with Girls and Women*.

Keywords: conformity to feminine norms, female norms, femininity, validity, reliability, factor structure

The American Psychological Association's (APA's) *Guidelines for Psychological Practice with Girls and Women* (2007) suggest that restrictive gender roles can have deleterious implications for women's mental health and well-being. These guidelines as well as feminist therapy principles and practice emphasize the importance of attending to the potential implications of gender norms for clients' experiences and presenting concerns (e.g., American Psychological Association, 2007; Chester & Bretherton, 2001; Moradi, Fischer, Hill, Jome, & Blum, 2000; Worell & Johnson, 2001). Thus, assessing conformity to feminine norms is important for research aimed at understanding women's experiences and for clinical interventions aimed at improving their mental health and well-being.

While numerous instruments have been developed to assess conformity to multiple dimensions of masculine norms (for a review, see Smiler, 2004), there is a paucity of measures to assess conformity to feminine norms. Consequently, researchers have used measures of related constructs, such as expressive personality characteristics (e.g., Bem Sex Role Inventory, Bem, 1974; and Personal Attributes Questionnaire, Spence, Helmreich, & Stapp, 1974); attitudes about the rights and roles of women and men (e.g., Feminine Role Norms Scale, Lefkowitz, Shearer, Gillen, & Espinosa-Hernandez, 2010; and Attitudes toward Women Scale, Spence & Helmreich, 1978); perceived stress associated with violating feminine gender role norms (e.g., Feminine Gender Role Stress Scale, Gillespie & Eisler, 1992); or single manifestations of feminine norms such as domestic cleanliness, modesty, and concern over physical appearance (e.g., Berg, Stephan, & Dodson, 1981; Cahill & Mussap, 2007; Harris & Sachau, 2005). These constructs may be related to conformity to feminine norms but do not directly reflect respondents' personal conformity to the multiple dimensions of such norms (Kite, 2001).

To address this gap, Mahalik et al. (2005) developed the Conformity to Feminine Norms Inventory (CFNI) as a multidimensional measure of conformity to dominant cultural norms of femininity. Specifically, Mahalik et al. sought to operationalize dominant U.S. cultural (e.g., White, heterosexual, middle class) construal of feminine norms, reasoning that such dominant cultural norms define the standards to which American women of various cultural backgrounds are held. That is, despite potential diversity in culture-specific feminine norms across cultural groups (e.g., Gibbons, Hamby, & Dennis, 1997; Williams & Best, 1990), members of socioculturally dominant and nondominant groups are exposed to and held to the standard of dominant cultural norms of femininity. Mahalik et al. employed a series of rational instrument development procedures (e.g., literature searches, focus groups, feedback from pilot samples) to generate the feminine norm domains and an initial item pool to assess those domains. Subse-

This article was published Online First June 27, 2011.

Mike C. Parent and Bonnie Moradi, Department of Psychology, University of Florida.

Correspondence concerning this article should be addressed to Mike C. Parent, Department of Psychology, University of Florida, Gainesville, FL 32611. E-mail: michael.parent@ufl.edu

quently, exploratory factor analysis of data from 731 mostly European American and heterosexual college women was used to reduce the initial item pool to an 84-item eight-factor measure to assess conformity to the following feminine norms: Nice in Relationships, Modesty, Domestic, Sexual Fidelity, Care for Children, Romantic Relationship, Thinness, and Invest in Appearance. One area of ambiguity, however, is that the Nice in Relationships factor was originally posited to reflect two separate norms: (a) Sweet and Nice and (b) Relational, but the examination of a nine-factor solution was not described, leaving some questions about optimal structure.

Since its development, research using the CFNI has identified a number of psychologically relevant correlates of conformity. Specifically, across samples of predominantly White college women, such research suggests that conformity to feminine norms is associated with lower levels of aggression (Reidy, Sloan, & Zeichner, 2009) and greater drive for leanness and disordered eating (Smolak & Murnen, 2008); may mediate the associations of feminist self-identification with greater self-esteem and lower body concerns, eating disorder symptoms, and depression (Hurt et al., 2007); and may be linked with women's channeling into Artistic, Social, and Realistic (i.e., working with objects or outdoors; Holland, 1973) career-related learning experiences (Tokar, Thompson, Plaufcan, & Williams, 2007). Thus, use of the CFNI has yielded findings that are consistent with theoretical conceptualizations and can inform research and clinical practice addressing mental health and career development issues.

Given the initial promise of the CFNI, Parent and Moradi (2010) aimed to evaluate the structural properties of the measure and optimize its length, which at 84 items may have been prohibitive for use in many research and clinical settings. They conducted a confirmatory factor analysis (CFA) of CFNI data from 243 mostly White Canadian college women. These analyses resulted in two major changes to the CFNI. First, Parent and Moradi tested both eight- and nine-factor models and found that a nine-factor model, reflecting Mahalik et al.'s (2005) initial conceptualization of Relational as a separate domain from Sweet and Nice, offered a superior fit to the data compared with the eight-factor model. Second, Parent and Moradi used item selection strategies that retained the highest and most distinctive indicators for each factor to optimize the length of the CFNI. The resultant nine-factor model with 45 items provided acceptable fit to the data, and the abbreviated subscale items yielded acceptable internal consistency reliabilities.

Despite these developments, further evaluation of the structural properties of CFNI–45 and evidence of convergent and discriminant validity are needed. Specifically, despite the multidimensionality of the CFNI, some researchers have employed total rather than subscale scores, and cautions against this practice are appearing in the literature (e.g., Parent & Moradi, 2010). Thus, the tenability of a general factor beyond the previously observed nine specific factors should be examined. In addition, neither the original CFNI nor the CFNI–45 was examined for evidence of measurement invariance across dominant and nondominant cultural groups. This comparison is important as it tests one of the theoretical assumptions of the CFNI framework—that dominant feminine norms apply as the standards against which members of socioculturally dominant and nondominant status groups are held. On the basis of this assumption, patterns of endorsement (or rejection) of dominant feminine norms should be similar across socioculturally dominant and nondominant status groups. Furthermore, only broad gender-related measures (e.g., the Bem Sex Role Inventory, Bem, 1974; and Eating Disorder Inventory–2, Garner, 1991) were used to assess the validity evidence for the original CFNI subscales, and Parent and Moradi did not present validity evidence beyond correlations of CFNI–45 subscale scores with their corresponding original-form CFNI subscale scores. Thus, evidence of construct-specific validity is needed for the CFNI–45 subscales. Similarly, the overlap of CFNI–45 subscale scores with impression management has not been evaluated, and it is important to demonstrate that CFNI–45 scores do not merely reflect socially desirable responding.

The Present Study

Our purpose in the present study was to offer additional psychometric data that can inform future use of the CFNI–45 in research and clinical practice. Specifically, several aspects of the structural properties of the CFNI–45, as well as reliability and validity evidence were examined. In terms of the structural properties of the CFNI–45, we assessed the posited nine-factor structure of the inventory using CFA. To test the tenability of total scale scores, we also tested a bifactor model that includes the nine specific factors that account for shared variance among domain-specific items and also a general factor that accounts for additional shared variance among all items (Chen, West, & Sousa, 2006); superior fit of the bifactor model would suggest that use of total CFNI–45 scores is a tenable addition to the use of subscale scores.[1] Next, given that the CFNI–45 was developed to reflect dominant cultural norms of femininity, we used CFA to explore measurement invariance (i.e., similarity of structural properties emerging from participant responses) across racial/ethnic groups that are socioculturally dominant (i.e., White) and nondominant (i.e., women of color); the theoretical supposition that the feminine norm standards enforced by dominant social groups apply to socioculturally dominant and nondominant groups suggests that invariance should be supported. In addition to these structural properties, the internal consistency reliability of the CFNI–45 subscale items was assessed. Finally, the convergent and discriminant validity of the CFNI–45 subscale scores was examined with construct-specific convergent validity indicators and impression management as a discrimination validity indicator.

