CROSS-CULTURAL ASSESSMENT OF EATING DISORDERS: PSYCHOMETRIC

PROPERTIES OF A SPANISH VERSION OF THE BULIMIA TEST-REVISED

A Thesis

by

MAYRA N. BERRÍOS-HERNÁNDEZ

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

May 2005

Major Subject: Psychology

CROSS-CULTURAL ASSESSMENT OF EATING DISORDERS: PSYCHOMETRIC

PROPERTIES OF A SPANISH VERSION OF THE BULIMIA TEST-REVISED

A Thesis

by

MAYRA N. BERRÍOS-HERNANDEZ

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Approved as to style and content by:

Antonio Cepeda-Benito	David Gleaves
(Co-Chair of Committee)	(Co-Chair of Committee)
Salvador Hector Ochoa	W. Steve Rholes
(Member)	(Head of Department)

May 2005

Major Subject: Psychology

ABSTRACT

Cross-Cultural Assessment of Eating Disorders: Psychometric

Properties of a Spanish Version of the Bulimia Test-Revised. (May 2005)

Mayra N. Berríos-Hernández, B.A., University of Puerto Rico

Co-Chairs of Advisory Committee: Dr. Antonio Cepeda-Benito
Dr. David Gleaves

The purpose of this study was to examine the psychometric properties of a Spanish version of the Bulimia Test-Revised (BULIT-R). The goal was to test the factor-structure equivalence of the BULIT-R across two samples of college students from two different cultures, Spain and the U.S. Researchers using Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) have reported different model solutions for the factor structure of the BULIT-R: a one-factor model (McCarthy et al., 2002); a four-factor model (Vincent et al., 2002), a five-factor model (Thelen et al., 1991) and a six factor model (Santos, 1996). For any of the two samples, CFA did not support any of the models previously reported in the literature. EFA supported a six and a four factor models for the US and Spanish samples, respectively.

.

ACKNOWLEDGEMENTS

I extend my gratitude to all of the people who have encouraged and supported me throughout the completion of this work. I would like to express heartfelt gratitude and appreciation to the co-chairs and members of my thesis committee. I am extremely grateful to Dr. Antonio Cepeda-Benito and Dr. David Gleaves for their critical insight, time and effort during the course of this project, and for their faith in my abilities and knowledge. I would also like to thank my thesis committee member, Dr. Hector Ochoa, for his disposition in being part of this project. I have truly enjoyed and appreciated all of their unique professional contributions.

I also extend my thanks and appreciation to Pedro Leiva and Sonia Rodriguez for their collaboration, and to my friend Marielisa Maysonet for her sincere and eternal friendship, understanding, encouragement and collaboration in this work. Finally, I want to recognize and thank my parents for their unyielding love, support and encouragement.

TABLE OF CONTENTS

Page

LIST OF TABLES

INTRODUCTION

America's cultural diversity is greater now than ever before, with Hispanics being one of the ethnic groups that is growing most rapidly. According to the Surgeon General Report and the Census of 2002, the Hispanic population has increased over the past decade by 56 % (to 35.5 million people; US Public Health Service Surgeon General, 2002). Census projections estimate that by the year 2050, the number of Hispanics will have increased to 97 million, almost one fourth of the U.S. population (Surgeon General's Report, 2002).

These figures reinforce the importance of conducting cross cultural research and the need to develop mental assessment tools and clinical interventions that are reliable and valid across multiple ethnic groups. The challenge of meeting the demands brought about by the increasing racial and ethnic diversity of the U.S. population will ever be present in our society (Prediger, 1993). Culturally sensitive psychological researchers working with diverse populations have the endeavor to use assessment techniques, data-generating procedures, and standardized instruments with validity, reliability, and measurement equivalence demonstrated across culturally diverse groups (CNPAAEMI, 2000; Helms, 1992; Marin & Marin, 1991; Padilla, 1995; Spengler, 1998). The current study will focus on cross-cultural assessment. The goal is to examine the factor structure of the English (Thelen et al., 1991) and a Spanish version of the Bulimia Test-Revised (BULIT-R).

_____

This thesis follows the style and format of the *Journal of Abnormal Psychology*.

As society has become more attuned to its ethnic and cultural diversity the interest in cross-cultural research in psychology and mental health has also gained prominence. Marsella and Yamada (2000) suggested that a perfect example of the current interest in cross-cultural psychology is the inclusion of the section on culture-bound disorders in the Diagnostic and Statistical Manual of Mental Disorders (4th ed.) (American Psychiatric Association [APA], 1994).

Researchers attempting to understand cross-cultural factors in research have used Etic and emic concepts. Etic refers to aspects of life that are consistent across cultures. Etic findings are universal truths or principles. Emic refers to aspects of life that are different across cultures. Emic findings are culture-specific truths or principles (Harris, 1976). Butcher (1996) points out that in order for cross-cultural science to proceed, it is first necessary to compare different cultures on the same or generally equivalent measures.

Investigators from the U.S. have developed a vast variety of instrument in the mental health area. There is thus the question of whether these instruments that were developed in the U.S., using mostly Caucasian participants, are also valid for use with US minorities or with people from other countries. If this work is not carried out, the practice of using measures not tested for cultural bias will continue to confound both research findings and clinical assessments of minority populations (Hinkle, 1994).

