Original Article

Construct Validity of the Fear of Negative Appearance Evaluation Scale in a Community Sample of French Adolescents

Christophe Maïano[1], Alexandre J.S. Morin[2], Johana Monthuy-Blanc[3], and Jean-Marie Garbarino[4]

[1]UMR 6233 "Institute of Movement Sciences, Etienne-Jules Marey," CNRS-University of Aix-Marseille II, Marseille, France, [2]University of Sherbrooke, Department of Psychology, Sherbrooke, Québec, Canada, [3]University of Montpellier I, EA 4206 "Addictive, Performance and Health Behaviors," Montpellier, France, [4]University of Nice Sophia-Antipolis, "Human Motricity, Education, Health," Nice, France

Abstract. The main objective of the present series of studies was to test the construct validity (i.e., content, factorial, and convergent validities) of the Fear of Negative Appearance Evaluation Scale (FNAES) in a community sample of French adolescents. A total sample of 683 adolescents was involved in three studies. The factorial validity and the measurement invariance of the FNAES were verified through a series of confirmatory factor analyses. The convergent validity of the FNAES was then verified through correlational analyses. The first study showed that the content and formulation of the French FNAES items were adequate for children and adolescents. The following two studies (Studies 2 to 3) provided (a) support for the factor validity, reliability, and convergent validity of a five-item French version of the FNAES, and (b) partial support for the measurement invariance of the resulting FNAES across genders. However, the latent mean structure of the FNAES did not prove to be invariant across genders, revealing a significantly higher latent mean score of FNAES in girls relative to boys. The present results, thus, provide preliminary evidence regarding the construct validity of the FNAES in a community sample of French adolescents. Recommendations for future practice and research regarding this instrument are outlined.

Keywords: fear of negative evaluation, adolescents, gender, confirmatory factor analysis, measurement invariance, nonclinical

Introduction

During the last 20 years, the prevalence of severe levels of body image disturbances in youth has increased to the point of becoming a very salient public health concern given the known role of such disturbances in the onset and maintenance of eating disorders (Levine & Piran, 2004). Among the potential risks factors associated with severe body image disturbances, Thompson and Stice (2001) recently suggested that adolescents "fear of negative appearance evaluation (FNAE) by others" may play an important role. They defined FNAE as the: (a) apprehension about receiving negative appearance evaluations, (b) avoidance of being physically evaluated, and (c) the expectation of being negatively evaluated physically. In order to be able to detect and to study FNAE more systematically, they developed the FNAE Scale (FNAES; Thomas, Keery, Williams, & Thompson, 1998) through an adaptation of the more generic Fear of Negative Evaluation Scale (FNES; Watson & Friend, 1969). In a preliminary study of the FNAES based on a single sample of 272 North American women, Thomas et al. (1998) found support (through principal component analyses [PCA]) for an eight-item unidimensional model of the FNAES. Subsequent analyses revealed moderate correlations between the FNAES and additional measures of body image and internalization of media images/messages (i.e., convergent validity), as well as a satisfactory internal consistency coefficient (Cronbach's $\alpha = .91$).

Very few scholars have attempted to replicate this initial preliminary study. A literature review conducted within several databases (i.e., Current contents, Medline, Psychology and Behavioral Science Collection, PsycINFO) revealed a single additional study of the FNAES. Lundgren, Anderson, and Thompson (2004) recently attempted to replicate Thomas et al.'s (1998) preliminary results in the context of two studies performed on a total of 325 North American participants. In the first study, which included 165 female undergraduate participants ($M_{age} = 19.60$, $SD_{age} = 3.00$; mean Body

DOI 10.1027/1015-5759/a000004

Mass Index [BMI; weight/height*height] = 23.80, SD_{BMI} = 4.80) the objective was to replicate Thomas et al.'s (1998) PCA results and to evaluate the convergent validity of the FNAES with a different set of measures. Findings from the PCA supported a truncated six-item, unidimensional version of the FNAES. In fact, two inverted items from the original version exhibited low factor loadings and were deleted. This version also presented a satisfactory level of internal consistency (Cronbach's α = .94). Subsequent analyses of the FNAES (i.e., convergent validity) revealed low to strong correlations with several instruments measuring fear of negative evaluation from others (r = .78, p < .01), negative feelings about body size and shape (r = .70, p < .01), body dissatisfaction (r = .42, p < .01), chronic unsuccessful dieting (r = .64, p < .01), short-term successful caloric restriction (r = .46, p < .01), depression (r = .43, p < .01), and trait anxiety (r = .45, p < .01). The second study relied on a sample of 160 undergraduate college students (68 male, 92 female; M_{age} = 20.50, SD_{age} = 5.00; M_{BMI} = 24.50, SD_{BMI} = 4.80) and aimed at evaluating the association of the FNAES with additional measures. The results revealed: (a) strong correlations between the FNAES and social physique anxiety (r = .76, p < .01), dietary restraint (r = .62, p < .01), and fear of fatness (r = .69, p < .01); and (b) low to moderate correlations between the FNAES and feelings of physical attractiveness (r = –.38, p < .01), efforts devoted to physical appearance (r = .54, p < .01), satisfaction with specific body areas (r = –.57, p < .01), overweight preoccupations (r = .54, p < .01), self-classified weight (r = .30, p < .01), depression (r = .46, p < .01), binge eating (r = .52, p < .01), purgative behaviors (r = .50, p < .01), avoidance of forbidden food (r = .43, p < .01), and physical self-esteem (r = –.45, p < .01).