With the present sample, we anticipated that (a) CFNI–45 data would fit the intended nine-factor structure (better than a bifactor structure) and demonstrate measurement invariance across White women and women of color; (b) CFNI–45 subscale items would yield adequate internal consistency reliability estimates as indi-

[1] A bifactor model posits that a general factor accounts for shared variance in the item set and that the specific factors account for additional, separate shared variance among their intended items. Support for such a model suggests the utility of overall scores as well as subscale scores. This differs from a hierarchical or second-order model approach in which a general factor is thought to account for the shared variance among the specific factors (rather than the full item set). A hierarchical model is not tenable a priori when the specific factors are known to correlate minimally, as is the case in the generally small or near zero intercorrelations of CFNI–45 factors (see Chen et al., 2006).

cated by values in the fair or better range of Ponterotto and Ruckdeschel's (2007) matrix for interpreting Cronbach's alpha; and (c) CFNI–45 subscale scores would yield medium to large correlations with their intended validity indicators and correlate minimally with impression management.

Method

Participants

Analyses were conducted with data from 520 college women. Participants ranged in age from 18 to 36 years ($M = 19.00$, $SD = 1.65$, $Mdn = 18.00$; 16 participants did not report their age). Regarding race/ethnicity, 55% of participants identified as White, 15% as Hispanic/Latina, 13% as African American/Black, 9% as Asian American/Pacific Islander, 4% as biracial or multiracial, 1% as Arab American/Middle Eastern, 1% as Haitian, less than 1% as American Indian/Native American, less than 1% as Pakistani, and 2% of participants declined to report a race or ethnicity. In terms of sexual orientation, 85% of the sample identified as exclusively heterosexual/straight, 9% identified as mostly heterosexual/straight, 2% identified as bisexual, 1% identified as mostly homosexual/lesbian, 1% identified as exclusively homosexual/lesbian, and 3% did not report a sexual orientation identity. Regarding social class, 51% of participants identified as middle class, 25% as upper middle class, 17% as lower class, 2% as upper class, 2% as working class, and 3% did not report a socioeconomic status. Finally, in terms of year at college, 55% of participants were first-year students, 16% were second-year, 14% were third-year, 12% were fourth-year, less than 1% were graduate or professional students, and 3% did not report their year in college.

Measures

The Conformity to Feminine Norms Inventory–45 (Parent & Moradi, 2010). The nine-factor, 45-item CFNI–45 was the focus of the present study. The nine subscales of the measure are Relational (sample item: "I believe that my friendships should be maintained at all costs"), Sweet and Nice (sample item: "Being nice to others is extremely important"), Invest in Appearance (sample item: "I spend more than 30 min a day doing my hair and make-up"), Domestic (sample item: "It is important to keep your living space clean"), Romantic Relationship (sample item: "Having a romantic relationship is essential in life"), Modesty (sample item: "I hate telling people about my accomplishments"), Sexual Fidelity (sample item: "I would feel guilty if I had a one-night stand"), Thinness (sample item: "I would be happier if I was thinner"), and Care for Children (sample item: "Taking care of children is extremely fulfilling"). Each subscale has five items; responses are made on a 4-point scale, with higher scores indicating greater conformity to each of the feminine norms assessed. Means, standard deviations, and Cronbach's alphas for the present sample are reported in Table 1.

International Personality Item Pool (IPIP; Goldberg et al., 2006). Four IPIP scales were used as convergent validity indicators for CFNI–45 subscales. The 10-item IPIP–Agreeableness scale (IPIP–Agree) assesses participants' endorsement of drive toward interpersonal positivity (sample item: "I have a good word for everyone") and was used as a validity indicator for the Sweet and Nice subscale of the CFNI–45. The 10-item IPIP–Friendliness scale (IPIP–Friend) assesses participants' reported openness toward close interpersonal relationships (sample item: "I make friends easily") and was used as a validity indicator for the Relational subscale of the CFNI–45. The 10-item IPIP–Modesty (IPIP–Modest) scale assesses participants' drive to downplay and avoid public acknowledgement of their own accomplishments (sample item: "I seldom toot my own horn") and was used as a validity indicator for the Modesty subscale of the CFNI–45. The 10-item IPIP–Need for Order and Cleanliness (IPIP–NOC) assesses participants' active engagement in maintaining a clean living area and an ordered life (sample item: "I like to tidy up") and was used as a validity indicator for the Domestic subscale of the CFNI–45. Responses to IPIP items are made on a 5-point scale; higher scores indicate greater endorsement of the construct assessed by the measure. In terms of validity, IPIP–Agree scores demonstrated a positive correlation with the revised NEO-Personality Inventory (Costa & McCrae, 1992) Agreeableness scores among an aggregated sample of men and women students and community members (race/ethnicity data were not reported for

Table 1
Conformity to Feminine Norms Inventory–45 Subscale Descriptive Statistics, Subscale Intercorrelations, and Latent Variable Correlations

Subscale	1	2	3	4	5	6	7	8	9	*M*	*SD*	α
1. Sweet and Nice		.38**	.04	.16**	.38**	.41**	.18**	.12*	.13*	2.22	0.44	.73
2. Relational	.30**		−.20**	.04	.03	.32**	.07	.05	.13*	1.90	0.44	.69
3. Modesty	.05	−.14**		−.18**	.11*	−.13*	−.19**	.08	−.02	1.26	0.43	.78
4. Domestic	.13**	.02	−.15**		.08	.22**	.23**	−.06	.08	2.15	0.48	.82
5. Sexual Fidelity	.32**	.04	.12**	.06		.26**	.06	−.09	−.07	2.06	0.68	.82
6. Care for Children	.32**	.25**	−.10*	.19**	.22**		.18**	−.10*	.05	2.11	0.65	.92
7. Romantic Relationship	.14**	.06	−.14**	.19**	.06	.15**		.16**	.15**	1.83	0.52	.75
8. Thinness	.10*	.08	.09*	−.02	−.07	−.10*	.17**		.20**	1.75	0.68	.87
9. Invest in Appearance	.09*	.10*	−.01	.09*	−.01	.03	.14**	.21**		1.59	0.60	.78
10. CFNI–45 total	.60**	.42**	.15**	.36**	.49**	.53**	.46**	.42**	.45**	1.87	0.24	

Note. $N = 520$. Values below the diagonal represent subscale correlations; values above the diagonal represent latent variable correlations for the Conformity to Feminine Norms Inventory–45 (CFNI–45) nine-factor model.
* $p < .05$. ** $p < .01$.

the samples; Gow, Whiteman, Pattie, & Deary, 2005); in a sample of mostly European Americans, IPIP–Friend scores were correlated positively with scores on the Hogan Personality Inventory Likeability subscale (Hogan & Hogan, 1995), which assesses openness to close relationships and interpersonal warmth (Church et al., 2006); and IPIP–Modest scores were correlated negatively with scores on a measure of competitiveness in a sample of mostly White college students (Fletcher & Nusbaum, 2008). In the present sample, Cronbach's alphas were .78 for responses to IPIP–Agree items, .89 for IPIP–Friend items, .79 for IPIP–Modest items, and .81 for IPIP–NOC items.

Brief Sexual Attitudes Survey–Permissiveness (BSAS–P; Hendrick, Hendrick, & Reich, 2006). The 10-item BSAS–P assesses participants' endorsement of noncommittant and nonrelationship-oriented sexual activity (sample item: "I do not need to be committed to a person to have sex with him/her") and was used as a validity indicator for the CFNI–45 Sexual Fidelity subscale. BSAS–P items are rated on a 5-point response scale; higher scores indicate less permissive sexual attitudes. In terms of validity, in a sample of mostly European American college students, scores on the BSAS–P were correlated in expected directions with scores on measures of commitment to relationships and "game-playing" style of love (Hendrick et al., 2006). In the present sample, responses to BSAS–P items yielded a Cronbach's alpha of .88.

Perceptions of Parenting Inventory–Enrichment (PPI–E; Lawson, 2004). The eight-item PPI–E measures the belief that raising children is rewarding (sample item: "Caring for the child would bring me happiness") and was used as a validity indicator for the Care for Children subscale of the CFNI–45. Responses are made on a 7-point scale; higher scores indicate greater endorsement of parenthood as rewarding. As evidence of validity, in a sample of college students (race/ethnicity data not reported), PPI–E scores were higher among participants who intended to have children than among those who did not (Lawson, 2004). In the present sample, responses to PPI–E items yielded a Cronbach's alpha of .93.

