

Vol. 20, No. 1, 2018

Systematic Validation of a Self-Administered Questionnaire to Assess Bullying: From Elementary School to High School and by Sex

Validación sistemática de un cuestionario autoadministrado para evaluar el *bullying*: desde primaria hasta preparatoria y por sexo

Arnulfo Ramos-Jiménez (1) aramos@uaci.mx
Óscar A. Esparza del Villar (1) oesparza@uaci.mx
Alberto Castro-Valles (1) alcastro@uaci.mx
Rosa P. Hernández-Torres (2) rphernant@yahoo.com
Miguel Murguía-Romero (3) miguelmurguia@gmail.com
Rafael Villalobos-Molina (3) rafaelvillalobosmolina@gmail.com

(1) Universidad Autónoma de Ciudad Juárez

(2) Universidad Autónoma de Chihuahua

(3) Universidad Nacional Autónoma de México

(Received: August 17, 2016; accepted for publishing: October 27, 2016)

How to cite: Ramos-Jiménez, A., Esparza, O. A., Castro-Valles, A., Hernández-Torres, R. P., Murguía-Romero, M. y Villalobos-Molina, R. (2018). Systematic validation of a self-administered questionnaire to assess bullying: from elementary school to high school and by sex. *Revista Electrónica de Investigación Educativa*, 20(1), 26-37. Retrieved from <https://doi.org/10.24320/redie.2018.20.1.1535>

Abstract

There is a need to develop and validate an instrument to assess school bullying in the Mexican population in order to evaluate the issue and intervene accordingly. This study validates a bullying assessment instrument named Bull-M in children from 5th to 12th grade. The Bull-M was administered to a sample size of 2,030 students from 5th to 12th grade in several schools in Ciudad Juárez, Mexico. A two-factor structure of the Bull-M was analyzed with confirmatory factor analyses across the whole sample and gender and educational levels. The confirmatory factor analyses indicate good model fits, strong factor loadings and adequate Cronbach's alpha values to assess internal reliability of the factors and scale. The Bull-M can be used to assess bullying in students from 5th to 12th grade in northern Mexico.

Keywords: Confirmatory factor analysis, school violence, internal reliability, emotional health.

Resumen

Es necesario desarrollar y validar un instrumento para valorar el acoso escolar en la población mexicana a fin de evaluar el problema e intervenir en consecuencia. Este estudio valida un instrumento de evaluación de intimidación llamado Bull-M, el cual se administró a una muestra de 2,030 estudiantes de 5o. a 12o. grado en varias escuelas de Ciudad Juárez (México). Se analizó una estructura de dos factores del Bull-M con análisis de factores confirmatorios en toda la muestra y los niveles de género y educación. Los análisis

factoriales confirmatorios indican buenos ajustes del modelo, fuertes cargas de factores y valores adecuados de alfa de Cronbach para evaluar la confiabilidad interna de los factores y la escala. El Bull-M se puede usar para evaluar la intimidación en estudiantes de 5o. a 12o. grado en el norte de México.

Palabras clave: Análisis factorial confirmatorio, violencia escolar, confiabilidad interna, salud emocional.

I. Introduction

In many countries, repeated aggressive behavior in students, known as bullying, has become a social health problem (Levine & Tamburrino, 2014). This is because victims (the bullied) and aggressors (the bullies) have higher incidences of health complaints, depression (Due, Damsgaard, Lund, & Holstein, 2009; Fleming & Jacobsen, 2009; Hanley & Gibb, 2011), and interpersonal relationship problems (Gilmartin, 1987), which later may result in different forms of violence, criminality, substance abuse, suicidal thoughts and suicide attempts (Bauman, Toomey, & Walker, 2013; Fleming & Jacobsen, 2009). The global prevalence of bullying in high schools has been found to range from 5% to 45%. This figure is higher in women (Craig et al., 2009), and an increase in violent behavior has also been reported (Bickmore, 1997). In addition, it has been reported that the prevalence of school bullying may differ depending on the educational level (Nansel, Overpeck, Pilla, Ruan, Simons-Morton, & Scheidt, 2001; Sawyer, Bradshaw, & O'Brennan, 2008), decreasing in older students but not disappearing (Alzahrani, 2012; Ozkal, 2011). The reduced prevalence in older students is primarily related to physical aspects, but there is an increase in indirect or social forms, like exclusion, verbal and cyberbullying, among others (Alzahrani, 2012; Bauman, Toomey, & Walker, 2013; Ozkal, 2011).