This single cross-validation study is insufficient to reach clear conclusions regarding the psychometric properties of this instrument. This is especially evident in samples that differ from those of the original studies, such as non-English samples or adolescent samples. This observation is dramatic given the fact that concerns about physical appearance often emerge in adolescence, following pubertal development, and appears to differ according to cultural backgrounds (Smolak, 2004). However, to systematically study these questions, one needs to be able to rely on instruments that are appropriate for the target population. The multiple methodological limitations present in the preceding studies reinforce this conclusion and the need for additional cross-validation efforts.

First, both studies examined the dimensionality of the FNAES through PCA. Despite the relative accuracy of exploratory methods, confirmatory factor analysis (CFA) appears to represent a more rigorous and complete approach to the verification of the construct validity of psychometric tools (Byrne, 2005). The advantage of CFA is that it allows for the a priori specification of a factor structure consistent with a model-based hypothesis-testing framework, as opposed to the post hoc labeling of extracted factors that is the norm in exploratory analyses. Since CFA gives the researcher the ability to verify the adequacy of the hypothetical factor structure (or of alternative hypothetical structures) against observations, and to directly model measurement errors, it is considered the gold-standard method for the evaluation of the construct validity of psychometric inventories (Byrne, 2005). Moreover, since both studies evaluated the factor structure of the FNAES in a single sample and relied on this sample to obtain a reduced version of the questionnaire, their results should clearly be replicated to avoid the risk of capitalizing on chance.

Second, most of the preceding studies have exclusively relied on samples of adult women, with the exception of Lundgren et al.'s (2004) second study (convergent validity), which included a mixed sample. Thus, it is currently unknown whether (a) there is a gender-based difference in FNAE, and (b) the factor structure of the FNAES is not only appropriate for adolescents, but is also invariant across gender. The lack of research on these two points is surprising for two main reasons. First, numerous studies (e.g., Levine & Piran, 2004; Smolak, 2004) have demonstrated that concerns about physical appearance or body image are prevalent in both boy and girl samples, and that girls tend to present a significantly higher level of self-reported body dissatisfaction in comparison to boys. It is possible to hypothesize that girls would have a significantly higher level of FNAE than boys, and that this difference would constitute evidence of the construct validity of the instrument. We, thus, hypothesized that gender-based differences in mean-levels of FNAE would be observed (with girls presenting higher levels than boys). Second, numerous studies have also showed that concerns in physical appearance or body image vary according to gender (e.g., Benjet & Hernández-Guzmán, 2002; Klomsten, Skaalvik, & Espnes, 2004). Consequently, the verification of the gender-based measurement invariance of the FNAES should also be verified to ensure that it measures FNAE in the same way in both genders. In addition, measurement invariance represents a prerequisite to the verification of gender-based differences in mean-level of FNAE (Vandenberg & Lance, 2000).

Considering this, the main objectives of the present series of studies were to: (1) develop a French version of the FNAES and test its applicability in samples of children and adolescents, and (2) examine the construct validity of this instrument in two independent samples (i.e., content validity, factor validity, measurement invariance, latent mean invariance, and convergent validity). Given the absence of a validated French version of the FNAES, the purpose of Study 1 was to develop a preliminary French version of the FNAES for adolescents and to verify the content clarity of the resulting items in a sample of adolescents. Studies 2 and 3 then sought to: (a) examine the factor structure of the FNAES in a sample of adolescents; (b) to assess the measurement invariance and latent mean invariance of the FNAES across gender; (c) to cross-validate the construct validity of the FNAES in an independent sample; (d) to investigate the measurement invariance of the FNAES between samples from studies 2 and 3; (e) to evaluate the temporal stability of the FNAES; and (f) to test the convergent validity of the FNAES by examining its relationships with measures of fear of negative evaluation, eating disorders, self-esteem, and social physique anxiety.

Materials and Methods

Sample and Procedures

Study 1

A sample of 22 children and adolescents (M_{age} = 10.73 years, SD_{age} = 1.55), composed of 11 boys (M_{age} = 10.64 years, SD_{age} = 1.63) and 11 girls (M_{age} = 10.82 years, SD_{age} = 1.54), aged between 9 and 13 years, and attending regular classes, was recruited from one elementary and one middle school located in Southern France. This age bracket was chosen in order to develop a questionnaire that is easily understandable and accessible to a wider audience comprising older children, "normal" adolescents, as well as adolescents with reading difficulties. Indeed, sometimes young adolescents with reading difficulties present a comprehension level similar to older children (Harter, 1999). This sample completed the preliminary version of the FNAES in standardized conditions (i.e., isolation, quiet classroom conditions, and groups of up to 10 students with help in reading if necessary). The original response format was replaced by a 5-point Likert scale (i.e., from 1 = *not at all clear* to 5 = *completely clear*) to assess item clarity. Following the completion of the questionnaires, individual interviews were used to investigate how unclear items could be clarified.