Relationship salience. Because no established and validated measure focusing on the importance of romantic relationships was available, a set of four items (RelaSali) drawn from the role salience literature (Super & Nevill, 1986) was used as a validity indicator for the Romantic Relationships subscale of the CFNI–45. Role salience defines the importance of a construct as a person's commitment in, participation in, and knowledge about the role (Super & Nevill, 1986). As such, items on the measure reflected devotion of time and energy to establishing romantic relationships (a typical item would be a statement regarding the priority that a respondent assigned to romantic relationships) Responses were made on a 5-point scale, higher scores indicating greater importance of romantic relationships. In terms of validity evidence for this measure, in the present sample, an analysis of variance (ANOVA), Welch $F(2, 206.71) = 30.20$, $p < .001$, and Bonferonni-corrected post hoc tests indicated that single women had lower RelaSali scores, $M = 2.59$, $SD = 0.94$, compared with women who were dating, $M = 2.98$, $SD = 0.74$; $t(439) = 4.53$, $p < .001$, $d = 0.46$ or married, $M = 3.27$, $SD = 0.55$; $t(340) = 5.53$, $p < .001$, $d = 0.88$. In the present sample, responses to RelaSali items had a Cronbach's alpha of .91.

Objectified Body Consciousness Scale (OBCS; McKinley & Hyde, 1996). Two subscales from the OBCS were used as convergent validity indicators for CFNI–45 subscales. The eight-item OBCS–Body Shame subscale (OBCS–Shame) assesses feelings of guilt and shame about one's body and weight (sample item: "When I can't control my weight, I feel like something must be wrong with me") and was used as a validity indicator for the Thinness subscale of the CFNI–45. The eight-item OBCS–Surveillance subscale (OBCS–Surv) assesses preoccupation with one's appearance and concern over how one looks to others (sample item: "During the day, I think about how I look many times") and was used as a validity indicator for the Invest in Appearance subscale of the CFNI–45. Responses to OBCS items are made on a 6-point scale; higher scores indicate greater body shame and body surveillance. In terms of validity, during development OBCS–Shame and OBCS–Surv items emerged as separate but correlated constructs in factor analyses of data from mostly White college women; these subscale scores were also correlated positively with dieting behaviors and negatively with body esteem (McKinley & Hyde, 1996). In the present sample, responses to items on the OBCS–Shame and OBCS–Surv each yielded a Cronbach's alpha of .86.

Balanced Inventory of Desirable Responding, Version 6–Impression Management subscale (BIDR–IM; Paulhus, 1994). The BIDR–IM has 20 items that reflect socially desirable but unrealistic responding (a typical item would be a statement in which the respondent endorsed never having littered on the street) and was used as the discriminant validity indicator for CFNI–45 subscales. BIDR–IM items are rated on a 7-point response scale, and in dichotomous coding, as described by Paulhus, the number of extreme responses (i.e., ratings of 5 and 6) are totaled, with higher scores indicating greater impression management. In the present sample, responses to items on the BIDR–IM had a Cronbach's alpha of .72.

Procedure

Participants were recruited from the undergraduate participant pool and from psychology courses at a large public university in the southeastern United States. The survey was administered online through a survey system hosted by the authors' institution. Participants were informed that the study was about gender role beliefs and self-perceptions. Participants first viewed a page containing the informed consent information and institutional review board approval of the study; they clicked a link indicating that they consented to participate prior to beginning the survey. Upon completing the survey, participants viewed a debriefing page thanking them for their participation and reiterating the researchers' contact information should they have any questions about the study. Participants received course credit as allowed by instructors.

Results

Five hundred and ninety-eight surveys were submitted. The data were screened to identify incomplete entries, and missing data were handled following guidelines set forth by Schlomer, Bauman, and Card (2010) and Schafer and Graham (2002). Specifically, 30 surveys reflected incomplete entries and were missing over one third of items included in the analyses. Another 31 entries were missing at least 20% of data on one or more measures of interest. Because this level of missing data was prohibitive to imputation, and because unequal sample sizes for validity comparisons may

have produced potentially misleading results, we excluded these 61 cases from the present analyses. We replaced the remaining missing data using expectation maximization in SPSS, an acceptable method of missing data replacement given that data appeared to be missing at random (i.e., missing data points were dispersed across the data set and were not localized in particular measures or items; see Schafer & Graham, 2002). Overall, 1 data point was replaced for 100 participants each and 2 data points were replaced for two participants. In addition, 17 cases were identified as multivariate outliers, having Mahalanobis distances significant at $p < .001$ (Tabachnick & Fidell, 1996). These cases were deleted from the data set because multivariate outliers violate assumptions of maximum likelihood estimation and because inspection of their response patterns indicated contradictory or invariant responses, both suggesting potentially careless responding. After removal of cases with missing data or cases that were identified as multivariate outliers, the final data set consisted of data from 520 women. The data for the final sample of 520 women met guidelines for univariate normality at the item level (for reliability analyses and CFA of CFNI–45 data) and subscale level (for validity analyses) as outlined by Weston and Gore (2006).

Factor Structure

To examine the factor structure of CFNI–45 data, we conducted CFAs using the raw data as input using Amos 6.0 (Arbuckle, 2005) and maximum likelihood estimation. Following prior recommendations (e.g., Martens, 2005; Worthington & Whittaker, 2006), we report the chi-square statistic with degrees of freedom, the comparative fit index (CFI), residual mean square error of approximation (RMSEA) with 90% confidence intervals (CIs), and standardized root-mean-square residual (SRMR). Criteria for acceptable fit range from values over .90 for CFI and below .10 for RMSEA and SRMR (by less conservative criteria) to values over .95 for CFI, below .08 for RMSEA, and below .06 for SRMR (by more conservative criteria; Hu & Bentler, 1999; Quintana & Maxwell, 1999; Weston & Gore, 2006). However, researchers have averred that model evaluation should take into account model complexity (i.e., degrees of freedom) and theoretical considerations rather than rely solely on fit index cutoffs as golden rules (e.g., Hu & Bentler, 1999; Quintana & Maxwell, 1999; Weston & Gore, 2006).

Sample sizes of 200 are typically adequate for CFA (e.g., Kline, 2005; Quintana & Maxwell, 1999), and models with greater degrees of freedom achieve high power at lower sample sizes (MacCallum, Browne, & Sugawara, 1996); thus, the size of the present sample was acceptable for the analyses.

Nine-factor model. All items on the CFNI–45 were included in the initial model and constrained to load only onto their intended latent factor, and latent factors were allowed to correlate. The model demonstrated adequate fit to the data, $\chi^2(909, N = 520) = 1841.30$, $p < .001$; CFI = .90; RMSEA = .04, 90% CI [.04, .05]; SRMR = .06. Although the CFI just met the .90 cutoff, CFI is known to be lower when degrees of freedom are large, as is the case in CFAs of measurement models with a large number of degrees of freedom (Kenny & McCoach, 2003).

Latent variable intercorrelations are presented in Table 1,and factor loadings are presented in Table 2. Examination of modification indices (MIs) indicated that the largest MIs were associated with freeing the covariance between the uniqueness (i.e., residual variances) of the Romantic Relationship items "My life plans do not rely on my having a romantic relationship" and "If I were single, my life would be complete without a partner" (MI = 44.83), and that between the Thinness items "I would be perfectly happy with myself even if I gained weight" and "I am terrified of gaining weight" (MI = 42.01). Freeing these two covariances, both individually and together, produced models that demonstrated slightly lower chi-square statistics but minimal changes in fit index values; the model with both covariances freed yielded: $\chi^2(907) = 1747.39$, $p < .001$; CFI = .91; RMSEA = .04, 90% CI [.04, .05]; SRMR = .06. Given that both of these covariances involved within-subscale uniqueness terms and did not reflect problematic

Table 2
Factor Loadings for the Conformity to Feminine Norms Inventory–45 Nine-Factor Model

