Original Article

A Confirmatory Factor Analysis of the Structure of Statistics Anxiety Measure: An examination of four alternative models

Shahram vahedi, PhD[1]
Farahman Farrokhi, PhD[2]
Hossein Bevrani, PhD[3]

1 Department of Educational Psychology, Faculty of Education and Psychology, University of Tabriz, Tabriz, Iran
2 English Departments, University of Tabriz
3 Department of Statistics, Faculty of mathematical sciences, university of Tabriz

Corresponding author:
Shahram Vahedi, PhD
Assistant Professor Department of Psychology, Department of Psychology, 29 bahman, The Tabriz University, Iran.
Tel: 98 411-3392090
Fax: 4113356009
Email: vahedi117@ yahoo.com

Objective: The aim of this study is to explore the confirmatory factor analysis results of the Persian adaptation of Statistics Anxiety Measure (SAM), proposed by Earp.
Method: The validity and reliability assessments of the scale were performed on 298 college students chosen randomly from Tabriz University in Iran. Confirmatory factor analysis (CFA) was carried out to determine the factor structures of the Persian adaptation of SAM.
Results: As expected, the second order model provided a better fit to the data than the three alternative models.
Conclusions: Hence, SAM provides an equally valid measure for use among college students. The study both expands and adds support to the existing body of math anxiety literature.

Keywords: *Anxiety, Psychometrics, Questionnaire, Statistical factor analysis*

Iran J Psychiatry 2011; 6:92-98

For many years, psychologists have been interested in finding variables that can predict academic performance (AP). In recent years, research on the relationships between personality and AP has not only analyzed the general relationships between the two variables but has also focused on the relationships between anxiety and performance in specific academic domains. As a result, several authors have investigated the predictive power of personality on performance in statistics courses.

It has been estimated that as many as 80% of graduate students experience uncomfortable levels of statistics anxiety, and statistics examinations are more anxiety-inducing than other types of examinations (1). Statistics anxiety may even hinder a student from completing a degree or deter a talented student from thinking about a career as a professor (2). Identifying individuals suffering from statistics anxiety and gaining a better understanding of the domains that contribute to such anxiety is a start to addressing the problem of statistical illiteracy today.

Statistics anxiety has been defined as anxiety that occurs because of encountering statistics in any form and at any level, involving a complex array of emotional reactions (apprehension, fear, nervousness, panic, and worry) that hinder the learning process (3).

oreover, statistics anxiety is situation-specific, inasmuch as the symptoms only emerge at a particular time and in a particular situation–specifically, when learning or applying statistics in a formal setting (4 and 2).

Research indicates that statistics anxiety is a multidimensional construct (5, 6 and 2). Using factor analysis, Earp (7) identified five components of statistics anxiety, namely: (a) anxiety, (b) performance, (c) attitude towards class, (d) attitude towards math, and (e) fearful behavior.

A growing body of research has documented a consistent negative relationship between statistics anxiety and course performance (8). In fact, statistics anxiety has been found to be the best predictor of achievement in research methods courses (9) and statistics courses (10). Moreover, a causal link between statistics anxiety and course achievement has been established. In particular, Onwuegbuzie and Seaman (11) found that graduate students with high levels of statistics test anxiety who were randomly assigned to a statistics examination which was administered under timed conditions tended to have lower levels of performance than did their low anxious counterparts who took the same test under untimed conditions.

Earp (7) established an instrument named 'SAM' to measure Statistics Anxiety in a community college. 'SAM' had high internal consistency reliability

(Cronbach's coefficient alpha = 0.82-0.95) and construct validity. SAM needs to be more adequately validated because counselors have used it extensively. Based on exploratory factor analysis (EFA) in Earp (7), in this study, we tested four models. Our research questions were as follows:

Do statistics anxiety items generated to reflect the five identified domains (Anxiety, Performance, Attitude Towards Class, Attitude Towards Math, and Fearful Behavior factor) fit appropriately into the five domains?

Answering this question required the evaluation of the fit of measurement models to SAM data.

Does the Statistics Anxiety Measure provide adequate evidence for reliability and validity? Estimation of reliability was performed under the framework of confirmatory factor analysis (CFA).