Translation of Instruments

Dana (1993) noted that translations of major assessment instruments are necessary in order to provide equivalence versions of the instruments to individuals

whose primarily language is not English. The translation of a test involves a professional translation of the instructions of the test and the test's items. The initial translation is then translated back to the source language. Separate translators should work independently in the initial and back translations, and then the translations should be compared for semantic equivalence. If necessary, the translation process is repeated until semantic equivalence between the two language versions of the instrument is achieved. Any instrument assessing a psychological construct across culturally different groups has the potential for containing items that lack linguistic equivalence across the groups, even if the instrument was developed within the context of a standard language (Butcher, 1996).

Fridja and Jahoda (1966) posited that translation and adaptation problems may never be resolved to provide perfectly equivalent measures and suggested that the creation of comparable forms of the items across languages might be a more realistic goal. For example, Campbell (1968) argued that there is not one correct translation of any given item into another language because there are a number of possible appropriate phrasings of the item in the target language. Likewise, the version of any item in the target language may have multiple equivalents in the source language. Thus, what it is clear is that the translated version of the test items must accurately convey the meaning of the items employed in original development research for the test (Butcher, 1996). Following content equivalence, the next concern in the use of assessment instruments cross-culturally is that of measurement equivalence (Allen & Walsh, 2000).

Establishing measurement equivalence reduces measurement bias confounds and facilitates comparative research across samples from different cultures. That is, comparing results across culturally different groups can be misleading in the absence of measurement equivalence because observed differences or similarities in mean levels or in the pattern of correlations between variables being potentially artifacts (Reise, Widaman, & Puhg, 1993).

Hinkle (1994) indicated that the potential problems associated with cross-cultural testing include difficulty establishing equivalence across cultures, lack of appropriate cultural norms, differences in response sets across cultures, lack of semantic equivalence of the items across cultures, and differing test-taking attitudes across cultures. Abbot, Snyder, Gleaves, (2002) proposed that cross-cultural equivalence is best understood as freedom from context bias. A test may be valid and unbiased in one context, but biased in a different context. Thus the question is the extent to which the test is valid and equivalent across different cultural contexts.

Cross-Cultural Research and Eating Disorders

Dounchis, Hayden and Wilfley (2001) pointed out that it is important to understand the development of ethnic identity and the process of cultural adaptation because the degree to which individuals depart from the values and behaviors of their culture of origin plays a crucial role in their risk for developing eating disorders. Striegel-Moore and Smolak (2001) argue that the failure to gather comprehensive data regarding eating disorders symptoms in ethnically diverse samples has contributed to the belief that minority girls and women do not experience symptoms of eating disorders.

However, new research suggests that eating disorders are present among individuals from a wide variety of ethnic backgrounds (Striegel-Moore & Smolak, 2001).

Many instruments have been developed in the US to assess eating disorder. Some of these instruments have been used in other countries and with diverse groups in the US to examine eating disorder symptomathology. Although many of these instruments have been translated for use in different countries, seldom any of these measures have been cross-culturally validated (Striegel-Moore & Smolak, 2001). That is, until otherwise demonstrated, emic measures that were designed for a dominant-culture population in the United States or Anglo-European countries are translated and used with people from different cultures (Dana, 1993). For example, researchers have translated into Spanish some of the most widely used instruments for the assessment of eating and body-image problems (see Table 1). Although these studies attempted to assess eating and body image problems in Spanish speakers, none of these studies examined the equivalence of the English and Spanish versions of the tests.

Research with the Bulimia Test-Revised

Smith and Thelen (1984) generated a pool of 36 items to measure bulimic symptoms according to the criteria listed in the Diagnostic Statistical Manual-III (DSM-III; APA, 1980) for the diagnosis of Bulimia Nervosa. Later, these authors revised the original version of the BULIT according to the Revised version of the DSM-III (APA, 1987) and created the BULIT-R. During the revision, the authors changed some items of the BULIT to cover with greater fidelity the new criteria listed in the DSM-III-R. Thelen et al., (1991) conducted Principal Factor Analysis (PCA) with data collected with

the BULIT-R during the cross-validation process. The sample during the cross-validation consisted of bulimic and control subjects. They reported that five factors with eigenvalues greater than 1.0 emerged, with the five factors accounting for 46.8% of the total variance. In another principal-factor analysis, with data collected from college women, five factors also emerged, accounting for 41.6% of the total variance. Factor 1 included 22 items pertaining to bingeing, control, and body image. Factor 2 included 6 items pertaining to related to extreme weight-loss measures and fasting. Factor 3 contained 2 items related to use of exercise. Factor 4 included 4 items related to vomiting and laxatives. Factor 5 included 2 items related to diuretics.

However, using a school-based sample of adolescent boys and girls, Vincent, McCabe, and Ricciardelli (1999) reported a four-factor structure for the BULIT-R. The four-factor model reported by Vincent et al. divided the items into the categories of bingeing (14 items), extreme weight loss behaviors (7 items), loss of control (7 items), and normative weight loss behaviors (9 items). Adding greater inconsistency across evaluations of the structure of the BULIT-R, McCarthy, Simmons, Smith, Tomlinson, and Hill (2002) found that a one-factor model provided an excellent fit for the scores of a sample of adolescent boys and girls.