Item	Loading	Uniqueness	Item	Loading	Uniqueness	Item	Loading	Uniqueness
	Sweet and Nice			Domestic			Romantic Relationship	
7	.68	.54	26	.82	.33	25	.82	.33
23	.63	.60	34	.69	.52	11	.78	.39
44	.62	.62	5	.68	.54	30	.54	.71
15	.55	.70	2	.67	.55	38	.49	.76
21	.50	.75	12	.60	.64	33	.40	.84
	Relational			Sexual Fidelity			Thinness	
39	.72	.48	32	.83	.31	1	.85	.28
28	.66	.56	42	.72	.48	17	.84	.29
35	.60	.64	22	.69	.52	31	.78	.39
13	.45	.80	18	.66	.56	37	.68	.54
45	.38	.86	9	.58	.66	41	.63	.60
	Modesty			Care for Children			Invest in Appearance	
19	.82	.33	43	.91	.17	14	.93	.14
29	.67	.55	16	.84	.29	27	.77	.41
10	.66	.56	8	.84	.29	6	.52	.73
24	.59	.65	40	.80	.36	20	.51	.74
4	.53	.72	36	.78	.39	3	.48	.77

Note. All loadings are significant at $p < .01$.

cross-loadings, and in light of their minimal impact on fit index values, the original model was retained.[2]

Bifactor model. Next, we assessed the fit of the bifactor model. To create this model, we added a "general" feminine norms factor, onto which all items were constrained to load, to the initial model with the nine specific factors. Thus, each CFNI–45 item was constrained to load onto both its intended specific latent factor and the general factor, and the nine specific factors and general factors were orthogonal to each other (i.e., all covariances between latent variables were constrained to zero; see Chen et al., 2006). Fit indices for the bifactor model were as follows: $\chi^2(900) = 1873.32$, $p < .001$; CFI = .89; RMSEA = .05, 90% CI [.04, .05]; SRMR = .07. To compare model fit with the initial nine factor model, we used the Akaike information criterion (AIC; Brown, 2006), which can be used to compare competing nonnest models. Lower values of AIC are indicative of better fit. The bifactor model (AIC = 2143.32) demonstrated slightly inferior fit compared with the initial model (AIC = 2093.30). Loadings on the general factor also suggested poor fit. Specifically, only 28 of the 45 items loaded significantly (at $p < .05$) onto the general factor, and 25 of these 28 significant general factor loadings were negative. By contrast, none of the loadings for the specific subscale factors were reduced from significance (at $p < .001$) by the introduction of the general factor. This pattern of findings does not support a cohesive general conformity to feminine norms factor beyond the nine specific factors.

Measurement invariance. Next, we examined measurement invariance across socioculturally dominant (i.e., White) and nondominant (i.e., women of color) racial/ethnic groups. Measurement invariance testing is used to examine similarity of structural properties across groups (Meredith, 1993; Steenkamp & Baumgartner, 1998; Steinmetz, Schmidt, Tina-Booh, Wieczorek, & Schwartz, 2009; Vandenberg & Lance, 2000). It is useful for determining whether groups have comparable response patterns at the following multiple levels of analysis. Configural invariance indicates that the number of factors and the pattern of item to factor loadings in the posited model hold across groups; it is tested by evaluating the fit of the original CFA model in both groups simultaneously. Configural invariance is a prerequisite for subsequent levels of invariance testing. Metric invariance suggests that the magnitude of factor loadings or the extent to which items reflect the underlying construct are similar across groups; it is tested by placing between-group equality constraints on corresponding factor loadings. Scalar invariance indicates that the intercept of the regression for each item on its intended factor is similar across groups and that there is not differential upward or downward response difference across groups; it is tested by placing between-group equality constraints on corresponding item intercepts (it is also necessary to constrain the latent variable means to zero for one of the groups in order for a scalar invariance model to be identified; this has no substantive impact on the model itself). Invariance of factor variances indicates that the factors have approximately equal variances across the groups; it is tested by placing between-group equality constraints on the corresponding latent variable variances. Invariance of factor covariances suggests that the factors are interrelated similarly across groups and is tested by placing between-group equality constraints on the corresponding latent factor covariances. Invariance of factor means indicates that mean scores on the latent factors are comparable across groups and is tested by placing between-group equality constraints on the means of the latent factors.

Each of these steps of invariance testing retains the constraints of the previous level; thus, constraints proceed in a stepwise fashion, and each step is compared with the one immediately preceding it. Invariance testing is further partitioned into full and partial invariance. Full invariance reflects the complete absence of any significant invariance and may be to too stringent a statistical demand to be feasible. If a particular level of invariance testing results in a significant rise in the chi-square statistic over the prior level, partial invariance can be tested by relaxing some of the constraints (identified through examining modification indices) that would improve model fit; such partial invariance is considered a practically sufficient indicator of invariance that warrants continuing to the next levels of invariance testing (Steinmetz et al., 2009).

Results of the levels of invariance testing are displayed in Table 3. To assess invariance across adjacent models, we used Chen's (2007) recommendations of change in CFI of equal to or greater than −.010 accompanied with change in RMSEA of equal to or greater than .015 as indicative of noninvariance; we did not use SRMR, as it has been shown to be differentially sensitive at different levels of invariance (Chen, 2007). As indicators of model fit, we present the chi-square statistic and chi-square/degrees of freedom ratio, CFI, RMSEA and 90% CI, SRMR, AIC, and chi-square difference significance tests between adjacent models (at $p < .001$). As indicated in Table 3, for all models, the chi-square/degrees of freedom ratio remained under the recommended cutoff of 2 (Schermelleh-Engel, Moosbrugger, & Müller, 2003), RMSEA remained under conservative criteria (i.e., <.08), and SRMR remained under less conservative criteria (i.e., <.10). Values of CFI dropped below .90, but these values should be interpreted in light of the reduced subsample sizes in the multiple-group analysis relative to the overall sample as well as the aforementioned reduction in CFI in models with large degrees of freedom.

None of the adjacent levels of invariance model comparisons resulted in model fit drop that met Chen's (2007) criteria for noninvariance. We note, however, that when constraints for scalar invariance (i.e., invariance of intercepts) were imposed, model chi-square dropped significantly (at $p < .001$) compared with the previous metric invariance model. Consistent with recommendations to test partial invariance under these conditions (Steinmetz et al., 2009), we examined modification indices for the scalar invariance model. Modification indices indicated that if three of the intercept equality constraints were relaxed, the model would no longer be significantly different from the previous model (at $p < .001$). These constraints were for Invest in Appearance Item 20 ("I

[2] We also evaluated the nine-factor model with two orthogonal method factors representing the variance associated with the direction of item coding on the basis of scoring instructions (that is, positively coded items loaded onto one factor, and reverse-coded items loaded onto a separate factor). The fit for this model was comparable to that of the model without method factors: $\chi^2(864) = 1606.38$, $p < .001$; CFI = .92; RMSEA = .04, 90% CI [.04, .04]; SRMR = .05. However, it is important to note that none of the items loaded significantly onto the positive or negative factors, indicating that the method factors did not account for any additional shared variance among the items.

Table 3
Results of Invariance Testing

Model		$\chi^2(df)$	χ^2/df ratio	CFI	RMSEA (90% CI)	SRMR	AIC	$\Delta\chi^2(df)$	$\Delta\chi^2$ Comp
A	Initial confirmatory factor analysis	1841.30 (909)	2.026	.895	.044 (.042, .047)	.057	2093.30		
B	Configural invariance	2815.35 (1818)	1.549	.888	.033 (.030, .035)	.070	3319.35	974.05 (909)	B − A
C	Loading invariance (metric)	2858.79 (1854)	1.542	.887	.032 (.030, .035)	.070	3290.79	43.44 (36)	C − B
D	Invariance of intercepts (scalar)	2952.43 (1889)	1.563	.880	.033 (.031, .035)	.071	3494.43	93.64 (35)*	D − C
D_{PS}	Invariance of intercepts (scalar) − Partial scalar	2915.50 (1886)	1.546	.884	.032 (.030, .035)	.071	3463.50	56.71 (32)	D_{PS} − C
E	Invariance of latent variance	2973.05 (1898)	1.566	.879	.033 (.031, .035)	.071	3497.05	20.62 (9)	E − D
E_{PS}	Invariance of latent variance − Partial scalar	2935.87 (1895)	1.549	.883	.033 (.030, .035)	.071	3465.87	20.37 (9)	E_{PS} − D_{PS}
F	Invariance of covariances	3007.45 (1933)	1.556	.879	.033 (.030, .035)	.074	3461.45	34.40 (35)	F − E
F_{PS}	Invariance of covariances − Partial scalar	2970.27 (1930)	1.539	.883	.032 (.030, .035)	.075	3430.27	34.40 (35)	F_{PS} − E_{PS}
G	Invariance of latent means	3090.44 (1942)	1.591	.871	.034 (.032, .036)	.075	3526.44	82.99 (9)*	G − F
G_{PS}	Invariance of latent means − Partial scalar	3050.48 (1939)	1.573	.875	.033 (.031, .035)	.075	3492.48	80.21 (9)*	G_{PS} − F_{PS}
G_{PM}	Invariance of latent means − Partial means	3021.92 (1939)	1.558	.878	.033 (.031, .035)	.075	3463.92	14.47 (6)	G_{PM} − F
$G_{PS + PM}$	Invariance of latent means − Partial scalar and partial means	2984.88 (1936)	1.542	.882	.032 (.030, .035)	.075	3432.88	14.61 (6)	$G_{PS + PM}$ − F_{PS}