Violence in schools in Mexico has been measured nationwide as part of an extensive survey that measured several child-related aspects in 2000 (Secretaría de Salud, 2006), 2012 (Instituto Federal Electoral, 2012), and 2015 (Instituto Federal Electoral, 2015). In the year 2000, information was obtained from 4,000,000 girls and boys with the objective of estimating violence reported by students in school. The results indicated that 32% of children between the ages of 6 and 9 and 13% of children between the ages of 10 and 13 reported having been victims of violence in school (Secretaría de Salud, 2006). A similar survey, which measured only school violence in the 13 to 15-year-old age group, was administered in 2012 to 2,256,532 children across all the states of Mexico (Instituto Federal Electoral, 2012). The results indicated that 12% of students reported being harassed or intimidated by peers in school and 4% reported being victims of sexual abuse in schools.

The most recent survey was reported in 2015 with a sample size of 2,916,686 children from all the states of Mexico, and violence in schools was only measured in the 10 to 13-year-old age group (Instituto Federal Electoral, 2015). The results showed that 14.9% reported physical violence, 26.3% verbal violence, 19.5% emotional violence, and 2.9% sexual violence in school. These indicators were also measured in the state of Chihuahua, where the study sample is from, and for the state the results were similar, as 14.5% reported physical violence in school, 28.6% verbal violence, 21.6% emotional violence, and 2.8% reported sexual violence in school.

Bullying in Mexico has become an important topic, particularly in places where there has been a lot of violence, like in Ciudad Juárez, which was named the most violent city in the world in 2010 due to the war between drug cartels and the Mexican government. Between 2007 and 2011, more than 9,000 people were murdered in Ciudad Juárez (Valencia & Chacon, 2013), producing social, cultural and psychological consequences. Violence was part of citizens' everyday lives, as it appeared in television broadcasts, radio, newspapers, social media, and the Internet. Children were aware of the problem since they had to live – and survive – in this environment full of violence and death. It is in this context that researchers in Mexico, and Ciudad Juárez specifically, decided to analyze this problem to intervene and attempt to minimize the negative consequences in children.

Qualitative observation, direct interviews, psychometric tests and self-report methods are accepted ways of evaluating bullying among school-based professionals (Casey, Hayward, & Gowen, 2001; Glover, Gough, Johnson, & Cartwright, 2000; Shapiro & Heick, 2004). Qualitative observation focuses on

elaborating on and explicating the experiences of bullied and bullies (Patton, Hong, Patel, & Kral, 2015); this method identifies the physical characteristic of the phenomena in situ better than verbal and emotional expressions, but its implementation and analysis are laborious processes. Direct interviews focus profoundly on the interpretation and perceptions of bullying phenomena, but it takes a long time for researchers to interpret responses while avoiding subjectivity and thinking errors (Creswell, 2003). Psychometric tests are a formal way to diagnose the psychological and emotional effects of bullying (Tyler, 1972), but are impossible to administer on a large scale. Finally, through a combination of closed and open inquiries, self-reports focus on the subjective perception of, or reflection on, a phenomenon to generate self-evaluations of a psychological condition (Alarcón, Pérez-Luco, Salvo, Roa, Jaramillo, & Sanhuesa, 2010).

Questionnaires are also useful tools to evaluate bullying behavior, because despite being less sensitive than the methods mentioned above, they are practical and economic (Crothers & Levinson, 2004). In Mexico, questionnaires to evaluate bullying have been conducted with no validity reported (Cerezo, 2006; Instituto Nacional de Salud Pública, 2006; Instituto Nacional de Salud Pública, 2012), so the results may not be reliable or valid. In addition, with the exception of the Bull-M, we do not know of validated questionnaires to evaluate bullying behavior in Mexico; the Bull-M has, however, only been validated in junior high schools (Ramos-Jimenez, Wall-Medrano, Esparza-Del Villar, & Hernández-Torres, 2013). Because bullying manifests itself differently with age and its prevalence changes, it becomes necessary to evaluate bullying at different school ages, and thereby implement better and effective prevention programs. The importance of the Bull-M is that it has been designed to assess school bullying anonymously and in a short time (10-15 min), thus increasing reliability when administered to larger groups of students and in areas with high rates of violence. In light of the foregoing, the aim of this study was to validate the Bull-M at three different school stages (from elementary school to high school) and by sex.