Acevedo-Cruz, Lebrón-Hernández and Reyes-Rodriguez (1996) translated into Spanish and adapted the BULIT-R for use in Puerto Rico. These authors administered the test to a sample of 156 female college students enrolled in a general psychology course at the University of Puerto Rico. Acevedo et al. scored the measure by clustering the items into four categories or constructs on the basis of a content-analysis of the

BULIT-R items. The four categories or constructs created were binge eating and loss of control over eating, body image concerns, inappropriate compensatory behaviors (vomiting, laxative use), and self-evaluation. Using the same translated version of the BULIT-R, Santos (1996) administered the questionnaire to a sample of 1,944 adult women and performed Exploratory Factor Analysis (EFA). However, this author found that the data fit a six-rather than a four-factor model. The six factors were (1) bingeing, (2) body and shape perception, (3) shape and control, (4) extreme behaviors to control weight, (5) vomiting, and (6) control capacity. Thus it seems that the factor structure of the BULIT-R is unstable across both English speaking and Spanish speaking samples.

Study Objectives

The literature reviewed above show that there is no consensus regarding the factor structure of the BULIT-R. The goal of this study was to investigate the psychometric properties of the BULIT-R across two different cultures. The present study used Confirmatory Factor Analysis (CFA) methodology in an attempt to use a more direct test of the factor structure of the BULIT-R across two samples. CFA allows for (a) simultaneous model fitting of a potential factorial structure in two or more groups, (b) tests of the cross-group equivalence of all reliable measurement parameters, and c) comparisons of latent means between groups. However, the CFA results were inconclusive and the statistical analysis strategy was switched to EFA.

METHOD

Participants

Participants were 200 Caucasian undergraduate psychology students from a southwestern public university and 204 Spanish college students from the University of Granada, Spain. Following informed consent, participants completed the BULIT-R, and then received a written debriefing form. The debriefing form explained in more detail the study's objectives and included information about resources available to them regarding eating disorders information and mental health services.

Measures

The Bulimia Revised-Test (BULIT-R; Thelen et al, 1991) is a 36 item self-report, multiple choice measure that assess Bulimia Nervosa symptoms based on the criteria outlined in the DSM III-R (APA, 1987). It assesses the frequency of bulimic behaviors such as bingeing, inappropriate compensatory behaviors (vomiting, laxative use), and a sense of loss of control while eating. The BULIT–R has 28 items that are scored and 8 items pertaining to weight control that, for diagnostic purposes are not used. Thelen et al. (1991) reported a 2-month test-retest reliability of .95 for the BULIT-R. These authors also found that the BULIT-R's scores differentiated between participants with bulimia and those without an eating disorder (Thelen et al., 1991).

Other research has reported high internal consistency (r = .97) with data collected with the BULIT-R (Williamson, Anderson Jackman & Jackson,1995). Thelen, Mintz, and Vander-Wal (1996) performed a validation of the BULIT-R using DSM-IV (APA,

1994) criteria for Bulimia Nervosa and found high internal consistency in their set of scores (.98). Using discriminant analysis, these authors found the BULIT-R to be highly sensitive for identifying bulimics in a sample of college women, and for discriminating between individuals with bulimia and individuals from nonclinical samples.

In the present study two Spanish versions of the BULIT-R were used to derive a single Spanish version. One of the versions was that used by Acevedo, Lebrón, and Reyes (1994). The second version was obtained from a Spanish eating disorders clinic. The authors of the Spanish translation are unknown. The Puerto Rican version was examined by an eating disorders investigator from Spain, who suggested alternatives for any words she judged to be awkward or rarely used colloquially by Spaniards. A Puerto Rican investigator examined the Spanish version and made changes to words that she judged to be awkward or rarely used colloquially by Puerto Ricans. Both versions were then compared and any items that were not identical across the two forms were further modified to find a common wording that would fit the Spanish of both countries. This latter step was conducted by a new pair of Spanish and Puerto Rican investigators. Finally, after the language of the Spanish version was judged to be adequate for use in both countries, the measure was back translated into English. The back translation was examined by three independent English speakers, who found the back translation to be accurate.

Analysis

EFA were conducted using Statistics Package for the Social Sciences software (SPSS for Windows Version 11.0, 2001) and CFA were conducted using LISREL 8.5

(Jöreskog & Sörbom, 1996). We tested the different BULIT-R models reported by various investigators: a one-factor model (McCarthy et al., 2002); a four-factor model (Vincent et al., 2002), five-factor model (Thelen et al., 1991) and a six-factor model (Santos, 1996). To test fit indices, we examined χ^2 statistics (and associated p value), the Goodness-of-Fit-index (GFI; Jöreskog & Sorbom, 1989), the Comparative Fit Index (CFI; Bentler, 1990), the Non-Normed Fit Indices (NNFI; see Marsh, Balia, & McDonald, 1988), and the Root Mean Square Residual (RMSR; Steiger, 1989). Values of the GFI, CFI, NNFI, range from zero to 1.00, with values over .90 indicating a good fit (Mulaik et al., 1989). For the RSMEA, values of less than .05 are considered a close fit, and less than .08, and adequate fit (Finch & West, 1997). As mentioned above, the BULIT-R consists of 36 items, but 8 of them are not used in the final score of the test. The analyses were conducted using both the 36-item and the 28- item versions. However, the results across both versions were very similar. Thus, for conciseness-sake and to facilitate comparisons across previous factorial analyses of the BULIT-R, we present only the data of and analyses conducted on the long version.