Study 2

A sample of 507 adolescents (M_{age} = 14.87 years, SD_{age} = 1.98; M_{BMI} = 19.99, SD_{BMI} = 2.82), composed of 303 boys (M_{age} = 15.68 years, SD_{age} = 2.25; M_{BMI} = 20.42, SD_{BMI} = 2.87) and 204 girls (M_{age} = 14.80 years, SD_{age} = 2.24; M_{BMI} = 19.35, SD_{BMI} = 2.61), aged between 11 and 18 years and attending regular classes, was recruited from four middle and high schools in Southern France. This sample completed the adolescent version of the FNAES in the same standardized conditions as in Study 1. In addition, 23 of those (M_{age} = 16.57 years, SD_{age} = 0.95; M_{BMI} = 20.01, SD_{BMI} = 2.38), comprising 11 boys (M_{age} = 16.64 years, SD_{age} = 1.03; M_{BMI} = 20.47, SD_{BMI} = 2.50) and 12 girls (M_{age} = 16.50 years, SD_{age} = 0.90; M_{BMI} = 19.60, SD_{BMI} = 2.29), were retested after two weeks.

Study 3

A sample of 155 adolescents (M_{age} = 15.34 years, SD_{age} = 2.29; M_{BMI} = 19.99 SD_{BMI} = 2.82), composed of 96 boys (M_{age} = 15.68 years, SD_{age} = 2.25; M_{BMI} = 20.85, SD_{BMI} = 3.14) and 59 girls (M_{age} = 14.80 years, SD_{age} = 2.24; M_{BMI} = 19.40, SD_{BMI} = 2.30), aged between 11 and 18 years and attending regular classes, was recruited from two middle and high schools in Southern France. This sample completed the FNAES and the French versions of the Rosenberg self-esteem inventory (Rosenberg, 1965), FNES (Watson & Friend, 1969), Eating Attitudes Test-26 (EAT-26; Garner, Olmstead, Bohr, & Garfinkel, 1982), and the Social Physique Anxiety Scale (SPAS; Hart, Leary, & Rejeski, 1989).

Measures

Fear of Negative Appearance Evaluation

The final version of the FNAES, taken from the more systematic Lundgren et al.'s (2004) study, was translated into French following the standardized back-translation techniques widely described in the literature (Van de Vrijver & Hambleton, 1996). Translation from English into French was done separately by two bilingual researchers and a bilingual translator. Thereafter, translation discrepancies between the three translated forms were discussed in order to develop an initial French version. A second bilingual translator whose native language was English, and who had not seen the original English version of the FNAES, translated this French version back into English. The back-translated version was then compared with the original English version and inconsistencies, errors, and biases were highlighted. The translation process was repeated until the back-translated versions were equivalent to the original English version. The final version exhibited no discrepancies with the original version when back-translated. As an additional check, the final version was independently reviewed by the translators to confirm that each item had kept its original meaning (Van de Vrijver & Hambleton, 1996). This preliminary French version comprises six items (see Table 1) rated by the participant on a 5-point Likert scale ranging from *not at all* (1) to *extremely* (5). The items are summed to obtain a total score for the FNAES.

Table 1. Items of the fear of negative appearance evaluation scale (English items versions are in *italics* in parentheses)

1. Je suis préoccupé(e) par ce que les gens pensent de mon apparence (*I am concerned about what other people think of my appearance*)
2. Cela me dérange si je sais que quelqu'un juge la forme de mon corps (*It bothers me if I know someone is judging my physical shape*)
3. Cela me tracasse que les gens puissent trouver des défauts à mon apparence (*I worry that people will find fault with the way I look*)
4. Lorsque je rencontre de nouvelles personnes, je me demande ce qu'elles pensent de mon apparence (*When I meet new people, I wonder what they think about my appearance*)
5. J'ai peur que les gens puissent remarquer mes défauts physiques (*I am afraid other people will notice my physical flaws*)
6. J'accorde trop d'importance à ce que les gens pensent de mon apparence (*I think that other people's opinions of my appearance are too important to me*)

Fear of Negative Evaluation

The French version of the FNES (Watson & Friend, 1969; Musa, Kostogianni, & Lépine, 2004) was used to measure apprehension about negative evaluations. This instrument contains 30 items rated on a true-false answer scale. This instrument was used in Study 3.

Eating Disorders

The French version of the EAT-26 (Garner et al., 1982; Leichner, Steiger, Puentes-Neuman, Perreault, & Gottheil, 1994) was used in Study 3 as a self-report inventory to evaluate the presence of disturbed eating attitudes and behaviors. This instrument comprises a global scale and three subscales: dieting, bulimia and food preoccupation, and oral control. The 26 items of this instrument are rated by the participants using a 6-point scale ranging from *always* to *never*. Their answers on the 6-point scale were recoded into a 4-point scale ranging from 0 through 3 in which 0 is assigned to the three responses that represent the least symptomatic answers, and 3 represents the most symptomatic answer (Garner et al., 1982).

Social Physique Anxiety

The French version of the SPAS (Hart et al., 1989) was used in the third study to determine the degree to which people become anxious because of the real or perceived evaluation of their physique by others. The 12 items are rated on a 5-point Likert-type scale ranging from *not at all* (1) to *extremely* (5). This instrument was used in Study 3.