Note. Partial invariance models are described in text. $\Delta\chi^2$ Comp denotes the specific comparison for the chi-square difference test. CFI = comparative fit index; RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean-square residual; AIC = Akaike information criterion; PS = partial scalar invariance; PM = partial mean invariance.
* $\Delta\chi^2$ is significant at $p < .001$.

get ready in the morning without looking in the mirror very much," reverse coded), for which women of color ($B = 1.82$, $SE = .06$) had a higher intercept (i.e., were more likely to disagree with the item) than White women ($B = 1.58$, $SE = .05$); Romantic Relationship Item 30 ("My life plans do not rely on my having a romantic relationship," reverse coded), for which White women ($B = 1.73$, $SE = .05$) had a higher intercept (i.e., were more likely to disagree with the item) than women of color ($B = 1.51$, $SE = .06$); and Sweet and Nice Item 44 ("I don't feel guilty if I lose contact with a friend," reverse coded), on which White women ($B = 2.00$, $SE = .05$) had a higher intercept (i.e., were more likely to disagree with the item) than women of color ($B = 1.78$, $SE = .05$). Given that these differences, though statistically significant, were small and involved only three of 45 items, we continued invariance testing both with and without the relaxed constraints on the scalar invariance model.

Furthermore, the latent mean invariance model also demonstrated deviation from invariance. In this case, modification indices suggested relaxing equality constraints on the Invest in Appearance, Relational, and Sweet and Nice factor means. For each of these factors in the final model with partial scalar invariance and partial latent factor mean invariance, White women had higher estimated latent factor means than women of color: Invest in Appearance, $M = .23$, $SE = .04$; Relational, $M = .20$, $SE = .05$; and Sweet and Nice, $M = .10$, $SE = .03$. When these constraints were relaxed, the model no longer differed from the adjacent model at $p < .001$.

Overall, measurement invariance was supported by the present analyses, given that changes in CFI and RMSEA between adjacent models were below Chen's (2007) cutoffs at every level of invariance testing. Furthermore, in the context of a 45-item and nine-factor model, the number and magnitude of observed differences (i.e., three items and three latent factor means) were small, and ultimately application of Chen's (2007) criteria for determining meaningful difference between levels of invariance testing indicated that there was not a meaningful difference between configural invariance (i.e., Model B) and the most stringent level of full invariance testing, latent mean invariance (i.e., Model G). Taken together, these results support invariance in responses between White women and women of color. Finally, for additional descriptive information, we provide group means for the CFNI subscales in Table 4. A series of one-way ANOVAs comparing these means across the two groups yielded results that were consistent with those of the aforementioned latent mean invariance testing. Specifically, four of the nine subscale means were significantly different between the two groups. The magnitude of group differences ranged from small to medium (partial η^2 values from .00 to .05, median = .01; see Table 4). The subscales with the largest differences corresponded with those that emerged as areas of partial invariance in the latent mean invariance testing.

Reliability

We assessed the internal consistency reliability of CFNI–45 subscale items using Cronbach's alpha coefficients. Cronbach's alphas in this sample ranged from .69 to .92, with a median value of .78 (see Table 1). These values fell within the fair to excellent range according to Ponterotto and Ruckdeschel's (2007) matrix for interpreting Cronbach's alpha.

Validity

We examined CFNI–45 subscale scores' correlations with validity indicators to assess convergent validity and those with im-

Table 4
Analysis of Variance Results for Comparison of Subscale Scores

Subscale	*F(dfs)*	Partial η^2	White women		Women of color	
			M	*SD*	*M*	*SD*
Sweet and Nice	18.40*** (1, 518)	0.03	2.29	0.43	2.13	0.43
Relational	16.17*** (1, 518)	0.03	1.97	0.42	1.82	0.44
Modesty	3.46 (1, 518)	0.01	1.29	0.43	1.22	0.42
Domestic	7.86** (1, 518)	0.02	2.09	0.47	2.21	0.49
Sexual Fidelity	0.02 (1, 518)	0.00	2.06	0.69	2.05	0.67
Care for Children	0.03 (1, 518)	0.00	2.11	0.62	2.10	0.68
Romantic Relationship	0.07 (1, 518)	0.00	1.83	0.54	1.82	0.49
Thinness	3.21 (1, 518)	0.01	1.80	0.73	1.69	0.73
Invest in Appearance	26.92*** (1, 518)	0.05	1.71	0.56	1.44	0.61
CFNI–45 total	13.03*** (1, 518)	0.03	1.91	0.23	1.83	0.23

Note. CFNI–45 = Conformity to Feminine Norms Inventory–45.
* $p < .05$. ** $p < .01$. *** $p < .001$.

pression management to assess discriminant validity. Following recommendations for correcting for measurement unreliability in validity testing (Hoyt, Warbasse, & Chu, 2006; Schmitt, 1996), we adjusted convergent and discriminant validity correlations using the correction for attenuation due to measurement error, given as

$$\frac{r_{ab}}{\sqrt{\alpha_a * \alpha_b}}$$

where r_{ab} is the raw correlation between each variable and its validity indicator, α_a is the Cronbach's alpha of the first component (i.e., the CFNI–45 subscale), and α_b is the Cronbach's alpha of the second component (i.e., the convergent or discriminant validity indicator). All correlations are presented in Table 5. Corrected correlations with validity indicators ranged from .49 to .93, with a median value of .68. According to Cohen's (1992) guidelines for interpreting effect sizes (.29 and less being considered small, .30 to .49 medium, and .50 and greater large), these values were near or in the large range. Absolute values of corrected correlations with impression management ranged from .01 to .44, with a median value of .12; these values were generally in the

Table 5
Corrected and Uncorrected (Raw) Correlations Among Conformity to Feminine Norms Inventory–45 subscales and Validity Indicators

Subscale	Correlation	IPIP–Agree	IPIP–Friend	IPIP–Modest	IPIP–NOC	BSAS	PPI	RelaSali	OBCS–Shame	OBCS–Surv	BIDR–IM
Sweet and Nice:	*r* (corrected)	**.63**	.35	.22	.09	.27	.43	.17	.09	.09	.44
	r (raw)	.47**	.26**	.17**	.07	.23**	.33**	.13**	.07	.07	.32**
Relational	*r* (corrected)	.25	**.49**	−.04	.05	.09	.30	.13	.09	.08	−.01
	r (raw)	.19**	.38**	−.03	.04	.08	.23**	.10*	.07	.07	−.01
Modesty	*r* (corrected)	.02	−.32	**.65**	−.13	.05	−.17	−.14	.17	−.04	.12
	r (raw)	.01	−.25**	.51**	−.10*	.05	−.13**	−.11*	.13**	−.03	.09*
Domestic	*r* (corrected)	.02	.18	−.07	**.82**	.16	.26	.15	−.01	.03	.06
	r (raw)	.02	.15**	−.06	.67**	.14**	.21**	.12**	−.01	.03	.05
Sexual Fidelity	*r* (corrected)	.21	.03	.17	.13	**.88**	.27	.09	−.02	−.05	.37
	r (raw)	.18**	.03	.14**	.12**	.75**	.23**	.08	−.02	−.04	.29**
Care for Children	*r* (corrected)	.35	.42	.00	.15	.29	**.68**	.17	−.11	−.04	.20
	r (raw)	.33**	.39**	.00	.14**	.25**	.63**	.16**	−.10*	−.04	.17**
Romantic Relationship	*r* (corrected)	−.03	−.01	−.06	.19	.13	.24	**.93**	.19	.28	−.07
	r (raw)	−.03	−.01	−.05	.16**	.11*	.20**	.77**	.16**	.23**	−.05
Thinness	*r* (corrected)	−.11	−.12	.05	.00	−.10	−.04	.17	**.81**	.50	−.21
	r (raw)	−.10*	−.10*	.05	.00	−.08	−.04	.15**	.70**	.43**	−.17**
Invest in Appearance	*r* (corrected)	.00	.12	−.04	.10	.04	.04	.12	.23	**.59**	−.12
	r (raw)	.00	.10*	−.04	.09	.03	.03	.10*	.19**	.48**	−.09*
M		2.79	2.80	2.18	2.61	0.85	5.01	2.79	1.99	3.20	5.91
SD		0.57	0.75	0.64	0.62	0.77	1.02	0.88	0.97	0.87	3.31