RESULTS

Confirmatory Factor Analysis

Kaiser's Measure of Sampling Adequacy (MSA; Kaiser, 1960) for the English version was .92 which Kaiser described as "marvelous". MSA values for individual items were over .82 for the items in the English version of the test used with the Caucasian sample. Reliability analysis indicated that the overall alpha for the English version of the BULIT-R was .95. Table 1 shows that item-total correlations ranged from .13 to .81. The Spanish data obtained an MSA value of .90. All MSA values for individual items were over .80. The overall alpha for the Spanish version was .94. As presented in Table 2, item-total correlations were high, ranging from .21 to .81.

Table 3 depicts the CFA conducted on the Spanish and English versions of the BULIT-R. All the models tested with the Caucasians (n = 200) and Spaniards (204) fit the data poorly. None of the fit values for the GFI, CFI and NNFI exceeded .76 for any of the models within any of the samples. Likewise, the RSMEA values were all above .12, confirming the model fit the data poorly (see Table 3). The four-factor model for the Spanish version did not even converge, an indication of the model's inadequacy for the data.

Exploratory Factor Analysis

In the absence of a model fit from these the previous analysis, EFA were performed for each sample separately. The analysis was conducted using Maximum Likelihood with promax rotation. The number of factors was selected according to the scree test. For the purpose of this study, factors were interpreted according to items that

have high specificity. High specificity was defined following two criteria: the item's highest loading had to be at least .40, and the difference between this loading the item's second highest loading had to be more than .25.

Caucasians. The scree test revealed 6 factors that accounted for 61.8 % of the variance (see table 4). Following rotation of the factors, Factor 1 accounted for 11.4% of the variance and contained 15 items about binge eating or lack of control over eating. Factor 2 accounted for 12.3% of the variance and contained 12 items regarding body concerns/preoccupation. Factor 3 accounted for 8.3% of the total variance and contained 3 items related to vomiting. Factor 4 accounted for 3.1% of the variance and contained 2 items concerned with the use of diuretics. Factor 5 included 2 items regarding the use of laxatives and accounted for 6.5% of the variance. Factor 6 contained 2 items concerned with exercise and accounted for 4.8% of the variance. Factors 1 and 2 contained 11 and 9 items, respectively, that met the criteria for high specificity. Each of Factors 3, 4, 5, and 6 included 2 items with high specificity. That is, 32 out of 36 items met the criteria for high factor specificity. However, most of the items (20 items) fell into two categories (binge eating and body concerns), with the remaining 8 items being equally distributed across the four remaining categories. These four categories were composed of items reporting different forms of compensatory behaviors (vomiting, use of diuretics, use of laxatives, and exercising).

Spaniards. The scree test revealed four factors that accounted for 48.7 % of the variance. As table 5 indicates, the first factor included 17 items, most of them related to body concerns, and accounted for a 10.5% of variance. Factor 2 accounted for 9.3% of

the variance and contained 13 items concerned with binge eating or lack of control with eating. Factors 1 and 2 contained 10 and 11 items, respectively, that met the criteria for high specificity. Each of Factors 3 and 4 accounted for 4.1% of the variance and was made up of 3 items. Factors 3 and 4 were concerned with use of diuretics and laxatives, respectively. Only two items in each of these two factors met the criteria for high specificity. That is, 25 out of 36 items met our criteria for high factor specificity. Like in the American data, most of the items (21 items) fell into two categories (binge eating and body concerns), with the remaining 4 items being equally distributed across categories that reflected different forms of compensatory behaviors (use of diuretics and use of laxatives). That is, the items about vomiting and exercising did not load high in any of the 4 factors.

Frequency analyses of item endorsement were performed to compare the vomiting- and exercise-related factors across the Spanish and Caucasian samples. Results showed that Caucasians and Spaniards did not differ in item endorsement for questions about vomiting. On the other hand, there were considerable differences between samples in item endorsement for the two items of the exercise factor. For example item (11) "I exercise in order to burn calories", 42% of Spanish versus a 58% of Caucasians endorsed this item. Likewise, for item (20), "I exercise vigorously and for long periods of time in order to burn calories" 12% of Spaniards versus 25% of Caucasians endorsed the item to some level.

CONCLUSION AND DISCUSSION

In past decades there has been an increased interest in evaluating how cultural factors relate to mental health problems such as eating disorders. The use of valid and reliable psychological instruments across cultures is necessary in order to obtain an accurate assessment. The current study evaluated the psychometric properties and factor structure of the BULIT-R across samples from two different countries, the USA and Spain. Two types of statistical analyses were used: CFA and EFA.

The results of this study indicated that the scores obtained with the BULIT-R were adequately reliable for both the American and Spanish samples of college students. These results are consisted with previous studies that indicated a good reliability of the BULIT-R scores in Caucasian samples (Brelsford, Hummel & Barrios, 1992; Thelen, Mintz, & Vander-Wal, 1996).