II. Methods

The sample consisted of 2,030 students from grades 5 to 12 (9 to 20-year-old age range) from the metropolitan area of Ciudad Juárez, in the state of Chihuahua, Mexico (see Table I): 1,001 girls and 1,029 boys from 17 elementary schools (1,110 students), six junior high schools (560 students), and six high schools (360 students). The students interviewed were those present in classrooms on the day the Bull-M was administered. According to the National System of School Statistics Information (SNIEE), there are 611 public schools in the city: 460 elementary schools, 106 junior high schools, 33 high schools and 12 universities, with approximately 400,000 students in total. In order to obtain a random and appropriate sample including both genders and all educational levels, a multi-step method was used for sampling. First, schools and educational levels were selected through a stratified probabilistic method using the SNIEE list of schools for Ciudad Juárez, and then students were selected randomly. In addition, no students younger than 5th grade were included as we considered that these students might not have been able to understand the questions asked.

Table I. Gender and age of sample by group and total

	n	Gender		Age Range
		Boys n (%)	Girls n (%)	
Total	2030	1029 (50.7%)	1001 (49.3%)	9-20
Elementary (1st to 6th grade)	1110	554 (49.9%)	556 (50.1%)	9-14
Junior high school (7th to 9th grade)	560	296 (52.9%)	264 (47.1%)	11-16
High school (10th to 12th grade)	360	179 (49.7%)	181 (50.3%)	15-20

The Bull-M was designed with the aim of understanding the phenomenon of bullying anonymously. It contains ten items, eight of which are divided into two theoretical constructs: the bullied factor (items 2-5), to identify if the participant is being bullied by peers; and the bully factor (items 6-9), to identify if the participant bullies others. There is also one introductory item about social relations, and at the end, one item about the physical symptoms related to bullying (see Table II for the content of items). Response options for this questionnaire follow a five-point Likert scale with the following values: never (0), seldom (1), sometimes (2), often (3), and every day (4). Because items 1 and 10 deal with the problem of bullying indirectly, only items 2 to 9 are included in the theoretical constructs for the confirmatory factor analyses as described in the two factors above. The Bull-M was designed initially by four researchers in the field of social and community health, with the support of seven elementary and junior high school teachers with over five years of direct in-class experience in schools with a high degree of violence, and where bullying often occurs. Other studies include a complete validation of the questionnaire design (Ramos-Jimenez, Wall-Medrano, Esparza-Del Villar & Hernández-Torres, 2013).

Once participating schools were selected, written informed consent was obtained from participating students, school administrators and parents prior to administering the questionnaire. If authorization was not given, an alternative school was randomly selected. Participants were allowed 20 minutes to complete the questionnaire, which was administered in a classroom without a teacher or other school authority figure present. Groups were comprised of up to 30 students, and the protocol was approved by school authorities and the Ethics Committee of the Universidad Autónoma de Ciudad Juárez (Autonomous University of Ciudad Juárez).

To analyze the internal reliability of the questionnaire between education levels and genders, Cronbach's alpha analyses were performed. Construct validity was assessed with a confirmatory factor analysis with a two-factor structure, as described previously (Hu & Bentler, 1999), using structural equation modeling with the maximum likelihood estimation (MLE) method. A previous study analyzed the factor structure of the scale using an exploratory factor analysis (EFA; Ramos-Jimenez et al., 2013). An EFA analyses the factors in which items group together without constraints, and suggests a specific number of factors for the scale according to eigenvalues greater than one (Yong & Pearce, 2013). Once a factor structure has been defined with the EFA, it must be validated to confirm it and the suggested procedure is to perform a confirmatory factor analysis (CFA), which compares observed variables (responses to items) to a proposed factor structure, usually defined by the structure found in the EFA (Kline, 2013). Since the Bull-M has been previously validated, in this study a CFA was used to confirm the factor structure of the scale (Kline, 2013).

Finally, measurement invariance was assessed between men and women. Measurement invariance assesses equivalence of a measurement across groups to verify if a scale measures the construct equally. The Bull-M is an indicator of a latent construct, bullying, and this study analyzed if the items relate to the latent variables in men and women equally. To assess this equivalence, the first step is to analyze the factor structure of the scale in both groups (men and women) and evaluate if the item responses of the samples fit the proposed two-factor model using structural equation modeling. If the model has an adequate fit, then the item loadings are constrained (metric invariance), assuming they are the same across both groups. If the new model with the constrained loading has an adequate fit, then the intercept equivalence, error variances and covariance equivalence are evaluated. The structural equation models were analyzed with the Amos 22.0 computer program; all other analyses were performed with spss 22.0.