Note. Bolded values represent intended validity correlations. IPIP–Agree = IPIP Agreeableness scale; IPIP–Friend = IPIP Friendliness; IPIP–Modest = IPIP Modesty; IPIP–NOC = IPIP Need for Order and Cleanliness; BSAS = Brief Sexual Attitudes Survey–Permissiveness; RelaSali = Relationship Salience; OBCS–Shame = Objectified Body Consciousness Scale (OBCS) Shame subscale; OBCS–Surv = OBCS Surveillance subscale; BIDR–IM = Balanced Inventory of Desirable Responding (BIDR) Impression Management subscale.
* $p < .05$. ** $p < .01$.

small range, with the exception of the medium correlations of impression management with Sweet and Nice and with Sexual Fidelity.

Discussion

In this study, we investigated the structural properties, reliability, and convergent and discriminant validity of data produced by the CFNI–45. With the present sample of college women, analyses of CFNI–45 data yielded support for structural validity and measurement invariance across socioculturally dominant and nondominant racial/ethnic status groups, acceptable internal consistency reliabilities, large positive correlations with convergent validity indicators, and generally small correlations (with two exceptions) with impression management. These findings offer supportive psychometric evidence for the CFNI–45 within the generalizability boundaries of the present sample. As such, these findings can inform future use of the CFNI–45 in research and clinical practice aimed at understanding the role of conformity to feminine norms in women's experiences and functioning.

First, with regard to the structural properties of the CFNI–45 data, confirmatory factor analysis results were consistent with the hypothesized nine-factor model of the CFNI–45. All fit indices fell within acceptable ranges, though CFI did not meet more conservative recommendations. However, CFI can be negatively impacted by minor deviations from simple structure (Beauducel & Wittmann, 2005). Indeed, Beauducel and Wittmann suggested that in CFAs of constructs for which indicators may be expected to have moderate factor loadings, which may exhibit minor deviations from simple structure and in which latent variables may demonstrate moderate correlations, "models would only have a chance to be accepted when incremental fit indexes . . . are not used for model evaluation" (p. 72). In addition, the CFI has been found to be negatively impacted when models contain a large number of variables, as is the case in the present study (Kenny & McCoach, 2003). Given this context for interpreting the value of CFI, the fit indices obtained in the present study supported the intended factor structure of the CFNI–45. Furthermore, measurement invariance testing supported the stability of the CFNI–45 structure across socioculturally dominant and nondominant racial/ethnic status groups (i.e., White women and women of color) in that model–data fit between the progressively constrained invariance models remained relatively stable. Our findings did indicate that three of the items and three of the latent factors of the CFNI–45 demonstrated differences between the two groups. However, these differences were small and may have minimal impact on the interpretability of CFNI–45 scores across groups of White women and women of color. Thus, consistent with the theoretical underpinnings of the CFNI–45, the present data suggest that the pattern of endorsement of dominant norms of femininity, as assessed by the CFNI–45, is generally similar across socioculturally dominant and nondominant racial/ethnic status groups. This does not negate potential diversity in culture-specific feminine norms across cultural groups (e.g., Gibbons et al., 1997; Williams & Best, 1990), but is consistent with the theoretical underpinning that members of socioculturally dominant and nondominant groups in the United States are exposed to and held to the standard of the dominant cultural norms of femininity.

Related to the factor structure and multidimensionality of the CFNI–45, we encourage caution in interpreting total scores on the CFNI–45. Correlations among the CFNI–45 subscale scores and among the latent factors indicated that there is little association among some of the dimensions of feminine norm conformity as measured by the CFNI–45. Furthermore, tests of the bifactor model did not support the presence of a general factor over and above the subscale factors. This means that total scores do not capture meaningful item variance beyond that captured by subscale scores. Thus, as was intended in the construction of the CFNI, it may behoove researchers and clinicians to take advantage of the multidimensionality of the CFNI–45 rather than use total scale scores. This is also consistent with recommendations to use subscale scores rather than total scale scores of multidimensional gender-related instruments (e.g., Smiler & Esptein, 2010).

In terms of internal consistency reliability, Cronbach's alphas for responses to CFNI–45 subscale items fell in the fair to excellent range according to Ponterotto and Ruckdeschel's (2007) matrix for interpreting Cronbach's alpha. The values obtained in the present sample were also at or near the corresponding subscale Cronbach's alphas obtained in Parent and Moradi's (2010) sample. The two lowest Cronbach's alphas in the present study, .69 for Relational items and .73 for Sweet and Nice items, correspond with the two lowest Cronbach's alphas in Parent and Moradi's (2010) sample; these are also the two subscales that were originally intended to reflect theoretically distinct constructs but were collapsed during development of the original CFNI (Mahalik, et al., 2005). Thus, one question to arise might be whether these two subscales should be collapsed back into a single subscale, but the convergent validity and factor analytic findings suggest the originally theorized distinctiveness of the Relational subscale and Sweet and Nice subscale. Furthermore, the Cronbach's alpha values obtained in the present study must be interpreted with subscale brevity in mind (five items each), given that Cronbach's alpha values are suppressed with fewer items. Overall, responses to CFNI–45 subscale items demonstrated adequate internal consistency reliability in this sample, and these values were consistent with a prior sample. In light of the abbreviated length of CFNI–45 subscales relative to the original form of the CFNI, the generally acceptable and consistent Cronbach's alpha coefficients are noteworthy and suggest that the efficiency of the abbreviated measure has limited costs in terms of internal consistency of subscales.

In terms of convergent validity, CFNI–45 subscale scores demonstrated strong corrected correlations with scores on measures intended to be their validity indicators (corrected correlations ranged from .49 to .93, $Mdn = .65$), and yielded relatively weaker correlations with measures not intended to be their validity indicators. The strong associations with the intended validity indicators and weaker associations with other measures suggest the construct specificity of CFNI–45 subscales and are consistent with Mahalik et al.'s (2005) originally hypothesized multifactorial structure of feminine norms.

With regard to discriminant validity, CFNI–45 subscale scores demonstrated generally small corrected correlations with impression management in this sample. Thus, CFNI–45 subscale scores appear to be relatively independent from socially desirable responding in this sample. The highest CFNI–45 correlations with impression management emerged for the Sweet and Nice subscale and the Sexual Fidelity subscale. One interpretation of this pattern

is that responses to these subscales were biased in a socially desirable manner. Another possibility is that these correlations with impression management reflect conceptually substantive findings. For example, perhaps social desirability pressures for women are particularly strong for the Sweet and Nice norm and the Sexual Fidelity norm. Furthermore, Paulhus (1994) noted moderate correlations of impression management scores with agreeableness and conscientiousness traits; conscientiousness was not measured in this study, but the correlation between impression management and agreeableness in this study was large, $r_{corrected} = .55$, $r_{raw} = .41$, $p < .001$. Perhaps women who are highly agreeable and conscientious would be expected to endorse Sweet and Nice items and Sexual Fidelity items. In fact, agreeableness was the validity indicator and strongest correlate of Sweet and Nice in the present study, and the correlation between Sexual Fidelity and agreeableness was modest but significant ($r_{corrected} = .23$, $r_{raw} = .18$, $p < .001$). Thus, the overlap of Sweet and Nice scores and Sexual Fidelity scores with impression management may reflect, at least in part, the meaningful overlap of these two constructs with agreeableness and, speculatively, conscientiousness. Nevertheless, the strong associations of Sweet and Nice scores and Sexual Fidelity scores with impression management suggest cautious interpretation of scores on these two CFNI–45 subscales.