The fit indices for all the models tested with CFA were poor. As a result, none of the factor solutions that were tested provided a good fit in any of the samples. Therefore, it was not appropriate to perform a test of measurement equivalence across samples using a sequence of multi-sample, "stacked" measurement models as was originally envisioned.

Results from the EFA using statistical criteria of eigen value higher than one indicated that a six factor solution fit the data for the Caucasian sample. Santos (1996) using a translation of the BULIT-R by Acevedo, Lebrón and Reyes (1993) in a Puerto Rican sample found that a six factors solution also provided a good fit for the data. However, with the exception of the first (binge), second (body image), and, to some

extent the third (control and shape) factors, the items that loaded in the other four factors were different across the present study and the study by Santos (1996).

The four-factor model proposed by Vincent, McCaabe and Ricciardelli (1999) did not pass the CFA test but the EFA in the Spaniard sample yielded a four factor model. However, Vincent et al (1999) only used 30 rather than all 36 items, leaving out the 6 items that refer to body image concerns. Only the binge factor in the Vincent et al. study looked similar to the binge factor of the Spanish sample.

The factor structures of Caucasians and Spaniards showed some similarities and some differences of item allocation in the common factors across samples. Nine of eleven items in the binge-eating factors were the same across both samples. The two remaining items with high specificity in the binge-eating factor within each sample also loaded high in the binge-eating factor of the other sample. With regards to the body-concerns factor, seven items were the same across samples.

Most differences were found for items related to exercise and vomiting, which loaded in separate factors in the Caucasian sample but loaded poorly across factors in the Spanish sample. However, while the frequencies of item endorsement across samples were similar for vomiting-related items, the frequencies of item endorsement were substantially different for the items related to burn calories.

Using the results from the present sample, and comparing our results to previous findings, it appears that the BULIT-R consistently measures two robust factors, binge eating and body concerns. The failure to replicate any of the factor solutions previously reported in the literature raises questions about the appropriateness of the BULIT-R to

capture a stable multifaceted construct. Several alternatives have been proposed to explain why such lack of generalizability might be associated with some assessment measures. Most often, though, when researchers find a bad fit they may decide to use only those items that in their sample appear to provide a better fit. That is, they may throw away items until an acceptable fit is accomplished irrespective of the theoretical merits of reaching such decisions (Ferrando & Lorenzo-Seva 2000). This situation leads to the development of instruments that are sample specific and do not replicate across samples.

Nonetheless, the present results appear to contribute new knowledge with regards to the universal or emic nature of the structure of eating disorder symptoms as they are manifested across college students from both countries. That is, symptoms were structured consistently into two factors, body-concerns and binge-eating items, in both samples.

REFERENCES

Acevedo, A., Lebrón, D., & Reyes, M. (1994). *Traducción y adaptación del instrumento EAT-26*. [Translation and adaptation of the EAT-26]. Manuscrito no publicado. Universidad de Puerto Rico, Recinto de Río Piedras, PR.

Allen, J., & Walsh, J. A. (2000). A construct-based approach to equivalence: methodologies for cross-cultural/multicultural personality assessment research. In Dana, R.H. (Ed.) *Handbook of cross-cultural and multicultural personality assessment*. (63-85). Publishers, Mahwah, NJ: Lawrence Erlbaum Associates.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders.* (3rd ed.,) Washington, DC: Author.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders.* (DSM-III, 3rd ed., Rev.) Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (DSM-IV, 4th ed.) Washington, DC: Author.

Bentler, P. M. (1990). Comparative indexes in structural models. *Psychological Bulletin, 107,* 238-246.

Brelsford, T.N., Hummel, R.M., & Barrios, B.A. (1992). The Bulimia Test Revised: A Psychometric investigation. Psychological Assessment, 4, 339-401.

Butcher, J. (1996). Understanding abnormal behavior across cultures: The use of objective personality assessment methods. (3-25) In Butcher, J.N. (Ed.) *International adaptations of the MMPI-2 research and clinical applications.* Minneapolis: University of Minnesota Press.

Campbell, D. T. (1968). A cooperative multinational opinion sample exchange. *Journal of Social Issues 24*, 245-258.

Castro, J., Toro, J., Salamero, M., & Quimera, E. (1991). The Eating Attitudes Test: Validation of the Spanish version. *Psychological Assessment, 7* (2), 175-190.

Council of National Psychological Associations for the Advancement of Ethnic Minority Interests (CNPAAEMI). (2000). *Guidelines for research in ethnic minority communities.* Washington, DC: American Psychological Association.

Dana, R.H. (1993). *Multicultural assessment perspective for professional psychology.* Boston: Allyn and Bacon.

Dounchis, J.Z, Hayden, H.A., & Wilfley, D.E. (2001). Obesity, body image, and eating disorders in ethnically diverse children and adolescents. In Thompson, J. Kevin Smolak, Linda (Eds). *Body image, eating disorders, and obesity in youth: Assessment, prevention, and treatment.* (67-98) Washington, DC: American Psychological Association.

Ferrando, P.J. & Lorenzo-Seva, U. (2000). Unrestricted versus restricted factor analysis of multidimensional test items: Some aspects of the problem and some suggestions. *Psicológica ,21*, 301-323.

Finch, J.F., & West, S.G. (1997). The investigation of personality structure: Statistical models. *Journal of Research in Personality, 31,* 439-485.