Materials and Method

Participants

The SAM was administered to 300 undergraduate students (133 males and 165 females) chosen from different disciplines of human sciences at Tabriz University in Iran who enrolled in entry-level statistic courses and voluntarily participated in the study. The sample consisted of 133 males and 165 females. College research examination Board approved the research protocol.

Assessment Measures

Statistics Anxiety Measure (7): The 43 items of this scale are rated on a 5-point scale ranging from 1= strongly disagree to 7= strongly agree (higher scores reflect greater statistics anxiety; see Appendix A for the list of items). SAM comprises of five discrete subscales: Anxiety, Performance, Attitude towards class, Attitude towards math, and Fearful behavior. The English versions of the scale show a multidimensional structure for students, and have good construct, and discriminate validity (7). The internal consistent reliability of the overall scale ($\alpha = 0.93$) as well as sub-scales generally ranged from high to excellent ($\alpha = .82–.95$).

The Persian version of the SAM was developed using the standard back-translation technique (12).The first author initially translated the SAM into Persian, and an independent translator unaffiliated with the study then translated this version back into English. Minor differences that emerged during this process were resolved between the translators.

Procedure

All participants were recruited opportunistically using a cluster-sampling technique initiated by three data collectors. All participants completed paper-and-pencil versions of the questionnaire anonymously, and returned the questionnaires to their contact person. All data were treated confidentially, and participants were provided with a debrief sheet following completion. All participants took part on a voluntary basis and were not remunerated for participation.

Data analysis

The analyses addressed two main questions. First, which existing factor structure (one, four and five factor structures) provides an acceptable measurement model for the 44-item SAM? To address this question, CFA was used to impose each of the three factor structures on two data sets to evaluate each model's goodness-of-fit. Second, is there measurement invariance with respect to gender? To address this question, multigroup CFA was used to test hypotheses about the invariance of the 41-item SAM across males and females. One-way analysis of variance (ANOVA) was also used to compare gender differences on the subscale of SAM. Data was analyzed using PASW Statistic18 and AMOS 16 (13 and 14). PASW was used to analyze descriptive statistics and the reliability of the SAM. AMOS was used to perform the CFAs of the SAM analyzing the fit of models and its respective parameter estimates in two distinct stages.

In stage 1, the four models were subjected to a maximum-likelihood CFA using AMOS 16. Model 1 specified a single factor model, the factor being Statistics Anxiety. We eliminated 9 items with non-significant factor loadings. Model 2 specified a correlated four-factor model with six items loading on the Performance factor (items y29-y34), eleven items loading on the Anxiety factor (items y1-y11), eight items loading on the Attitude towards math factor (items y21 to y28), and seven items loading on the Attitude towards class factor (items y12 to y20).

Model 3 specified a correlated five-factor model with seven items loading on the Anxiety factor (items y1-y7), six items loading on the Performance factor (items y29- y34), four items loading on the Fearful behavior factor (items y8 to y11), nine items loading on the attitude towards class factor (items y12 to y20), and eight items loading on the attitude towards math factor (items y21 to y28). Model 4 specified the same factor structure as Model 3 but included a second-order factor labeled Statistics Anxiety. This model was used to determine the existence, or robustness, of the five first-order factors in the presence of a general factor. The models are shown in Figure 1 and 2. 1.

Results

Between-Group Differences

In order to examine possible between-group differences in responses to the SAM, we ran a one-way analysis of variance (ANOVA) with the subscales of SAM (Anxiety factor, Performance factor, Fearful behavior factor, the attitude towards class factor, attitude towards math factor) as the dependent variable and participants' sex as independent variable.