Self-Esteem

The French version of the Rosenberg Self-Esteem Inventory (RSEI; Rosenberg, 1965; Vallières & Vallerand, 1989) was used to assess overall feelings of self-worth or self-acceptance. The 10 items from this instrument are rated on a 4-point Likert scale ranging from *strongly agree* (4) to *strongly disagree* (1). This instrument was used in Study 3.

Data Analysis

Study 1

Analyses of the clarity of the items from the preliminary French version of the FNAES were performed following Vallerand's (1989) recommendations: An item with a clarity score of less than 4 on a 5-point Likert scale was considered unsatisfactory (Vallerand, 1989). Follow-up interviews were then conducted with participants to identify the problems.

Study 2

In this study, analyses were conducted in two stages. In the first stage, a CFA was done on the complete data set to verify the factor structure of the preliminary version of the French FNAES. This CFA model hypothesized that: (1) answers to the FNAES could be explained by one factor, (2) each item would have a nonzero loading on the FNAES factor, and (3) measurement error terms would be uncorrelated. Given the significant multivariate nonnormality of the data (normalized coefficient values for kurtosis: 15.489), the analyses were performed using bootstrapped maximum likelihood (ML) estimation with AMOS 7.0 (Arbuckle, 2006). All fit indices are, thus, based on the Bollen-Stine bootstrap *p*-value and bootstrap adjusted χ^2 and goodness-of-fit indexes (Yuan & Hayashi, 2003). In case of inadequate fit for the initially hypothesized model, modifications to the CFA model were performed based on analyses of items': intercorrelations, factor loadings, square multiple correlations, *t* values, and modification indices. The CFA was then reproduced to determine whether the modification resulted in an improved fit. This process was continued until a reasonable fit was obtained for the model. Finally, the temporal stability of the resulting questionnaire was also estimated using test/retest reliability correlations on the data from the 23 adolescents who were retested after 2 weeks.

In the second stage, the French version of the FNAES was used to test the measurement and latent mean invariance of the CFA model across gender. CFA models were first estimated separately in all gender-related subsamples and then measurement invariance tests across gender were performed in the sequential order recommended by Meredith (1993), always using the preceding model as the reference for model comparisons.

Assessment of fit of the CFA models was based on multiple indicators: χ^2, comparative fit index (CFI), the Tucker-Lewis index (TLI), the standardized root mean square residual (SRMR), the root mean square error of approximation (RMSEA), and the 90% confidence interval of the RMSEA (RMSEA 90% CI). Values greater than .90 for CFI and TLI are considered to be indicative of adequate model fit, although values greater than .95 are preferable (Byrne, 2005; Hu & Bentler, 1999). Values smaller than .08 or .05 for the RMSEA and smaller than .10 and .08 for the SRMR support, respectively, acceptable and good model fits (Hu & Bentler, 1999; Vandenberg & Lance, 2000). Critical values for the tests of gender and multigroup measurement invariance in CFA models were evaluated relative to several criteria: χ^2 difference tests, and CFI and RMSEA changes resulting from the application of additional invariance constraints (Chen, 2007; Cheung & Rensvold, 2002; Vandenberg & Lance, 2000). A CFI difference of .01 or less and RMSEA differences of .015 or less between a baseline model and the resulting model indicate a lack of measurement variance. Finally, reliability was computed from the model's standardized parameters estimate, using