Fridja, N., & Jahoda, G. (1966). On the scope and methods of cross-cultural research. *International Journal of Psychology, 1*, 109-127.

Garner, D.M., & Garfinkel, P.E. (1979). The Eating Attitudes Test: An index of the

symptoms of anorexia. *Psychological Medicine, 9,* 273-279

Harris, M. (1976). History and significance of the emic/etic distinction. *Annual Review of Anthropology, 5,* 329-350.

Helms, S.A. (1992). Why is there no study of cultural equivalence in standardized cognitive ability testing? *American Psychologist, 47,* 1083-1101.

Hinkle, J.S. (1994). Assessment in action. Practitioners and cross-cultural assessment: A practical guide to information and training. *Measurement and Evaluation in Counseling and Development, 27,* 103-115.

Jöreskog, K. G., & Sörbom, D. (1989). *LISREL 7-A guide to the program and applications.* Chicago: SPSS Publications.

Jöreskog, J.S., & Sorbom, D. (1996). *LISREL 8.5 Users' referennce guide.* Chicago: Scientific Software.

Kaiser, H.F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20,* 141-151.

Lebrón, D.A. & Reyes, M. (1993). *Estudio sobre conductas bulímicas en mujeres estudiantes de la Universidad de Puerto Rico.* [Study about bulimics behaviors in female collage students from University of Puerto Rico]. Recinto de Rio Piedras. Tesis de maestría sin publicar: Universidad de Puerto Rico, Rio Piedras.

Marin, G., & Marin, B.V (1991). *Research with Hispanic populations.* Thousand Oaks, CA: Sage.

Marsella, J.A., & Yamada, A.M. (2000). Culture and mental health: An

introduction and overview of foundations, concepts, and issues. In Paniagua F. A. (Ed.), *Handbook of Multicultural Mental Health*. (3-23). San Diego: Academic Press

Marsh, H.W., Balia, R., & McDonald, R.P. (1988). Goodness-of-fit Indices in confirmatory factor analysis. The effect of sample size. *Psychological Bulletin, 103,* 391-410.

McCarthy, D. M., Simmons, J. R., Smith, G. T., Tomlinson, K., L. & Hill, K. L. (2002). Reliability, stability, and factor structure of the Bulimia Test-Revised and Eating Disorder Inventory-2 scales in adolescence. *Assessment, 9* (4), 382-389

Mulaik, S.A., James, L.R., Alstine, J.V., Bennet, N., Lind, S., & Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin, 10,* 430-445.

Padilla, A. M. (Ed. (1995). *Hispanic psychology: Critical issues in theory and research.* Thousand Oaks, CA: Sage.

Prediger, D. J. (1993). *Multicultural assessment standards: A compilation for counselors.* Alexandria, VA: Association for Assessment in Counseling.

Raich, E. R., Mora, G.M., Torras, C., J., & Sánchez, C.D. (2000). Uso de la entrevista EDE para evaluar los trastornos alimentarios en población universitaria [Use of the EDE interview to evaluate eating disorders in a university population]. *Psicología Contemporánea, 7* (1), 36-47.

Reise, S.P., Widaman, K.F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin, 114,* 552-566.

Santos, D. (1996). *Prevalencia de bulimia nerviosa en una muestra probabilística de mujeres que residen en áreas pobres en Puerto Rico*. [Prevalence of bulimic symptoms in a probabilistic simple of women who live in poor areas of Puerto Rico]. Disertación doctoral sin publicar. Universidad de Puerto Rico, Rio Piedras

Smith, M. C., & Thelen, M. H. (1984). Development and validation of a test for bulimia. *Journal of Consulting and Clinical Psychology, 52,* 863-872.

Spengler, P.M. (1998). Multicultural assessment and a scientist and a scientist-practitioner model of psychological assessment. *Counseling Psychologist, 6*, 930-938.

Scientific Software. (2001). Statistical package for the social sciences (SPSS) [Version 11.1]. Chicago: Author.

Steiger, J.H. (1989). *EzPATH: Causal modeling*. Evanston, IL: SYSTAT.

Striegel-Moore, R.H. & Smolak, L., (2001). *Eating disorders. Innovative directions in research and practice.*Washington, DC: APA.

Thelen, M. H., Farmer, J., Wonderlich, S., & Smith, M. (1991). A revision of the Bulimia Test: The BULIT-R. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3,* 119-124.

Thelen, M, H., Mintz, L. B., & Vander Wal, J.S. (1996). The Bulimia Test-Revised: Validation with DSM-IV criteria for bulimia nervosa. *Psychological Assessment, 8,* 219-221.

US Public Health Service Surgeon General (2002). *Mental health: Culture, race, and ethnicity*: A report of the Surgeon General. Rockville MD: Author.

Vincent, M. A., McCabe, M. P., & Ricciardelli, L.A. (2002). Factor validity of the

　　Bulimia Test-Revised in adolescent boys and girls. *Behavior Research and*

　　Therapy, 37, 1129-1140.

Williamson, D.A., Anderson, D.A., Jackman, L. P., & Jackson, S.R. (1995). Assessment

　　of eating disordered thoughts, feelings, and behaviors. In D.B. Allison (Ed.),

　　Handbook of assessment measures for eating behaviors and weight-related

　　problems: Measures, theory, and research (pp. 347-386). Thousand Oaks, CA:

　　Sage.