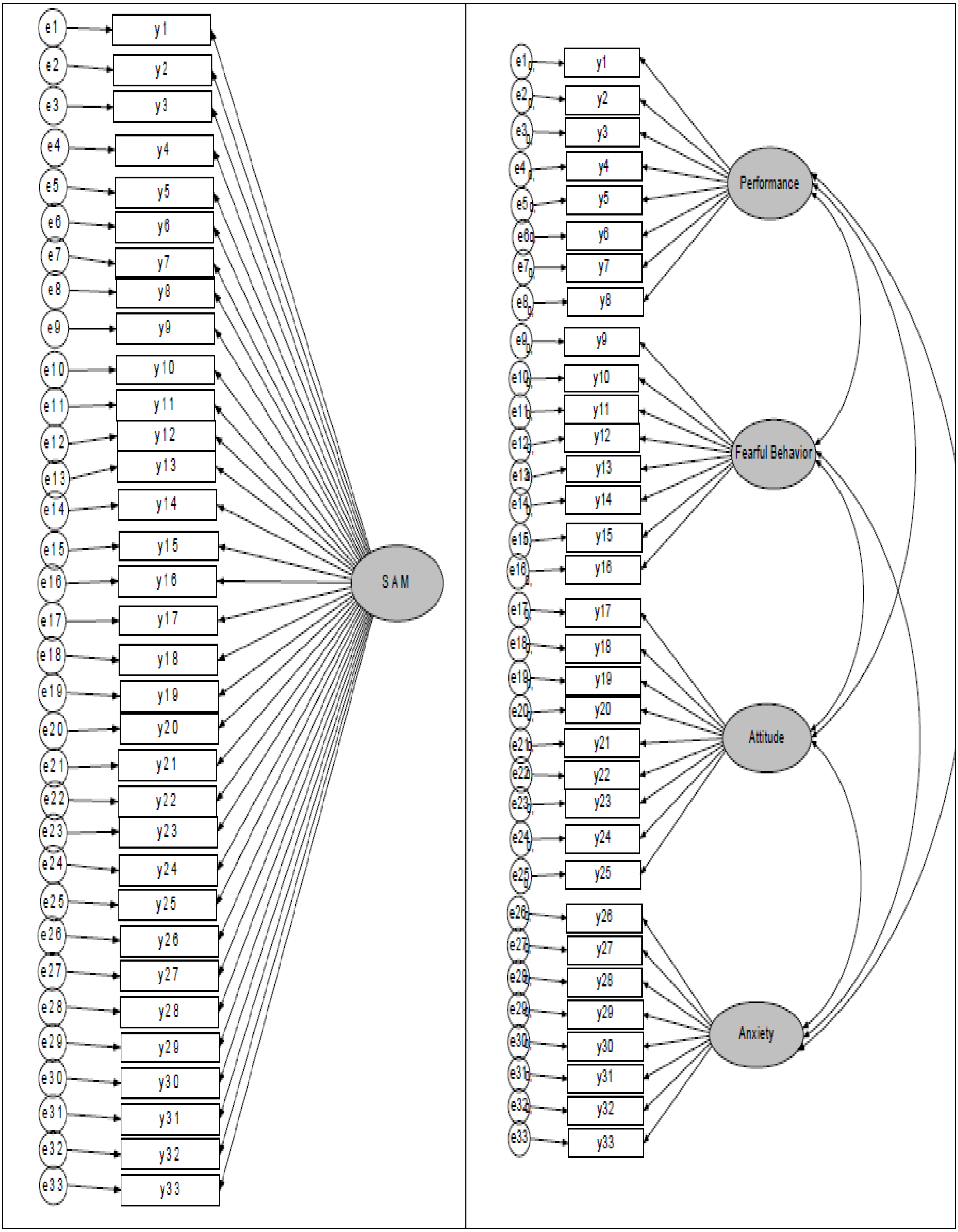

Figure 1. Alternative factor models of Statistics Anxiety Measure: Model 1, Model 2