III. Results

The factor structure analyzed in this study has been slightly modified in comparison to a previous study wherein the instrument was validated in Mexico (Ramos-Jimenez et al., 2013). The bullied factor includes items 2 to 5, and measures if the student is bullied by others. The bullies factor evaluates whether the student bullies others. For this structure, items 1 and 10 were taken off the bullied factor since they measure bullying indirectly; item 1 is related to social relations and item 10 is related to physical symptoms. The Cronbach's alpha if item is deleted and item-total correlations table supported our decision to evaluate item 1 by itself.

The first step was to analyze the internal reliability of the total scale and its factors by calculating Cronbach's alpha indices for the total scale, the bullied factor, and the bullies factor (see Table II). The internal reliabilities, including all participants, were $\alpha=0.82$ for the total scale, $\alpha=0.76$ for the bullied factor, and $\alpha=0.78$ for the bullies factor. The internal reliability was also calculated for each segment of the educational levels and by sex (see Table I). Eliminating any of the items would decrease the value of Cronbach's alpha for each of the factors, indicating that all of the items should be kept. When analyzing the internal reliability of the 10 items, the only increase in the alpha value would be if item 1 was deleted. This is the reason why item 1 is scored separately. Most of the internal reliability values were acceptable except for alpha values less than 0.70: total scale for middle school ($\alpha=0.68$), bullied factor for high school ($\alpha=0.67$), and bullies factor for middle school ($\alpha=0.67$). According to the item-total correlation table, all of the items have Pearson correlations greater than 0.30 (Cristobal, Flavián, & Guinalú, 2007), except for item 1, which has a very low correlation and is another reason why it is scored separately.

Table II. Internal reliability values using Cronbach's alpha, Cronbach's alpha if item is deleted, and item-total correlations for all participants and by educational level and sex

Item	Cronbach's alpha if item is deleted						Item-total correlation					
	All	Elementary	Middle school	High school	Men	Women	All	Elementary	Middle school	High school	Men	Women
1. Frequency of invitations from peers to participate with them	0.84	0.87	0.75	0.79	0.84	0.84	0.19	0.41	-0.02	0.02	0.20	0.18
2. Frequency of exclusion by peers.	0.80	0.85	0.64	0.72	0.80	0.79	0.55	0.59	0.46	0.41	0.52	0.59
3. Frequency of being obligated or threatened by peers	0.80	0.85	0.65	0.73	0.80	0.80	0.54	0.59	0.41	0.40	0.54	0.55
4. Frequency of being scoffed at, insulted or hurt by peers	0.79	0.84	0.63	0.71	0.80	0.79	0.58	0.62	0.47	0.46	0.57	0.59
5. Frequency of being blamed or accused by peers	0.79	0.84	0.65	0.70	0.79	0.79	0.60	0.65	0.42	0.51	0.59	0.60
6. Frequency of excluding peers	0.79	0.84	0.65	0.69	0.79	0.79	0.57	0.61	0.40	0.57	0.58	0.56
7. Frequency obligating or threatening peers	0.79	0.84	0.65	0.70	0.79	0.79	0.61	0.65	0.43	0.58	0.62	0.60
8. Frequency of scoffing at, insulting or hurting peers	0.79	0.84	0.64	0.69	0.79	0.79	0.59	0.64	0.45	0.55	0.62	0.55
9. Frequency of blaming or accusing peers	0.79	0.85	0.64	0.69	0.80	0.79	0.58	0.60	0.47	0.58	0.58	0.58
10. Frequency of physical symptoms (e.g. headache) in the last 4 weeks	0.82	0.86	0.69	0.74	0.82	0.81	0.38	0.46	0.22	0.30	0.39	0.41
Cronbach's α	0.82	0.86	0.68	0.74	0.82	0.82						
Bullied factor												
Item 2	0.70	0.73	0.60	0.57	0.69	0.71	0.58	0.57	0.56	0.51	0.57	0.58
Item 3	0.72	0.73	0.67	0.62	0.70	0.74	0.55	0.57	0.45	0.49	0.55	0.55
Item 4	0.70	0.71	0.61	0.63	0.68	0.71	0.59	0.61	0.54	0.44	0.58	0.59
Item 5	0.71	0.72	0.68	0.60	0.72	0.70	0.55	0.58	0.43	0.46	0.51	0.60
Cronbach's α	0.76	0.78	0.71	0.67	0.75	0.77						
Bullies factor												
Item 6	0.73	0.76	0.59	0.73	0.75	0.70	0.59	0.61	0.47	0.63	0.60	0.58
Item 7	0.72	0.74	0.60	0.75	0.74	0.70	0.61	0.63	0.48	0.63	0.62	0.59
Item 8	0.73	0.74	0.60	0.76	0.74	0.72	0.60	0.64	0.45	0.62	0.63	0.56
Item 9	0.74	0.76	0.62	0.74	0.76	0.72	0.57	0.59	0.42	0.61	0.59	0.55
Cronbach's α	0.78	0.80	0.67	0.80	0.80	0.77						