Table 2. Goodness-of-fit statistics of FNAES models

Study	No. items	Model	Description	χ^2	*df*	CFI	TLI	SRMR	RMSEA	RMSEA 90% CI	$\Delta\chi^2$	Δdf	|ΔCFI|	|ΔRMSEA|
Study 2[a]	6	CFA	1-factor	13.892*	9	.977	.961	.036	.085	.060–.112				
	5	CFA	1-factor	8.358	5	.996	.992	.019	.029	.000–.084				
	5	CFA, invariance tests	Boys (n = 303)	8.218	5	.993	.980	.027	.055	.000–.018				
			Girls (n = 204)	6.337	5	1.00	.999	.021	.016	.000–.099				
			A – No invariance	14.656	10	.996	.992	.027	.031	.000–.062				
			B – λs invariant	18.709	14	.990	.985	.036	.041	.014–.065	4.053	4	.006	.010
			C – λs, τs invariant	22.756*	18	.976	.976	.042	.052	.032–.072	4.047	4	.014	.011
			C' – λs, τs (τ_1 free) invariant	21.803	17	.984	.981	.037	.047	.025–.068	3.094	3	.006	.006
			D – λs, τs (τ_1 free), δs invariant	29.744*	22	.976	.979	.042	.050	.031–.068	7.941	5	.008	.003
			E – λs, τs (τ_1 free), δs, ξ invariant	30.730*	23	.972	.976	.045	.052	.035–.070	0.986	1	.004	.002
			F – λs, τs (τ_1 free), δs, ξ, η invariant	31.444*	24	.950	.959	.052	.069	.053–.086	0.714	1	.022	.017
Study 3[b]	5	CFA	1-factor	7.651	5	.968	.936	.048	.145	.085–.211				
	5	CFA, invariance tests	Boys (n = 96)	7.317	5	.991	.982	.034	.073	.000–.172				
			Girls (n = 59)	7.931	5	.947	.894	.034	.207	.107–.317				
			A – No invariance	14.103	10	.971	.942	.034	.099	.051–.148				
			B – λs invariant	17.992	14	.963	.947	.046	.094	.053–.136	3.889	4	.008	.005
			C – λs, τs invariant	22.300*	18	.943	.937	.053	.103	.068–.139	4.308	4	.020	.009
			C' – λs, τs (τ_1 free) invariant	21.263*	17	.953	.945	.054	.097	.059–.135	3.271	3	.010	.003
			D – λs, τs (τ_1 free), δs invariant	28.247*	22	.921	.929	.060	.110	.078–.142	6.984	5	.032	.013
			D' – λs, τs (τ_1 free), δs ($\delta_{1,3}$ free) invariant	24.607*	20	.944	.944	.051	.097	.062–.132	3.344	3	.009	.000
			E – λs, τs (τ_1 free), δs ($\delta_{1,3}$ free), ξ invariant	25.262*	21	.945	.948	.057	.094	.060–.128	1.331	1	.001	.003
			F – λs, τs (τ_1 free), δs ($\delta_{1,3}$ free), η invariant	30.659*	22	.906	.925	.062	.113	.083–.143	5.397*	1	.039	.019
Study 2–3[c] comparisons	5	CFA, invariance tests	A – No invariance	15.375*	10	.988	.975	.019	.057	.035–.080				
			B – λs invariant	19.420	14	.989	.984	.020	.045	.025–.066	4.05	4	.001	.012
			C – λs, τs invariant	23.037*	18	.988	.987	.020	.042	.023–.060	3.62	4	.001	.003
			D – λs, τs, δs invariant	31.064	23	.988	.990	.020	.037	.019–.053	8.027	5	.000	.005
			E – λs, τs, δs, ξ invariant	31.838*	24	.987	.989	.028	.037	.021–.053	0.774	1	.001	.000
			F – λs, τs, δs, ξ, η invariant	32.949*	25	.986	.989	.029	.038	.022–.054	1.111	1	.001	.001

Note. CFA: confirmatory factor analytic model; χ^2: chi-square; *df*: degrees of freedom; CFI: comparative fit index; TLI: Tucker-Lewis index; RMSEA: root mean square error of approximation; RMSEA 90% CI = 90% confidence interval for the RMSEA point estimate; SRMR: standardized root mean square residual; λ: factor loading; τ: intercept; δ: uniquenesses; ξ: factor variance; η: factor mean; $\Delta\chi^2$: change in goodness-of-fit χ^2 relative to the preceding model; Δdf: change in degrees of freedom relative to the preceding model; ΔCFI: change in comparative fit index relative to the preceding model; ΔRMSEA: change in root mean square error of approximation relative to the preceding model; *$p < .05$, [a]n = 507, [b]n = 155, [c]n = 662.

the following formula (Bagozzi and Kimmel, 1995): $\rho = (\Sigma\lambda i)^2/([\Sigma\lambda i]^2 + \Sigma\delta ii)$ where λi are the factor loading and δii the error variances.

Study 3

In this study, CFA analyses of the FNAES model developed in Study 2 were performed in four stages using bootstrapped ML estimation with AMOS 7.0 for nonnormal data (normalized coefficient values for kurtosis: 9.750). In the first stage, the CFA model was applied to this sample to cross-validate the factor structure of the FNAES obtained in the preceding study. In the second stage, the measurement and latent mean invariance of this model were verified across gender and samples from Study 2 and 3, following the same aforementioned procedures. Finally, in the fourth stage the convergent validity of the resulting version of the FNAES was evaluated through correlations with measurements of self-esteem (RSEI), fear of negative evaluation (FNES), eating disorders (EAT-26) and social physique anxiety (SPAS). A Bonferroni correction was applied (alpha error was, thus, set at .05/4 = .01).

Results

Study 1: Item Content Clarity of the Preliminary Version of the French FNAES

Items from the French and English version of the FNAES are reported in Table 1. Analyses of item clarity showed that all items were considered to be satisfactory using the cutoff criteria proposed by Vallerand (1989). In fact, the results ranged from $M = 4.36$ ($SD = 0.58$) for Item 4 to $M = 4.77$ ($SD = 0.43$) for Item 5. This first study, thus, provided support for the appropriateness of the translated items for younger and older adolescents.

Study 2: Factor Validity of the Preliminary Version of the French FNAES

Stage 1

The goodness-of-fit statistics of the CFA models of the FNAES are displayed in Table 2. The results showed that the estimated CFA model exhibited (Table 2) significant bootstrapped χ^2 values, CFI and TLI exceeding .95, and SRMR indicator under .05. However, the RMSEA was over the .08 criteria. Examination of the specific results revealed that Item 4 was highly correlated with other items and presented elevated modification indices (i.e., correlated uniqueness). This item was removed from the original version of the FNAES and a CFA was re-estimated based on the remaining five items. As illustrated in Table 2, this reduced model presented an optimal level of fit to the data (CFI, TLI > .95, SRMR < .02, RMSEA < .03) and was characterized by significant and substantial loadings (i.e., ranging from .715 to .810). These results support the factor validity of the alternative five-item measurement model for French adolescents. This version exhibited a scale mean score of 10.38 ($SD = 4.57$), an acceptable reliability coefficient ($\rho = .83$), and a satisfactory test/retest reliability correlation coefficient ($r^{tt} = .77$).