This study's findings must be interpreted in light of a number of limitations. In particular, the characteristics of the present sample of college women forms the boundaries for interpretation of the present results, and the results are not immediately generalizable outside this population. Thus, research is needed to explore the psychometric properties of the CFNI–45 with more diverse populations, including specific racial, ethnic, and sexual minority groups, mid-life and older women, and samples collected internationally. The present findings of partial measurement invariance across White and women of color provide useful groundwork in this regard.

In addition to evaluating the psychometric properties of the CFNI–45 with diverse populations, it is important to recall that the intended purpose of this measure was to assess conformity to dominant cultural norms of femininity in the United States (and potentially other Eurocentric cultures), with the reasoning that women of various backgrounds are held to those dominant cultural norms. Thus, the CFNI–45 does not assess potentially unique manifestations of feminine norms in various subcultures or populations within or outside the United States. (e.g., norms of femininity among sexual, racial, or ethnic minority subcultures or in non-Eurocentric cultures). Efforts to operationalize other cultural manifestations of feminine norms continue to be important directions for future research. It also seems useful to consider that conformity to feminine norms may not be exclusively applicable to women. The CFNI–45 might be useful in exploring individual differences in conformity to feminine norms among men and transgender individuals. Thus, investigations of the psychometric properties of the CFNI–45 might also include these populations.

Relatedly, it is important to reiterate that the present findings do not negate potential diversity in culture-specific norms of femininity. For example, research has demonstrated some cross-cultural variability in meaningfulness of some gender role constructs (Gibbons et al., 1997) and in perceptions of some characteristics as masculine or feminine (e.g., "arrogant" was perceived as a male-associated trait in most countries but as female-associated in Nigeria; Williams & Best, 1990). Rather, the present findings suggest that patterns of endorsement of dominant norms of femininity (within the United States) are similar across dominant and nondominant racial/ethnic status groups. Evaluating measurement invariance across other socioculturally dominant and nondominant status groups (e.g., sexual orientation, socioeconomic status) and operationalizing culture-specific norms of femininity are important avenues for further research.

Despite these limitations and directions for future research, the present findings may make researchers allow to be more confident in the conceptual clarity and factor structure of the CFNI–45, including its applicability to women of color in the United States. In addition, the present study suggests caution against use of CFNI–45 total scores as a broad indicator of feminine norm conformity. Thus, we recommend that researchers utilize the CFNI–45 as a multidimensional measure of feminine norm conformity, as it was originally intended, and not as a broad assessment of a single construct. This clarity along with the abbreviated length of the CFNI–45 can make the measure suitable for inclusion in research as well as in practice attending to the role of conformity to feminine norms in women's lives and functioning. Indeed, the APA's *Guidelines for Psychological Practice with Girls and Women* (2007) and the theory and practice of feminist therapy (e.g., Chester & Bretherton, 2001; Moradi et al., 2000; Worell & Johnson, 2001) emphasize the importance of assessing the impact of gender roles in women's lives. To that end, the CFNI–45 may be a useful tool for gathering data to form case conceptualizations and treatment plans that take into account clients' conformity to dominant cultural norms of femininity. The CFNI–45 may also be used as a process tool in collaborative discussion with clients' to facilitate exploration of the effects of feminine norms in their lives. Such an approach is consistent with Mahalik, Talmadge, Locke, and Scott's (2005) description of using the Conformity to Masculine Norms Inventory to assess the role of gender conformity or nonconformity in the client's life and to explore potential costs and benefits of that conformity or nonconformity. The CFNI–45 offers clinicians an abbreviated tool that can facilitate enactment of the APA guidelines. Overall, the present results provide support for the reliability, validity, and factor structure of the CFNI–45 and speak to its utility as a measure of conformity to feminine norms. Fruitful avenues for future research include investigations of the psychometric properties of the CFNI–45 with additional subgroups of women and further examination of its potential utility with men and transgender populations.

References

American Psychological Association. (2007). *Guidelines for psychological practice with girls and women: A joint task force of APA Divisions 17 and 35.* Retrieved from www.apa.org/practice/guidelines/girlsandwomen.pdf

Arbuckle, J. L. (2005). Amos Version 6.0 [Computer Program]. Chicago: SPSS.

Beauducel, A., & Wittmann, W. W. (2005). Simulation study on fit indexes in CFA based on data with slightly distorted simple structure. *Structural Equation Modeling, 12,* 41–75. doi:10.1207/s15328007sem1201_3

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42,* 155–162. doi:10.1037/h0036215

Berg, J. H., Stephan, W. G., & Dodson, M. (1981). Attributional modesty in women. *Psychology of Women Quarterly, 5,* 711–727. doi:10.1111/j.1471-6402.1981.tb01095.x

Brown, T. A. (2006). *Confirmatory factor analysis for applied research.* New York, NY: Guilford.

Cahill, S., & Mussap, A. J. (2007). Emotional reactions following exposure to idealized bodies predicts unhealthy body change attitudes and behaviors in women and men. *Journal of Psychosomatic Research, 62,* 631–639. doi:10.1016/j.jpsychores.2006.11.001

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14,* 464–504. doi:10.1080/10705510701301834

Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research, 41,* 189–225. doi:10.1207/s15327906mbr4102_5

Chester, A., & Bretherton, D. (2001). What makes feminist counseling feminist? *Feminism & Psychology, 11,* 527–545. doi:10.1177/0959353501011004006

Church, A. T., Katigbak, M. S., del Prado, A. M., Valdez-Medina, J. L., Miramontes, L. G., & Ortiz, F. A. (2006). A cross-cultural study of trait self-enhancement, explanatory variables, and adjustment. *Journal of Research in Personality, 40,* 1169–1201. doi:10.1016/j.jrp.2006.01.004

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155–159.

Costa, P. T., & McCrae, R. R. (1992). *NEO Personality Inventory–Revised*. Lutz, FL: Psychological Assessment Resources.

Fletcher, T. D., & Nusbaum, D. N. (2008). Trait competitiveness as a composite variable: Linkages with facets of the big-five. *Personality and Individual Differences, 45,* 312–317. doi:10.1016/j.paid.2008.04.020

Garner, D. M. (1991). *Eating Disorder Inventory–2 manual*. Odessa, FL: Psychological Assessment Resources.

Gibbons, J. L., Hamby, B. A., & Dennis, W. D. (1997). Researching gender-role ideologies internationally and cross-culturally. *Psychology of Women Quarterly, 21,* 151–170. doi:10.1111/j.1471-6402.1997.tb00106.x

Gillespie, B. L., & Eisler, R. M. (1992). Development of the feminine gender role stress scale: A cognitive–behavioral measure of stress, appraisal, and coping for women. *Behavior Modification, 16,* 426–438. doi:10.1177/01454455920163008

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. C. (2006). The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality, 40,* 84–96. doi:10.1016/j.jrp.2005.08.007

Gow, A. J., Whiteman, M. C., Pattie, A., & Deary, I. J. (2005). Goldberg's 'IPIP' Big-Five factor markers: Internal consistency and concurrent validation in Scotland. *Personality and Individual Differences, 39,* 317–329. doi:10.1016/j.paid.2005.01.011

Harris, P. B., & Sachau, D. (2005). Is cleanliness next to godliness? The role of housekeeping in impression formation. *Environment and Behavior, 37,* 81–101. doi:10.1177/0013916504266803

Hendrick, C., Hendrick, S. S., & Reich, D. A. (2006). The Brief Sexual Attitudes Scale. *Journal of Sex Research, 43,* 76–86. doi:10.1080/00224490609552301

Hogan, R., & Hogan, J. (1995). *Hogan Personality Inventory manual* (2nd ed.), Tulsa, OK: Hogan Assessment Systems.