The second step was to analyze the confirmatory factor analyses for all of the participants and for each of the educational levels and sexes (see Table 3). The model fit was analyzed with the following indices: chi-square (χ^2), chi-square divided by degrees of freedom (χ^2/df), goodness of fit index (GFI), norm fit index (NFI), comparative fit index (CFI), and the root mean square error of approximation (RMSEA). Since the χ^2 analyses are sensitive to large sample sizes, most times the results will be statistically significant, so we use other fit indices that are not sensitive to large sample sizes, as proposed by Hu and Bentler (1999), to

evaluate model fit. The cut-off points value per index for a good model fit are the following: $GFI \geq .90$, $NFI \geq .90$, $CFI \geq .95$, and $RMSEA \leq .06$. All of the models had good model fits in most of the indices, except for the RMSEA, which saw values ranging from 0.074 to 0.115. It can be concluded that the models are acceptable when including all of the participants or for each group.

Table III. Fit indices of the confirmatory factor analyses performed across all participants and by educational level and gender groups

	All	Elementary	Middle School	High School	Men	Women
χ^2	294.848	133.239	97.084	109.811	179.201	132.768
df	19	19	19	19	19	19
p	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
χ^2/df	15.518	7.013	5.110	5.780	9.432	6.988
GFI	0.969	0.970	0.958	0.931	0.958	0.967
NFI	0.953	0.960	0.896	0.878	0.936	0.951
CFI	0.956	0.965	0.914	0.896	0.942	0.957
RMSEA	0.079	0.074	0.086	0.115	0.091	0.077

Note. χ^2 =Chi-square; df=Degrees of Freedom; GFI=Goodness of Fit Index; NFI=Norm Fit Index; CFI=Comparative Fit Index; RMSEA=Root Mean Square Error of Approximation.

Factor loadings and uniqueness values were calculated for each item (see Table IV). A factor loading refers to how much an item contributes to a factor (Yong & Pearce, 2013) and uniqueness is the variance of an item that is due to measurement error and specific variance (Kline, 2013). An item should have a minimal factor loading value of 0.320; if the item has a lower value, then it should be discarded (Tabachnick & Fidell, 2007). The factor loadings for all of the items in all of the analyses were good, ranging from 0.495 to 0.757.

Table IV. Factor loadings and uniqueness values of the confirmatory factor analyses, and correlations between the two factors performed in all participants and by educational level and gender groups

	All	Elementary	Middle School	High School	Men	Women
Bullied Factor						
Item 2	0.650	0.644	0.701	0.581	0.638	0.660
Item 3	0.653	0.663	0.567	0.596	0.655	0.657
Item 4	0.686	0.703	0.654	0.563	0.688	0.682
Item 5	0.691	0.720	0.548	0.668	0.664	0.713
Bullies Factor						
Item 6	0.673	0.688	0.568	0.715	0.686	0.660
Item 7	0.717	0.736	0.597	0.757	0.719	0.711
Item 8	0.691	0.732	0.571	0.666	0.725	0.651
Item 9	0.692	0.685	0.602	0.726	0.699	0.677
Uniqueness						
Item 2	0.464	0.570	0.331	0.375	0.481	0.445
Item 3	0.291	0.382	0.300	0.128	0.351	0.222
Item 4	0.505	0.567	0.421	0.493	0.516	0.493
Item 5	0.397	0.421	0.436	0.316	0.452	0.344
Item 6	0.512	0.562	0.560	0.368	0.517	0.504
Item 7	0.306	0.377	0.318	0.185	0.325	0.286
Item 8	0.554	0.541	0.547	0.633	0.516	0.591
Item 9	0.419	0.506	0.416	0.292	0.447	0.383
Correlation between factors	0.828	0.882	0.675	0.723	0.798	0.858

The final step was to analyze the measurement invariance for the factor structure across men and women. Measurement invariance was assessed by calculating the model fit of the scale for both men and women. Since both samples had good model fit indices for the scale, both scales were analyzed by constraining parameters across both samples (see Table V).