APPENDIX

Table 1

Translations of eating disorders measures

Author	*Instrument*	*Psychometric Properties*
Castro, Toro, Salamero & Quimera (1991)	Eating Attitude Test (EAT; Garner & Garfinkel, 1979)	Validity coefficient: .61 Internal consistency: .93 - .92
Lebrón & Reyes (1993)	BULIT-R (Thelen, Farmer, Wonderlich & Smith, 1991)	Internal consistency: .93
Acevedo, Lebrón, & Reyes (1994)	EAT-26 (Garner Olmsted, Bohr, & Garfinkel, 1982)	Internal consistency: .80, Overall; 0.82, women; .76 men.
Raich, Mora, Torras, Sánchez (2000)	EDE (Garner, Olmsted, & Polivy, 1983)	Internal consistency : .80 - .85

Table 2

Item-total correlations and measures of sampling adequacy for Caucasians and Spaniards samples

BULIT Items	U.S. Sample		Spanish Sample	
	Corrected Item-Total Correlations	Kaiser MSA	Corrected Item-Total Correlations	Kaiser MSA
1	.614	.960	.447	.916
2	.705	.948	.568	.939
3	.615	.958	.373	.869
4	.678	.937	.529	.891
5	.740	.925	.598	.914
6	.129	.539	.208	.610
7	.735	.945	.652	.957
8	.543	.924	.569	.907
9	.673	.940	.563	.890
10	.365	.906	.408	.906
11	.332	.734	.321	.765
12	.694	.943	.604	.936
13	.662	.938	.670	.935
14	.749	.946	.673	.913
15	.483	.824	.526	.894
16	.572	.943	.605	.938
17	.634	.943	.384	.930
18	.700	.952	.379	.949
19	.726	.929	.779	.956
20	.418	.837	.692	.847
21	.721	.949	.399	.904
22	.689	.940	.571	.861
23	.692	.933	.382	.923
24	.685	.938	.551	.941
25	.756	.953	.710	.954
26	.476	.854	.719	.789
27	.207	.636	.314	.663
28	.801	.961	.345	.941
29	.532	.899	.677	.936
30	.374	.907	.626	.915
31	.282	.746	.507	.760
32	.797	.929	.370	.936
33	.628	.931	.813	.901
34	.704	.958	.588	.927
35	.668	.934	.644	.939
36	.255	.665	.374	.703

Table 3

Goodness of fit indices for one factor, four factors, five factors, six factors for Caucasians and Spaniards samples

Dimensions	χ2	df	p value	GFI	CFI	NNFI	RMSEA
C 1 Factor 36 items	3505.95	594	<0.01	0.51	0.61	0.55	0.16
C 4 Factors 36 items	4180.56	587	<0.01	0.46	0.46	0.52	0.17
C 5 Factors 36 items	2643.57	584	<0.01	0.58	0.73	0.71	0.13
C 6 Factors 36 items	1690.70	335	<0.01	0.62	0.72	0.68	0.14
S 1 Factor 36 items	1828.96	350	<0.01	0.66	0.70	0.68	0.14
S 4 Factors (did not converge)							
S 5 Factors 36 items	2216.77	584	<0.01	0.62	0.74	0.72	0.12
S 6 Factor 36 items	1376.81	335	<0.01	0.67	0.71	0.69	0.12

Note: C= Caucasian, S= Spaniards, GFI= Goodness of Fit Index, NNFI= Non-Normed Fit Index, CFI= Comparative Index, RMSEA= Root Mean Square Error or Approximation. The p values are not zero. If you are going to include them, list them as some value

Table 4

Factor loadings for 6 factors extracted from the BULIT-R for Caucasians

Items	Factor					
	1	2	3	4	5	6
Binge						
(23)	**.87***	-.11	.12	.06	-.09	.02
(8)	**.85***	-.15	-.05	-.04	.07	-.05
(21)	**.79***	.05	-.10	-.08	.17	-.01
(34)	**.79***	-.00	.13	-.07	-.04	-.00
(16)	**.76***	.09	-.10	.02	-.02	-.15
(9)	**.71***	.00	-.09	-.01	.07	.17
(2)	**.68***	-.10	.31	.02	-.08	.08
(22)	**.66***	.27	-.11	.07	-.11	.00
(30)	**.53***	-.07	.01	.02	-.02	-.05
(32)	**.52***	.28	-.04	-.02	.13	.10
(10)	**.51***	-.21	.09	-.02	.12	-.04
(3)	.47	.35	-.12	-.03	-.02	-.01
(33)	.45	.32	-.05	.01	-.10	.07
(17)	.45	.22	.04	.06	.07	-.11
(13)	.26	.11	.21	.09	.19	.09
Body concerns						
(4)	.07	**1.0***	-.24	-.01	-.23	-.00
(14)	-.08	**.94***	.06	-.00	-.07	-.04
(7)	-.03	**.92***	-.11	-.09	.09	-.02
(24)	.01	**.87***	-.15	-.01	.06	-.07
(1)	.02	**.86***	-.05	-.07	-.20	.00
(12)	-.07	**.69***	.11	.11	.09	-.06
(29)	-.28	**.60***	.19	-.01	.15	.02
(19)	.04	**.55***	.10	.04	.11	.07
(28)	.29	**.52***	.15	.12	-.07	-.05
(25)	.10	.45	.28	.10	.01	-.17
(35)	-.10	.41	.40	-.02	.18	.11
(5)	-.10	.41	.40	-.02	.18	.11
Vomiting						
(15)	.05	-.12	**.99***	-.05	-.16	-.06
(26)	-.01	.03	**.88***	.02	-.17	-.12
(18)	.15	.23	.33	-.19	.10	.23
Diuretics						
(27)	-.02	-.03	-.03	**.98***	.00	.00
(36)	.00	-.01	-.01	**.95***	-.01	.01