Stage 2

The results from the measurement and latent mean invariance tests across gender are reported in Table 2. These analyses showed that the five-item model performed relatively well in the separate samples of boys and girls. The first and second steps of invariance testing (i.e., Hypotheses A and B) resulted in nonsignificant bootstrap χ^2, acceptable goodness-of-fit indices (i.e., CFI, TLI > .95; SRMR, RMSEA < .05), nonsignificant $\Delta\chi^2$, and ΔCFIs and ΔRMSEAs that did not exceed .01 and .015, respectively. The third level of measurement invariance (i.e., Hypothesis C) provided significant bootstrap χ^2, acceptable goodness-of-fit indices (i.e., CFI, TLI > .95; SRMR, RMSEA < .06) a nonsignificant $\Delta\chi^2$, and a ΔRMSEA that did not exceed .015. However, this model resulted in a ΔCFI exceeding .01 and modification indices suggesting that the gender-group equality constraint for the intercept of Item 1 should be relaxed. The fourth model (i.e., Hypothesis C') freely estimated the parameter across gender while maintaining the other constraints and provided evidence of partial strong measurement invariance (i.e., nonsignificant $\Delta\chi^2$; ΔCFI < .01; ΔRMSEA < .015). The fifth model (i.e., Hypothesis D) added the equality constraints on items' uniqueness. These results support the measurement invariance of the items' uniqueness across gender (i.e., nonsignificant $\Delta\chi^2$; ΔCFI < .01; ΔRMSEA < .015). The sixth model (i.e., Hypothesis E) supports the invariance of the latent factor variance across groups (i.e., nonsignificant $\Delta\chi^2$; ΔCFI < .01; ΔRMSEA < .015). Finally, the last model (i.e., Hypothesis F) revealed that the latent factor mean was gender variant (ΔCFIs .02): Girls (latent mean = .39, $t = 4.93$, $p = .001$, $d = 0.45$) score significantly higher than boys (latent mean fixed at zero) on the FNAES.

Study 3: Cross-Validation and Convergent Validity of the French FNAES

Stage 1

The CFA models of the five-item version of the FNAES (Table 2) yielded results highly similar to those found in the preceding study and showed acceptable goodness-of-fit indices (i.e., CFI, TFLI > .95; SRMR < .05; significant loadings ranging from .744 to .859), with an exception for the RMSEA, which was higher than .10. It should be noted

that RMSEA values tend to be inflated in small sample sizes (Curran, Bollen, Chen, Paxton, & Kirby, 2003) and that this study's sample (n = 155) is much smaller than the one from the preceding study (n = 507). Examination of the modification indices revealed no significant cross-loadings or correlated uniqueness, further supporting its adequacy. The five-item version exhibited a scale mean score of 11.32 (SD = 5.37) and an acceptable reliability coefficient (ρ = .83). These results were similar to those found in Study 2 and cross-validated this version on a second independent sample.

Stage 2

The results from gender-based measurement invariance tests are reported in Table 2. They paralleled, in most cases, those found in the second stage of Study 2 and showed that the FNAES model and factor loadings overlapped across gender and that the item intercepts and uniqueness were partially invariant across gender. Again, most of the model noninvariance could be attributed to the first item. Finally, the results from the latent means invariance test showed that girls (latent mean = .38, t = 2.568, p < .01, d = 0.43) presented significantly higher scores than boys (latent mean fixed to 0) on the FNAES.

Stage 3

The results from the multiple-group measurement and latent mean invariance tests done on results from Studies 2 and 3 are reported in Table 2. These results showed that: (1) χ^2 tests were significant for most models, with an exception for Hypothesis B and D (i.e., invariant loadings and uniqueness); (2) χ^2 difference tests were nonsignificant; (3) the CFIs, TLIs, SRMRs, and RMSEAs values all indicate adequate model fit; and (4) ΔCFIs and ΔRMSEAs did not exceed .010 and .015, respectively. These results suggest that the measurement model of the five-item version fully overlapped between samples from Studies 2 and 3.

Stage 4

First, the internal consistencies of the different instruments used in this study were acceptable (i.e., EAT: Cronbach's α = .83; FNES: α = .86; RSE: α = .84; SPAS: α = .83). Second, the five-item version of the FNAES was positively and significantly correlated with the global scale score of the EAT-26 (r = .47, p < .001), the SPAS (r = .45, p < .001), and the FNES (r = .45, p < .001). A significant negative correlation between the FNAES and the RSEI (r = –.31, p < . 001) was also observed. These results confirm the convergent-related validity of the French five-item version of the FNAES for adolescents.

Discussion

The objectives of the first study were to develop a preliminary French version of the FNAES for adolescents and to verify the content clarity of the items in a sample of adolescents. Results supported that the translated items were successfully understood and that the vocabulary of the French FNAES was suitable for community samples of adolescents.