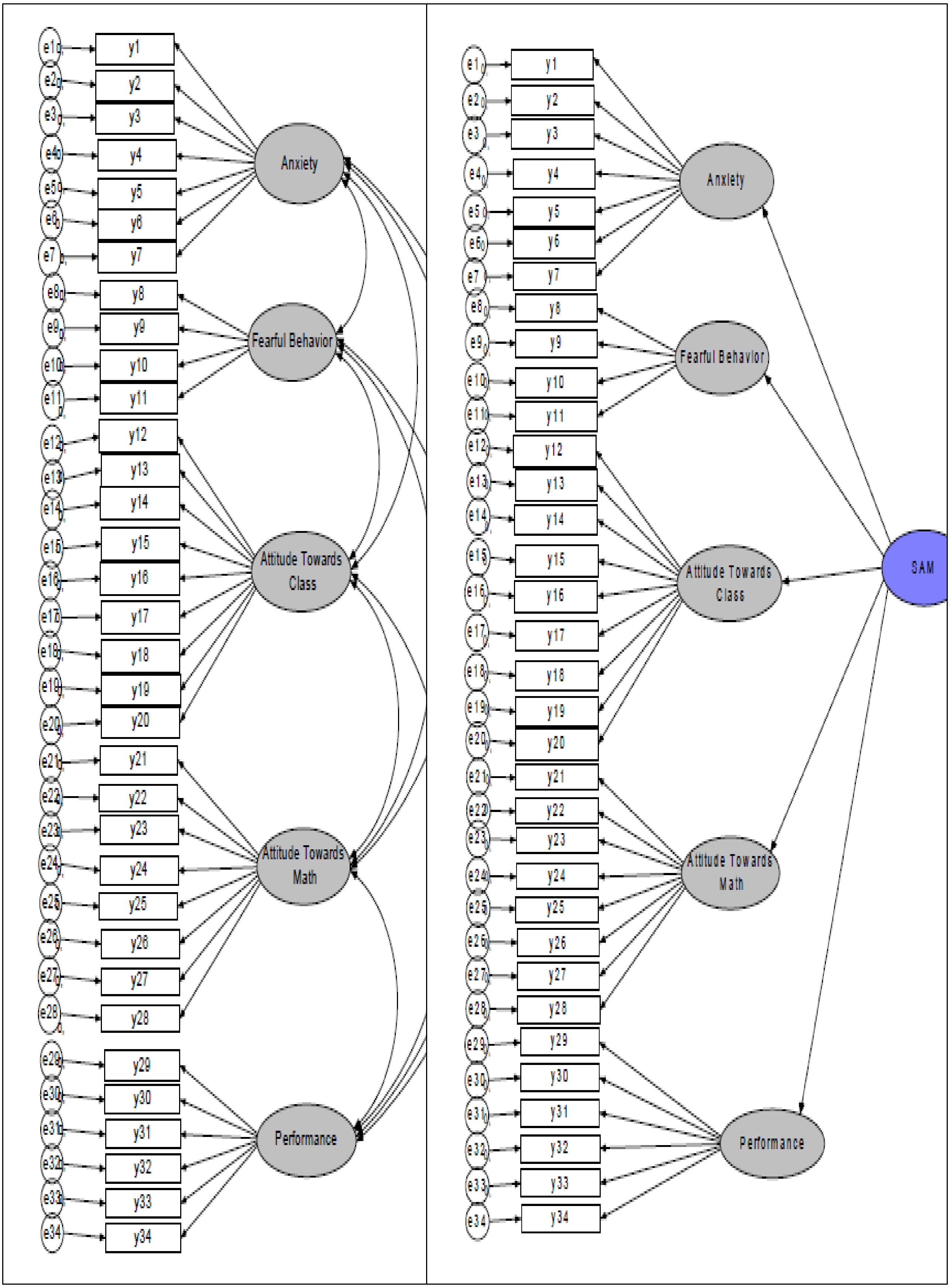

Figure 2. Alternative factor models of Statistics Anxiety Measure: Model 3, Model 4

Published by "Tehran University of Medical Sciences" (www.tums.ac.ir)

Results showed participants' sex was significantly related to the attitude towards class factor, $F (1, 298) =4.45$, $p< .05$) such that female student had lower scores (M=20.52, SD=15.37) than male student (M= 18.75, SD =13.57). However, participants' sex was not significantly related to Anxiety, Performance, Fearful behavior, attitude towards math ($p>.05$). In other words, female students reported more negative attitude towards class than male students.

To evaluate the goodness-of-fit of four alternative measurement models for the SAM, CFA was first run for a one-factor solution in which all 34 items loaded on to a single general strengths factor (Model 1) and subsequently run for other models suggested by Earp (7).