Table V. Fit indices for nested models across men and women

Model	χ^2	df	GFI	NFI	CFI	ΔCFI^a	RMSEA
Confirmatory Factor Analysis (Men Sample).	179.201**	19	0.958	0.936	0.942	--	0.091
Confirmatory Factor Analysis (Women Sample).	132.768**	19	0.967	0.951	0.957	--	0.077
Model 1: Both models unconstrained.	311.969**	38	0.962	0.943	0.950	--	0.060
Model 2: Measurement weights constrained.	323.137**	44	0.961	0.941	0.949	0.001	0.056
Model 3: The constraint of model 2 plus measurement intercepts, structural weights, structural intercepts, structural means, and structural covariances constrained.	332.692**	47	0.959	0.939	0.947	0.003	0.055
Model 4: All the constraints of model 3 plus structural residuals and measurement residuals constrained.	407.504**	55	0.951	0.926	0.935	0.015	0.056

Note. χ^2 =Chi-square; df=Degrees of Freedom; GFI=Goodness of Fit Index; NFI=Norm Fit Index; CFI=Comparative Fit Index; ΔCFI =Difference in between baseline model and nested models; RMSEA=Root Mean Square Error of Approximation.

--=Value not calculated.

^a= Nested models were compared to Model 1.

** $p < .01$

For the first model, the model fit of the scale was analyzed without constraints between the two samples. For the second model the measurement weights were constrained across both samples. In the third model included the constraint of model two and in addition, measurement intercepts, structural weights, structural intercepts, structural means, and structural covariances were constrained. In the fourth and final model, the constraints of model three were present, and in addition, structural residuals and measurement residuals were constrained. In these analyses Cheung & Rensvond (2002) recommend using the difference in the comparative fit index (ΔCFI) as an index to assess invariance between nested models and the baseline model. The CFI index is not sensitive to sample size or non-normal data. They suggest that a ΔCFI of less than or equal to 0.01 indicate a similar model fit (Cheung & Rensvond, 2002). Our results indicate partial measurement invariance since the difference in the comparative fit indices of the fourth model compared to the baseline model was 0.015, which is greater than the 0.01 suggested.

IV. Discussion and conclusions

Bullying is a social problem that originates primarily in homes with an abuse of power, violence, and communication problems (Fatima & Khatoon, 2015), and its consequences can result in criminality, substance use and abuse, and suicide (Fleming & Jacobsen, 2009). According to the National Survey for Health and Nutrition (Instituto Nacional de Salud Pública, 2006; 2012), in Mexico there has been an increase in bullying, and therefore this study evaluates the self-report Bull-M questionnaire to evaluate the prevalence of bullying among students. This instrument was previously tested and validated in a previous study with students from 7th to 9th grade from northern Mexico (Ramos-Jimenez et al., 2013). This study extends the sample age, from 5th to 12th grade, and was administered to a gender-balanced sample. The sample was randomly chosen considering the all schools in Ciudad Juárez as listed by the SNIEE.

The main results indicate that the Bull-M questionnaire is a valid instrument to evaluate anonymously the prevalence of aggressors and victims of bullying among students from different grades and by gender. The internal reliability analyses in our sample show adequate Cronbach's alpha values (0.77 to 0.87) for the total scale across all groups; however, when the reliabilities are analyzed by factor, there is a decrease in the values, which range from 0.67 to 0.80.

The confirmatory factor analyses performed with the total sample and all other groups indicated adequate model fit indices. All of the chi-square analyses were significant, which translates into a bad model fit, but this can be explained by the large sample sizes, according to Hu & Bentler (1999); for this reason there are other fit indices included in the analyses. The factor loadings for all of the items in both factors are strong, ranging from 0.55 to 0.76.

There is a high correlation between the bullied and bullies factor ($r=0.828$), which might suggest a unidimensional solution. A two-factor solution was kept, as exploratory and confirmatory factor analyses in a previous validation study suggested and confirmed a two-factor solution. The two-factor structure was validated in this study but with a high correlation between factors. The two factors were kept separately because according to the theory of bullying, it identifies two groups of people: the bullies and the bullied. Another reason to keep the two-factor structure is that in recent years Ciudad Juárez has had high levels of social violence due to the war between drug cartels, which has had an impact on society. This violence has been reflected in some school settings, with people becoming victims or perpetrators of violence. The schools sampled for this study were in the most affected areas of Ciudad Juárez and for this reason, the high correlation between two factors indicates that people who report higher levels of being bullied also report higher levels of bullying.

Measurement invariance across boys and girls was assessed but the results indicated partial measurement invariance. The instrument is invariant in the factor structure and loadings, measurement intercepts, structural means and structural covariances, but it was not invariant in structural and measurement residuals, even though it has been argued that invariance across measurement residuals is an overly restrictive test (Byrne, 2004); for this reason, we suggest it can be used to compare results between boys and girls since it has the same representation of factors in both groups. Measurement invariance across educational levels was not assessed since the sample for middle school was small in comparison with the rest of the groups.