Holland, J. L. (1973). *Making vocational choices: A theory of careers*. Englewood Cliffs, NJ: Prentice–Hall.

Hoyt, W. T., Warbasse, R. E., & Chu, E. Y. (2006). Construct validation in counseling psychology research. *The Counseling Psychologist, 34,* 769–805. doi:10.1177/0011000006287389

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

Hurt, M. M., Nelson, J. A., Turner, D. L., Haines, M. E., Ramsey, L. R., Erchull, M. J., & Liss, M. (2007). Feminism: What is it good for? Feminine norms and objectification as the link between feminist identity and clinically relevant outcomes. *Sex Roles, 57,* 355–363. doi:10.1007/s11199-007-9272-7

Kenny, D. A., & McCoach, D. B. (2003). Effect of number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling, 10,* 333–351. doi:10.1207/S15328007SEM1003_1

Kite, M. E. (2001). Changing times, changing gender roles: Who do we want women and men to be? In R. K. Unger (Ed.), *Handbook of the psychology of women and gender* (pp. 215–227). Hoboken, NJ: Wiley.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford.

Lawson, K. L. (2004). Development and psychometric properties of the Perceptions of Parenting Inventory. *The Journal of Psychology, 138,* 433–455.

Lefkowitz, E. S., Shearer, C. L., Gillen, M. M., & Espinosa-Hernandez, G. (2010). How gendered attitudes and traits relate to women's and men's sexual behaviors and beliefs. Manuscript submitted for publication.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure models. *Psychological Methods, 1,* 130–149. doi:10.1037/1082-989X.1.2.130

Mahalik, J. R., Morray, E., Coonerty-Femiano, A., Ludlow, L. H., Slattery, S. M., & Smiler, A. (2005). Development of the Conformity to Feminine Norms Inventory. *Sex Roles, 52,* 417–435. doi:10.1007/s11199-005-3709-7

Mahalik, J. R., Talmadge, W. T., Locke, B. D., & Scott, R. P. (2005). Using the Conformity to Masculine Norms Inventory to work with men in a clinical setting. *Journal of Clinical Psychology, 61,* 661–674. doi:10.1002/jclp.20101

Martens, M. P. (2005). The use of structural equation modeling in counseling psychology research. *The Counseling Psychologist, 33,* 269–298. doi:10.1177/0011000004272260

McKinley, N. M., & Hyde, J. S. (1996). The Objectified Body Consciousness Scale: Development and validation. *Psychology of Women Quarterly, 20,* 181–215. doi:10.1111/j.1471-6402.1996.tb00467.x

Meredith, W. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika, 58,* 525–543. doi:10.1007/BF02294825

Moradi, B., Fischer, A. R., Hill, M. S., Jome, L. M., & Blum, S. A. (2000). Does "feminist" plus "therapist" equal "feminist therapist"? An empirical investigation of the link between self-labeling and behaviors. *Psychology of Women Quarterly, 24,* 285–296. doi:10.1111/j.1471-6402.2000.tb00211.x

Parent, M. C., & Moradi, B. (2010). Confirmatory factor analysis of the Conformity to Feminine Norms Inventory and development of an abbreviated version: The CFNI–45. *Psychology of Women Quarterly, 34,* 97–109. doi:10.1111/j.1471-6402.2009.01545.x

Paulhus, D. L. (1994). *Balanced Inventory of Desirable Responding: Reference manual for BIDR Version 6.* Unpublished manuscript, Department of Psychology, University of British Columbia, Vancouver, BC, Canada.

Ponterotto, J. G., & Ruckdeschel, D. E. (2007). An overview of coefficient alpha and a reliability matrix for estimating adequacy of internal consistency coefficients with psychological research measures. *Perceptual and Motor Skills, 105,* 997–1014.

Quintana, S. M., & Maxwell, S. E. (1999). Implications of recent developments in structural equation modeling for counseling psychology research. *The Counseling Psychologist, 27,* 485–527. doi:10.1177/0011000099274002

Reidy, D. E., Sloan, C., & Zeichner, A. (2009). Gender role conformity and aggression: Influence of perpetrator and victim conformity on direct physical aggression in women. *Personality and Individual Differences, 46,* 231–235. doi:10.1016/j.paid.2008.10.008

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state

of the art. *Psychological Methods, 7,* 147–177. doi:10.1037/1082-989X.7.2.147

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research, 8,* 23–74.

Schlomer, G. L., Bauman, S., & Card, N. A. (2010). Best practices for missing data management in counseling psychology. *Journal of Counseling Psychology, 57,* 1–10. doi:10.1037/a0018082

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment, 8,* 350–353. doi:10.1037/1040-3590.8.4.350

Smiler, A. P. (2004). Thirty years after the discovery of gender: Psychological concepts and measures of masculinity. *Sex Roles, 50,* 15–26. doi:10.1023/B:SERS.0000011069.02279.4c

Smiler, A. P., & Epstein, M. (2010). Measuring gender: Options and issues. In J. C. Chrisler and D. R. McCreary (Eds.), *Handbook of gender research in psychology* (pp. 133–157). New York, NY: Springer. doi: 10.1007/978-1-4419-1465-1_7

Smolak, L., & Murnen, S. K. (2008). Drive for leanness: Assessment and relationship to gender, gender role, and objectification, *Body Image, 5,* 251–260. doi:10.1016/j.bodyim.2008.03.004

Spence, J. T., Helmreich, R., & Stapp, J. (1974). The Personal Attributes Questionnaire: A measure of sex-role stereotypes and masculinity-femininity. *JSAS: Catalog of Selected Documents in Psychology, 4,* 43–44.

Spence, J. T., & Helmreich, R. L. (1978). *Masculinity and femininity: Their psychological dimensions, correlates and antecedents.* Austin: University of Texas Press.

Steenkamp, J.-B. E. M. & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research, *Journal of Consumer Research, 25,* 78–90. doi:10.1086/209528

Steinmetz, H., Schmidt, P., Tina-Booh, A., Wieczorek, S., & Schwartz, S. H. (2009). Testing measurement invariance using multigroup CFA: Differences between groups in human values measurement. *Quality and Quantity: International Journal of Methodology, 43,* 599–616. doi: 10.1007/s11135-007-9143-x

Super, D. E., & Nevill, D. D. (1986). *The Salience Inventory: Theory, application, and research.* Palo Alto, CA: Consulting Psychologists Press.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics.* New York, NY: Harper Collins.

Tokar, D. M., Thompson, M. N., Plaufcan, M. R., & Williams, C. M. (2007). Precursors of learning experiences in social cognitive career theory. *Journal of Vocational Behavior, 71,* 319–339. doi:10.1016/j.jvb.2007.08.002

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3,* 4–69. doi:10.1177/109442810031002

Weston, R., & Gore, P. A., Jr. (2006). A brief guide to structural equation modeling. *The Counseling Psychologist, 34,* 719–751. doi:10.1177/0011000006286345

Williams, J. E., & Best, D. L. (1990). *Measuring sex stereotypes: A multinational study.* Newbury Park, CA: Sage.

Worell, J. P., & Johnson, D. M. (2001). Therapy with women: Feminist frameworks. In R. Unger (Ed.), *Handbook on the psychology of women and gender* (pp. 317–329). New York, NY: Wiley.

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34,* 806–838. doi:10.1177/0011000006288127

Received January 3, 2011
Revision received April 18, 2011
Accepted April 20, 2011 ■

Correction to Ebesutani et al. (2011)

In the article "A Psychometric Analysis of the Positive and Negative Affect Schedule for Children–Parent Version in a School Sample" by Chad Ebesutani, Kelsie Okamura, Charmaine Higa-McMillan, and Bruce F. Chorpita (*Psychological Assessment*, Vol. 23, No. 2, pp. 406–416), there was an error in the Appendix caption (pg. 416). The Appendix caption should have included the note, "Adapted from Watson, D. & Clark, L.A. (1999). The PANAS-X: Manual for the Positive and Negative Affect Schedule—Expanded form-Revised. Copyright 1994 by D. Watson and L. A. Clark; all rights reserved. PANAS-X adapted with permission."

DOI: 10.1037/a0026420