In order to determine which of the four proposed models is the `best' model we can use both statistical criteria and information from the parameter estimates from each of the models. The results of the CFAs for each model are shown in Table 1. In all the analyses, the chi-square goodness of fit statistic is large and significant beyond the 0.001 level, rather than being small and associated with a high probability, which would indicate a close fit between model and data. However, this statistic is sensitive to sample size and does not provide a realistic test of the fit of models (16). In model 1, items y29-y34 did not get hold of a statistical significant loading (.40). Following the removal of these factors loadings, indices for this model improved. The results of the initial estimation of the one factor model did not provide a satisfactory result with a chi-square value of 3473.316 (df =495), which was significant at the $P < .001$ level. Other fit indices revealed a low fit (RMSEA =.14; TLI=.53; CFI=.56; IFI = .56). To justify a unidimensional construct, we compared the four-factor model with a unidimensional model.

In accordance with Bentler and Bonett (15), data from modification indices in models of four, five factors and second order suggested that six paths reflecting covariance be added between error terms to improve the fit of the model that was obtained. These paths involved pairs of items that shared variance from variance accounted for by various factors. These items included "Developing conclusions based on mathematical solutions", "Solving mathematical equations", "Calculating probabilities" and "Developing conclusions based on mathematical solutions"' I do not expect to enjoy this class 'I expect this class to be boring, "My ability to calculate statistics will not affect my chances of getting a job in my chosen field", "Taking this class will have little impact on my life" " I lack motivation to learn or continue learning statistics", "There is no room to be creative in statistics", and "how satisfied do you think your child has felt about looks and appearance". On the basis of the criteria associated with RMSEA and CFI, IFI, TLI, the four, five -factor and second order models provide a better fit of the data than the one-factor model. In other words, they account for more variance than the one-factor model.

We also directly compared the models with the $\Delta\chi2$ statistics. Both statistics directly compare the fit of the models after adjusting for differences in the degrees of freedom. In every case, the $\Delta\chi2$ was significant at .001. These results again strongly support the superiority of the second order model over the one factor and four-factor model; thus, this model was considered optimal. Therefore, Model 4 is preferred to Model 3 because of parsimony; it provides an adequate description of the sample data, and provides a better description than the three alternative models.

Descriptive statistics and factor correlations for second order model are reported in Table 3.

magnitudes of the factor loadings to be equal for male and female students, and the other omitting this invariance constraint .

Table 2 presents the fit statistics for the models. Several fit indices were examined to evaluate the overall fit of each model: $\chi2$ tests the hypothesis that an unconstrained model fits the covariance or correlation matrix as well as the given model; ideally values should not be significant); Comparative Fit Index (CFI; 15; comparison of the hypothesized model with a model in which all correlations among variables are zero, and where values around .90 indicate very good fit; Root-Mean-Square Error of Approximation (RMSEA; values of .08 or below indicate reasonable fit for the

model; Tucker-Lewis index (TLI) and the incremental fit index (IFI), with values close to .95 being indicative of good fit; Akaike Information Criterion (AIC), AIC close to zero reflects good fit and between two AIC measures, the lower one reflects the model with the better fit) (cited in 14)

Table 1. Goodness-of-fit statistics and their Comparisons for four alternative measurement models SAM

Models and Comparisons	χ2	df	χ2/df	CFI	TLI	IFI	RMSEA	AIC	χ2 difference
Model 1	3694.43	527	7.01	.53	.50	.53	.14	3898.431	
Model 2	1539.28	515	2.99	.85	.83	.85	.8	1767.282	
Model 3	972.727	511	1.90	.93	.93	.93	.055	1208.727	
Model 4	982.54	52	1.90	.93	.93	.93	.055	1208.536	
M1–M2									2155.15*
M2– M3									544.55*
M3–M4									9.81*

Note: Model 1= One factor; Model 2=four-factor model; Model 3= five-factor model; Model 4= second order * P < 0.001.

Table2.