In this study only the factor structure and internal reliability were evaluated. There is a need to evaluate other types of validity like concurrent validity or discriminant validity. There is a need to analyze the correlation of the Bull-M with other scales that measure related or similar constructs. Second, the RMSEA values for all groups were not ideal, even though the other model fit indices showed ideal values. Third, the internal reliability alpha values were acceptable in most of the sample except for middle school and high school, which produced alpha values below 0.70. Fourth, the total sample size was 2,030 but when the sample was divided by education levels, high school had only 360 students. There is a need for a larger sample in high school and middle school, especially as larger sample sizes are necessary when calculating structure equation models. Fifth, there is a need to have samples obtained from different places in Mexico, including the south and center of the country. This sample was obtained from Ciudad Juárez, in northern Mexico.

According to the results, the Bull-M can be used at different educational levels, from 5th grade to 12th grade. The confirmatory factor analyses performed in the total sample, and by educational level and gender, indicate a good model fit with the bullied and bullies factor. Most of the model fit indices are adequate, the factor loadings for all of the items are strong and the Cronbach's alpha values indicate adequate internal reliability for both factors and the total scale. The Bull-M is an anonymous and brief instrument that can be used to diagnose bullying problems in schools in northern Mexico.

Validation of the Bull-M in samples that include children from 5th grade to 12th grade was necessary to evaluate bullying in schools in Ciudad Juárez. After the violence escalated in Ciudad Juárez, it became a considerable problem with social and psychological consequences. Schoolchildren started to accept violent behavior as normal, and teachers reported that sometimes students would pretend to be drug

dealers in recess at school. School administrators, government institutions and in particular civil society organizations started programs to intervene in schools to reduce violence among students.

When these interventions began, bullying assessments were not used or were very informal. The development and validation of the Bull-M provided a low-cost and rapid evaluation of violence at school. This scale has been used in several schools to assess bullying and will be promoted in the city and state government so that it can be applied in all schools to assess violence among students. This tool will provide quantitative information about the problem so that school authorities may decide if it is necessary to intervene with programs targeted at reducing violence. This scale can be validated in other populations to serve as a quick and low-cost tool to detect violence in schools.

References

- Alzahrani, H. A. (2012). Bullying among medical students in a Saudi medical school. *BMC Research Notes*, 5(335), 1-3.
- Bauman, S., Toomey, R. B., & Walker, J. L. (2013). Associations among bullying, cyberbullying, and suicide in high school students. *Journal of adolescence*, 36(2), 341-350.
- Bickmore, K. (1997). Teaching conflict and conflict resolution in school: (Extra-) curricular considerations. Sydney, Australia: Connections '97 International Social Studies Conference. Retrieved from <https://eric.ed.gov/?id=ED421587>
- Byrne, B. M. (2004). Testing for multigroup invariance using AMOS graphics: A road less traveled. *Structural Equation Modeling: A Multidisciplinary Journal*, 11(2), 272-300. doi:10.1207/s15328007sem1102_8
- Casey, S., Hayward, C., & Gowen, K. (2001). Middle-school girls' reports of peer victimization: Concerns, consequences, and implications. *Professional School Counseling*, 5, 138-147.
- Cerezo, F. (2006). Violencia y victimización entre escolares. El bullying: estrategias de identificación y elementos para la intervención a través del Test Bull-S [Violence and victimization among students. Bullying: Identification strategies and elements for intervention through the Bull-S test]. *Electronic Journal of Research in Educational Psychology*, 4(9), 333-352.
- Cheung, G. W. & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233-255.
- Craig, W., Harel-Fisch, Y., Fogel-Grinvald, H., Dostaler, S., Hetland, J., Simons-Morton, B., & Pickett, W. (2009). A cross-national profile of bullying and victimization among adolescents in 40 countries. *International Journal of Public Health*, 54(2), 216-224.
- Creswell, J. W. (2003). *Research design: qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage.
- Cristobal, E., Flavián, C., & Guinalú, M. (2007). Perceived e-service quality (PeSQ): Measurement validation and effects on consumer satisfaction and web site loyalty. *Managing Service Quality: An International Journal*, 17(3), 317-340. doi:10.1108/09604520710744326
- Crothers, L. M. & Levinson, E. M. (2004). Assessment of bullying: a review of methods and instruments. *Journal of Counseling and Development*, 82(4), 496-503.