The objectives of the second and third studies, performed within two independent samples, were to: (1) examine the factor validity and reliability of the FNAES in adolescents, (2) assess the measurement and latent mean invariance of the FNAES across gender and samples, and (3) assess the convergent validity of the FNAES. The present findings first indicated the inadequacy of the Lundgren et al. (2004) model with a community sample of adolescents. That model was carefully examined and one item (Item 4) was excluded from further analyses. The resulting five-item unidimensional model fit the data well in the samples from Study 2 and Study 3 and proved to be perfectly invariant (factor structure, factor loadings, item intercepts, item uniqueness, factor variance, and factor mean) across both samples. These results offset the possibility that the satisfactory results from Study 2 could have been the result of capitalization on chance (i.e., overfitting the model to a single sample). The results also confirmed the fact that this model possesses satisfactory internal consistency coefficients and test/retest reliability correlation coefficients. Clearly, these results provide strong evidence in favor of the FNAES construct validity. Researchers can, thus, be quite confident in their use of this instrument among nonclinical French adolescents.

Subsequent CFAs analyses were performed on the two independent samples with the five-item version of the FNAES to verify the measurement and latent means invariance of the model across gender. First, although none of the previous studies (Lundgren et al., 2004; Thomas et al., 1998) verified the measurement invariance of the FNAES across gender – which is problematic given the known gender difference regarding concerns about physical appearance – the present findings partially support the measurement invariance of the five-item model of the FNAES across gender and age groups. Indeed, the results revealed that the intercept and uniqueness of Item 1 lacked invariance between boys and girls. This suggests that the average level of boys' and girls' answers differs on these items and that these differences are independent of the observed differences in the overall FNAES factor score. This also suggests that this item have measurement error that differ across gender. A more refined inspection of this result revealed that boys presented a higher intercept and a lower uniqueness than girls on this item. It is interesting to note that this item is also the one with the lowest factor loading in the final FNAES measurement models. This evokes the possibility that this item may also tap into another construct

in addition to FNAES and that this construct (e.g., social dependency and desirability, etc.) may differ across gender. Clearly, the present results provide evidence of the construct similarity of a four-item version of FNAES across gender and suggest that meaningful comparisons can be made across gender, at least on the basis of the four invariant items. Although the uniqueness associated with Item 3 also proved to be noninvariant across gender in Study 3, full uniqueness invariance has never been considered as a prerequisite for meaningful group comparisons (Vandenberg & Lance, 2000).

In a related way, and consistent with what is known of gender differences regarding concerns about physical appearance and body image, the present results showed that the FNAES latent mean differed across gender, with girls presenting higher levels than boys. This clearly supports the construct validity of the FNAES. It is interesting to note that this gender-based difference in the latent mean of the FNAES could also be observed in the context of additional analyses relying on the invariant items in the context of both studies.

Third, the convergent validity of the five-item FNAES was also directly evaluated in the third study. Moderate correlations between the five-item FNAES model and convergent measures of fear of negative evaluation, eating disorders, self-esteem, and social physique anxiety were observed. These findings were in the expected range and very similar to those found in previous studies, lending support to the convergent validity of the five-item FNAES (Lundgren et al., 2004; Thomas et al., 1998). Nevertheless, the results from the correlation between the FNAES and the FNES found in this study (r = .45) and Lundgren et al. (2004) (r = .78) are of a different magnitude and remain insufficient to conclude whether the FNAE is simply a more specific manifestation of FNE or a separate and distinct construct. Another issue that remains unsolved is the role of "objective" physical appearance in FNAE: Could FNAE be simply a manifestation of FNE in people who have reasons to fear such evaluations? Therefore, future research might seek to assess the correlation between the FNAES and physical attractiveness (as judged by others) to test whether or not, and to what degree, FNAE is simply a reflexion of "objectively" measured qualities.

Three limitations of the current series of studies must be taken into account when interpreting these findings. First, the factor structure and measurement invariance analyses of the French FNAES were based on a mixed (boy and girl) sample of nonclinical and normally achieving adolescents and late-adolescents, which might not be considered representative of the French adolescent population. This indicates that the use of this instrument should be limited to samples similar to this one. Therefore, examining the factor structure and measurement invariance of the French FNAES across a more diverse sample of adolescents should be a future research priority. Such research should be performed using various clinical samples of adolescents (e.g., adolescents with anorexia nervosa, bulimia nervosa, or social phobias) as well as samples from other cultural or linguistic backgrounds.

Second, the reliance on a cross-sectional sample also precludes the verification of the developmental stability or change of the FNAES. Although the present study allowed for the verification of the 2-week test/retest reliability of the instrument, a complete test of its construct validity would involve testing the developmental continuity and changes throughout the adolescent years. This issue should clearly be addressed in the context of longitudinal studies.

Third, the discriminant validity of the French FNAES was not tested in this series of studies. It is, thus, still unknown whether this instrument could discriminate adolescents with high levels of FNAE (e.g., adolescents with anorexia nervosa, bulimia nervosa, or social phobias) from normal controls. Therefore, as a consequence of these limitations of examining this discriminant property of the instrument, using clinical samples and establishing cutoff scores should be a future research priority.

In conclusion, the psychometric properties of the French FNAES were tested within two independent and heterogeneous adolescent samples and found to be adequate. This instrument may, thus, be used in research and practice to assess FNAE in nonclinical populations of French adolescents. However, regarding the multiple limitations of these studies, it would be premature to recommend the use of the FNAES in clinical samples.