Second-order factor loadings from the factor Model 5 of Statistics Anxiety Measure (SAM) (estimates for Model 4 in parenthesis)

			Statistics anxiety	attitude towards math	attitude towards class	anxiety	performance	Fearful behavior
anxiety	<---	Statistics anxiety	.48					
fear	<---	Statistics anxiety	.90					
performance	<---	Statistics anxiety	-.39					
math	<---	Statistics anxiety	.69					
attitude	<---	Statistics anxiety	.90					
Y1	<---	anxiety				.69 (.69)		
Y2	<---	anxiety				.80 (.80)		
Y3	<---	anxiety				.87 (.87)		
Y4	<---	anxiety				.86 (.86)		
Y5	<---	anxiety				.70 (.71)		
Y6	<---	anxiety				.75 (.75)		
Y7	<---	anxiety				.74 (.74)		
Y8	<---	Fearful behavior						.81(.81)
Y9	<---	Fearful behavior						.67(.66)
Y10	<---	Fearful behavior						.75(.76)
Y11	<---	Fearful behavior						.83(.83)
Y12	<---	attitude towards class			.67 (.67)			
Y13	<---	attitude towards class			.69 (.69)			
Y14	<---	attitude towards class			.77 (.77)			
Y15	<---	attitude towards class			.72 (.72)			
Y16	<---	attitude towards class			.74 (.74)			
Y17	<---	attitude towards class			.53(.53)			
Y18	<---	attitude towards class			.77 (.79)			
Y19	<---	attitude towards class			.57 (.58)			
Y20	<---	attitude towards class			.56 (.55)			
Y21	<---	attitude towards math		.73 (.73)				
Y22	<---	attitude towards math		.88 (.88)				
Y23	<---	attitude towards math		.88 (.88)				
Y24	<---	attitude towards math		.88 (.88)				
Y25	<---	attitude towards math		.83 (.83)				
Y26	<---	attitude towards math		.91 (.91)				
Y27	<---	attitude towards math		.83 (.83)				
Y28	<---	attitude towards math		.79 (.79)				
Y29	<---	performance					.51(.51)	
Y30	<---	performance					.59 (.59)	
Y31	<---	performance					.70 (.70)	
Y32	<---	performance					.84 (.85)	
Y33	<---	performance					.67 (.67)	
Y34	<---	performance					.76 (.76)	

Reliability Estimates and factor correlations

Table 3: Factor correlations and Reliability from the second order model of statistic anxiety measure

	1	2	3	4	5	6	α	Mean	Std. Deviation
1) statistic anxiety							.90	84.08	14.65
2) performance	-.39						.86	22.37	4.02
3) fearful behavior	.90	-.35					.85	8	2.67
4) anxiety	.48	-.29	.43				.91	15.86	5.33
5) attitude towards class	.91	-.35	.81	.44			.89	20.22	5.56
6) attitude towards math	.69	-.27	.62	.33	.62		.95	17.63	6.53

All correlations were significant $p < 0.05$.

On the this basis of this, Model 4 should be preferred to Model 3 based on parsimony, ; it provides an adequate description of the sample data, and is provides a better description than the four alternative models. We reported the standardized factor loadings for Models 3 and 4 in Table 2 (estimates for Model 4 in parenthesis). For Models 3 and 4 all the factor loadings are positive, high and statistically significant. The factor correlations for Model 4 are reported in Table 2.

Discussion

The primary purpose of this study was to use confirmatory factor analytic techniques in a sample of young adult college students to explore the fit of the five-factor model of the SAM proposed by Earp. The second aim of the present study was to evaluate the psychometric properties of a Persian version of the SAM and to test measurement invariance across sex.

The present study showed that there was a significant gender difference for attitude towards class factor but not for other subscales of SAM. This is consistent with the research that reported significant gender differences in women who experienced higher levels of statistics anxiety (4 and 11). On the other hand, the other parts of the results of the present study are in support of the previous studies (e.g. 6, 17 and 18) and failed to find such gender differences. Future research is needed to further clarify statistics anxiety-gender relationships.

Confirmatory factor analyses on a validation sample showed inadequate support for either the one-or four-factor model (Performance, Fearful behavior, the attitude and Anxiety factor). Consistent with previous findings in Earp (7), our findings also indicate that a five factor model compared to a unifactorial solution best described statistic anxiety, with separate components such as Anxiety factor, Performance factor, Fearful behavior factor, the attitude towards class factor, attitude towards math factor.