- Due, P., Damsgaard, M. T., Lund, R. & Holstein, B. E. (2009). Is bullying equally harmful for rich and poor children?: A study of bullying and depression from age 15 to 27. *European Journal of Public Health, 19*(5), 464-469.
- Fatima, S. & Khatoon, S. (2015). Causes of students' aggressive behavior at secondary school level. *Journal of Literature, Languages and Linguistics, 11*, 49-65. Retrieved from <http://www.iiste.org/Journals/index.php/JLLL/article/view/23537/23818>
- Fleming, L. C. & Jacobsen, K. H. (2009). Bullying and symptoms of depression in Chilean middle school students. *The Journal of school health, 79*(3), 130-137.
- Gilmartin, B. G. (1987). Peer group antecedents of severe love-shyness in males. *Journal of personality, 55*(3), 467-489.
- Glover, D., Gough, G., Johnson, M., & Cartwright, N. (2000). Bullying in 25 secondary schools: Incidence, impact, and intervention. *Educational Research, 42*(2), 141-156.
- Hanley, A. J., & Gibb, B. E. (2011). Verbal victimization and changes in hopelessness among elementary school children. *Journal of clinical child and adolescent psychology, 40*(5), 772-776.
- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55.
- Instituto Federal Electoral [Federal Electoral Institute]. (2012). *Consulta infantil y juvenil 2012* [Infant and juvenile consultation 2012]. Mexico: Author.
- Instituto Federal Electoral [Federal Electoral Institute]. (2015). *Consulta infantil y juvenil 2015* [Infant and juvenile consultation 2015]. Mexico: Author.
- Instituto Nacional de Salud Pública [National Institute of Public Health] (2006). *Encuesta nacional de salud y nutrición 2006* [National Survey of Health and Nutrition]. Mexico: Secretaría de Salud.
- Instituto Nacional de Salud Pública [National Institute of Public Health]. (2012). *Encuesta nacional de salud y nutrición 2012: Resultados nacionales* [National Health and Nutrition Survey 2012: National results]. Mexico: Secretaría de Salud.
- Kline, R. B. (2013). Exploratory and confirmatory factor analysis. In Y. Petscher, C. Schattschneider, & Compton, D. (Eds.), *Applied quantitative analysis in the social sciences* (pp. 171-207). New York: Routledge.
- Levine, E. & Tamburrino, M. (2014). Bullying among young children: strategies for prevention. *Early Childhood Education Journal, 42*(4), 271-278.
- Nansel, T. R., Overpeck, M., Pilla, R. S., Ruan, W. J, Simons-Morton, B., & Scheidt, P. (2001). Bullying behaviors among us youth: Prevalence and association with psychosocial adjustment. *Journal of the American Medical Association, 285*(16), 2094-2100.
- Ozkal, U. U. (2011). Bullying among university students and the relationship between bullying and psychological problems (Master's thesis). Middle East Technical University, Ankara, Turkey.
- Patton, D. U., Hong, J. S, Patel, S. & Kral, M. J. (2015). A systematic review of research strategies used in qualitative studies on school bullying and victimization. *Trauma Violence Abuse, 18*(1), 3-16. doi:10.1177/1524838015588502

Ramos-Jimenez, A., Wall-Medrano, A., Esparza-Del Villar, O., & Hernández-Torres, R. P. (2013). Design and validation of a self-administered test to assess bullying (bull-M) in high school Mexicans: a pilot study. *BMC Public Health*, 13(334), 2-7. doi:10.1186/1471-2458-13-334.

Sawyer, A. L., Bradshaw, C. P., & O'Brennan, L. M. (2008). Examining ethnic, gender, and developmental differences in the way children report being a victim of "bullying" on self-report measures. *The Journal of Adolescent Health*, 43(2), 106-114.

Secretaría de Salud [Secretary of Health]. (2006). Extracto del informe nacional sobre violencia y salud [Extract from the national report on violence and health]. Mexico: Author.

Shapiro, E. & Heick, P. (2004). School psychologist assessment practices in the evaluation of students referred for social/behavioral/emotional problems. *Psychology in the Schools*, 41(5), 551-561.

Tyler, L. (1972). *Tests and measurements*. Cliffs, NJ: Prentice-Hall.

Valencia, N. & Chacon, A. (2013, January 5). Juarez shedding violent image, statistics show. CNN. Retrieved from <http://edition.cnn.com/2013/01/05/world/americas/mexico-juarez-killings-drop/index.html>

Yong, A. G. & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*, 9(2), 79-94. doi: 10.20982/tqmp.09.2.p079