Authors' Note

The first two authors contributed equally to the preparation of this paper. Their order of appearance was determined at random: both should be considered first authors. This manuscript was prepared while A.J.S. Morin was a visiting scholar at the University of Aix-Marseille II, France.

References

Arbuckle, J. (2006). *AMOS 7.0 user's guide*. Chicago, IL: SPSS.

Bagozzi, R.P., & Kimmel, S.K. (1995). A comparison of leading theories for the prediction of goal directed behaviors. *British Journal of Social Psychology, 34*, 437–461.

Benjet, C., & Hernández-Guzmán, L. (2002). A short-term longitudinal study of pubertal change, gender, and psychological well-being of Mexican early adolescents. *Journal of Youth and Adolescence, 31*, 429–442.

Byrne, B.M. (2005). Factor analytic models: Viewing the structure of an assessment instrument from three perspectives. *Journal of Personality Assessment, 85*, 17–32.

Chen, F.F. (2007). Sensitivity of goodness-of-fit indexes to lack of measurement. *Structural Equation Modeling, 14*, 464–504.

Cheung, G.W., & Rensvold, R.B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255.

Curran, P.J., Bollen, K.A., Chen, F., Paxton, P., & Kirby, J.B.

(2003). Finite sampling properties of the point estimates and confidence intervals of the RMSEA. *Sociological Research and Methods, 32*, 208–252.

Garner, D.M., Olmstead, M.P., Bohr, Y., & Garfinkel, P.E. (1982). The Eating Attitude Test: Psychometric features and clinical correlates. *Psychological Medicine, 12*, 871–878.

Hart, E.H., Leary, M.R., & Rejeski, W.J. (1989). The measurement of social physique anxiety. *Journal of Sport and Exercise Psychology, 11*, 94–104.

Harter, S. (1999). *The construction of the self: A developmental perspective*. New York: Guilford.

Hu, L.-T., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.

Klomsten, A.T., Shaalvik, E.M., & Espnes, G.A. (2004). Physical self-concept and sports: Do gender differences still exist? *Sex Roles, 50*, 119–127.

Leichner, P., Steiger, H., Puentes-Neuman, G., Perreault, M., & Gottheil, N. (1994). Validation d'une échelle d'attitudes alimentaires auprès d'une population québécoise francophone [Validation of an eating attitude scale in a French-speaking Quebec population]. *Revue Canadienne de Psychiatrie, 39*, 49–54.

Levine, M.P., & Piran, N. (2004). The role of body image in the prevention of eating disorders. *Body Image, 1*, 57–70.

Lundgren, J.D., Anderson, D.A., & Thompson, J.K. (2004). Fear of negative appearance evaluation: Development and evaluation of a new construct for risk-factor work in the field of eating disorders. *Eating Behaviors, 5*, 75–84.

Meredith, W. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika, 58*, 525–543.

Musa, C., Kostogianni, N., & Lépine, J.-P. (2004). Échelle de peur de l'évaluation négative (FNE): Propriétés psychométriques de la traduction française [The Fear of Negative Evaluation scale (FNE): Psychometric properties of the French version]. *Encéphale, 30*, 517–524.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Smolak, L. (2004). Body image in children and adolescents: Where do we go from here? *Body Image, 1*, 15–28.

Thomas, C.M., Keery, H., Williams, R., & Thompson, J.K. (1998, November). *The fear of negative appearance evaluation scale: Development and preliminary validation*. Paper presented at the annual meeting of the Association for the Advancement of Behavior Therapy, Washington, DC.

Thompson, J.K., & Stice, E. (2001). Thin-ideal internalization: Mounting evidence for a new risk factor for body-image disturbance and eating pathology. *Current Directions in Psychological Science, 10*, 181–183.

Vallerand, R.J. (1989). Vers une méthodologie de validation transculturelle de questionnaires psychologiques: Implications pour la recherche en langue française [Towards a methodology of transcultural validation of psychological questionnaires: Implications for research in the French language]. *Canadian Psychology, 4*, 662–680.

Vallières, E.F., & Vallerand, R.J. (1990). Traduction et validation canadienne-française de l'échelle de l'estime de soi Rosenberg [Canadian French-speaking translation and validation of the Ronsenberg Self-Esteem Inventory]. *International Journal of Psychology, 25*, 305–316.

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*, 4–70.

Van de Vrijver, F.J.R., & Hambleton, R.K. (1996). Translating tests: Some practical guidelines. *European Psychologist, 1*, 89–99.

Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology, 33*, 448–457.

Yuan, K.-H., & Hayashi, K. (2003). Bootstrap approach to inference and power analysis based on three statistics for covariance structure models. *British Journal of Mathematical and Statistical Psychology, 56*, 93–110.

Christophe Maïano

UMR 6233 "Institute of Movement Sciences, Etienne-Jules Marey"
CNRS-University of Aix-Marseille II
163 Avenue de Luminy
CP 910
F-13288 Marseille Cedex 9
France
Tel. +33 491 759-653
Fax +33 491 170-415
E-mail christophe.maiano@univmed.fr