Table 4 (continued)

Items			Factor			
	1	2	3	4	5	6
(20)	.10	-.14	-.06	.03	**.93***	-.13
(11)	-.00	.04	-.28	-.03	**.89***	-.06
Laxatives						
(31)	-.07	-.02	-.17	.19	-.01	**.90***
(6)	-.01	-.08	-.05	-.11	.-17	**.81***

* Indicates items that load highly on each respective factor

Table 5

Factor loadings for 4 factors extracted from the BULIT-R for Spaniards

Item	Factor			
	1	*2*	*3*	*4*
Body Concerns				
(14)	**.99***	-.13	-.02	-.17
(7)	**.86***	-.10	.00	-.04
(12)	**.86***	-.24	.03	.067
(29)	**.80***	-.08	-.05	.00
(19)	**.78***	.02	-.09	.03
(4)	**.77***	-.09	.04	-.21
(25)	**.74***	.03	.04	.03
(24)	**.67***	.23	-.07	-.08
(5)	**.66***	-.14	-.03	.34
(18)	**.55***	.23	.13	.04
(32)	.51	.45	-.04	-.01
(13)	.51	.17	.04	.10
(28)	.44	.28	.01	.05
(10)	.44	.27	-.00	.07
(1)	.28	.25	.03	-.09
(11)	.28	.09	.00	-.00
(20)	.20	.17	.03	.10
Binge				
(2)	-.14	**.87***	-.04	.00
(9)	-.05	**.78***	-.00	-.07
(34)	-.11	**.74***	-.00	.07
(16)	.01	**.74***	.01	-.09
(23)	-.04	**.66***	.02	.03
(10)	-.15	**.62***	-.02	.11
(8)	.09	**.62***	.00	-.08
(22)	-.03	**.55***	-.01	-.10
(21)	.08	**.55***	-.05	.14

Table 5 (continued)

Item	Factor			
	1	2	3	4
(33)	.18	**.51***	-.00	-.04
(3)	.01	**.49***	.01	-.17
(30)	*.17*	*.39*	*-.09*	*.12*
(15)	.03	.37	.28	.10
Diuretics				
(27)	.00	-.03	**1.0***	-.05
(36)	-.02	-.04	**.88***	.05
(26)	.11	.16	.21	-.02
Laxatives				
(31)	.05	-.05	-.02	**.87***
(6)	.05	-.06	.02	**.82***
(17)	.19	.11	.05	**.19**

* Indicates items that load highly on each respective factor

VITA

Permanent Address
Mayra N. Berríos-Hernández
PO Box 7733
College Station, TX 77844

American Psychological Association, Minority Mental Health and Substance Abuse Services Fellow, 2002-2005

Education:
Bachelor of Arts, Psychology. University of Puerto Rico, May 2002
Master of Science, Psychology. Texas A&M University, May 2005

Publications:
Rosselló, J. &, Berríos, M. N. (2004). Ideación suicida, depresión, auto-estima, actitudes disfunciones y eventos de vida negativos en una muestra de adolescentes puertorriqueños/as. (Suicide ideation, depression, self-esteem, dysfunctional attitudes and negative life events in a sample of Puerto Rican adolescents). *Revista Interamericana de Psicología, 38,* (2).

Roselló, J. & Berríos, M.N. (2003). DMID y autoestima en niños/as y adolescentes puertorriqueños/as).(IDDM and self-esteem in Puerto Rican children and adolescescents). *Revista Asociación Puertorriqueña de Psicología, 14, 107-126.*

Selected Presentations:
Berríos-Hernández, M.N., Cepeda-Benito, A, Gleaves, D., Rodríguez-Ruiz, S. Maysonet M. (November, 2004). Cross-cultural assessment of eating disorders: Psychometric Properties of a Spanish version of Bulimia Test-Revised. Paper presented at the 38th Annual Convention of the Association for Advancement of Behavior Therapy, New Orleans, LA

Berríos-Hernández, M.N., Susabda, A., Reig-Ferrer, A., Cepeda-Benito, A. (October, 2004). Evaluation of smoking expectancies in nonsmokers. Poster presented a the Presented at the National Hispanic Science Network on Drug Abuse Conference, San Antonio, TX

Berríos-Hernández, M.N., Susabda, A., Reig-Ferrer, A., Cepeda-Benito, A. (November, 2003). Evaluation of smoking expectancies in nonsmokers. Poster presented a the Presented at the 37th Annual Convention of the Association for Advancement of Behavior Therapy, Boston, MA