Furthermore, the factor structure also indicated the presence of a higher order general statistic anxiety factor. Examination of the association of the lower order factors and the higher order factor to measures of statistic anxiety symptoms indicated that the lower order factor of Fearful behavior and attitude towards class made the largest unique contribution to the prediction of statistic anxiety measure.

In addition, results showed that the SAM has high internal consistency, with Cronbach are a reaching 0.90. These data are further supported by the parameter estimates of the CFAs, and is generally consistent with previous works showing that SAM has high internal consistency (e.g., 7).

Several limitations of this study are as follows: First, the construct validity of the results reported in this article is mainly derived from the student sample (the University of Tabriz). Further research is necessary to replicate the results scale in other geographical settings to validate the Persian version of SAM. A second limitation is the relatively small sample size. Due to limited sample size, structures found in this study may not hold in future administrations given larger sample sizes.

Acknowledgements

We are very grateful to Mrs. Frahnaz Gahramani for assistance in data collection. Thanks are expressed to all personnel who took part in the research and to those who assisted us at various stages of the study. We also thank Khalil Esmailpor for his helpful comments on an earlier draft of this paper.

References

1. Onwuegbuzie, A. J, Wilson, V. A. Statistics Anxiety: Nature, Etiology, Antecedents, Effects, and Treatments-A Comprehensive Review of Literature. Teaching in Higher Education 2003; 8: 195-209.
2. Onwuegbuzie, A. J, DaRos D, Ryan, J. M. The components of statistics of statistics anxiety: a phenomenological study. Focus on Learning Problems in Mathematics 1997; 19: 11–35.
3. Onwuegbuzie AJ, Daley C. Perfectionism and statistics anxiety. Pers Individ Dif 1999; 26: 1089-1102.
4. Zeidner, M. Statistics and Mathematics Anxiety in Social Science Students: Some Interesting Parallels. Br J Educ Psychol 1991; 61: 319-328.
5. Cruise R J, Wilkins EM. STARS: Statistical Anxiety Rating Scale. Unpublished manuscript, Andrews University, Berrien Springs, MI 1980.
6. Cruise JR, Cash RW, Bolton L D. Development and Validation of an Instrument to Measure Statistical Anxiety, in American Statistical Association Proceedings of the Section on Statistical Education. Washington, D. C: American Statistical Association; 1985.
7. Earp, M. Development and Validation of the Statistics Anxiety Measure. Doctorate of Philosophy Dissertation, Presented to the College of Education University of Denver 2007.
8. Onwuegbuzie AJ. Academic procrastination and statistics anxiety. Assessment and Evaluation in Higher Education 2004; 29: 1-19.
9. Onwuegbuzie AJ, Slate J, Paterson F, Watson M, Schwartz R. Factors associated with underachievement in educational research courses. Research in the Schools 2000; 7: 53- 65.
10. Fitzgerald SM, Jurs S, Hudson LM. A model predicting statistics achievement among graduate students. *Coll Stud J* 1996;30: 361-366.
11. Onwuegbuzie AJ, Seaman M. The effect of time constraints and statistics test anxiety on test performance in a statistics course. J Exp Educ 1995; 63: 115–124.
12. Breslin RW. Back-translation for cross-cultural research. J Cross Cult Psychol 1970; 1:185–216.
13. Arbuckle JL. Amos™ 18 User's Guide. Chicago: Amos Development Corporation; 2009.
14. Byrne BM. Structural equation modeling with AMOS: Basic concepts, applications, and programming, 2th. New York: Routledge; 2010.
15. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. Psychol Bull 1980; 88: 588-606.
16. Byrne BM. Structural equation modeling with AMOS: Basic concepts, applications, and programming. Mahwah: Routledge; 2001.
17. Baloglu M. Individual differences in statistics anxiety among college students. Pers Individ Dif 2003; 34: 855- 865.
18. Baloglu M. An application of structural equation modeling techniques in the prediction of statistics anxiety among college students. Unpublished doctoral dissertation. Texas A&M University-Commerce; 2001.