

Validez concurrente de la escala Bullying GT

Yvette Bolaños Gramajo¹ & J. Andrés Gálvez-Sobral A.²

¹Estudiante de Maestría, Universidad del Valle de Guatemala y Analista de datos en Ministerio de Educación de Guatemala, ²Director del Centro de Investigaciones Educativas, Universidad del Valle de Guatemala
jagalvez@uvg.edu.gt

Bullying o acoso escolar

El Bullying es un fenómeno actual que tiene lugar en el contexto educativo y que merece especial atención en virtud de los efectos que puede ocasionar en el aprendizaje, rendimiento y salud mental de los estudiantes. La importancia de validar pruebas psicométricas se fundamenta en la premisa que una prueba que cumpla con los criterios básicos de confiabilidad y validez permitirán la generalización de resultados. En el campo de investigación, contar con instrumentos con estas características garantiza que se recogen datos que representan una muestra del constructo que se pretende medir y sobre la cual se realizarán generalizaciones que aportan al conocimiento que se posee sobre el fenómeno o constructo en Guatemala. Este documento se propone analizar la validez concurrente de la escala Bullying GT y a la investigación en marcha de este fenómeno en Guatemala.

El Bullying es reconocido por las manifestaciones de agresividad en el contexto escolar, es por esta razón que también se conoce como "acoso escolar". Investigaciones recientes siguen utilizando la definición que brinda Dan Olweus como precursor del término e investigador del fenómeno y a partir de dicha definición se inician algunos debates más recientes. El estudio sistemático de este fenómeno es reciente y data de la década de 1970 en los países Nórdicos. En la década siguiente, el fenómeno fue estudiado por Japón, Gran Bretaña, Holanda, Canadá, Estados Unidos y España. Posteriormente se reconoció como una manifestación que debía considerarse en el diseño de políticas públicas de casi todo el mundo (Olweus, 1993).

Como definición se propone que "Un estudiante es acosado o victimizado cuando está expuesto de manera repetitiva a acciones negativas por parte de uno o más estudiante". También es Bullying cuando un estudiante es golpeado, pegado, amenazado o encerrado en una habitación. Estas situaciones pueden tener

lugar frecuentemente y es difícil para el estudiante intimidado defenderse por sí mismo. También es Bullying cuando un estudiante es molestado repetidamente de forma negativa. Pero no es Bullying cuando dos estudiantes de la misma fuerza discuten o pelean" (Olweus, 1999). Gálvez-Sobral (2008), se refiere al Bullying como "...un fenómeno propio de niños y niñas en edad escolar, se refiere a una situación que involucra una víctima que por sus propios medios se considera indefensa ante acciones negativas que provocan otros alumnos agresores, la persecución es considerada por la víctima como sistemática y puede ser de carácter físico y psicológico."

La *acción negativa* se define como una acción que tiene lugar "...cuando alguien infringe, de manera intencionada, o intenta infligir mal o malestar a otra persona... el actor o autor utiliza su propio cuerpo o un objeto externo (incluso una arma) para infligir una lesión o un daño, relativamente grave, a otro individuo." Los elementos fundamentales que caracterizan este fenómeno son la presencia de una relación asimétrica donde un menor ejerce poder sobre otro. El poder se expresa en forma de agresión que puede tener lugar sin causa aparente. En esta relación asimétrica el sujeto en desventaja posee dificultad para defenderse (Olweus, 1993).

Recientemente, autores como Felix et. al (2011), promueven la importancia de retomar la definición original de Olweus para reconocer que el "desequilibrio de poder" es vital en el fenómeno. Los autores ponen en cuestionamiento la forma de evaluar el Bullying ya que consideran que las encuestas en papel y lápiz obvian este tema tan importante. A pesar de lo descrito, sigue siendo importante evaluar los fenómenos de agresión entre pares para poder reducir el impacto negativo que produce.

Tratándose de un fenómeno que involucra agresión, violencia y la interacción entre víctima-agresor, se caracteriza por ser de naturaleza reincidente y cíclica. Los ciclos pueden tener diferente

magnitud, alcance y duración y propiciar relaciones que cambian de expresión continuamente afectando la estabilidad emocional de los involucrados. Como consecuencia, personas que han sido víctimas pueden transformarse en agresores o fungir en ambos roles al mismo tiempo. Las causas de esta dinámica son al mismo tiempo las consecuencias provocadas por su incidencia y reincidencia. Como causas del fenómeno se mencionan las características físicas así como la falta de respaldo social, el afán del agresor de expresar su conducta agresiva así como la intimidación, la falta de supervisión en el aula, reglas ambiguas o generales, favoritismo, crianza y castigo de los menores en la familia (Gálvez-Sobral, 2011).

Las características de víctimas y agresores se tienden a generalizar, aunque existe cierto consenso en algunas generalidades para los sujetos en distintos roles. Las características de las víctimas son la prudencia, sensibilidad, introversión, timidez, personas inquietas, inseguras, con baja autoestima, depresivos, con ideas suicidas, a menudo solitarios sin buenos amigos y que se relacionan mejor con adultos que con sus iguales. Referente a los agresores una fuerte necesidad de dominar y someter a otros compañeros y salirse siempre con la suya, impulsividad, enfado fácil, no solidarios, desafiantes, agresivos, involucrados en actividades antisociales y delictivas, en general los estudiantes de género masculino son más fuertes que sus víctimas (Solberg & Olweus, 2003).

El estudio de las características de los sujetos en distintos roles dejó de ser relevante para Olweus ya que después de varias investigaciones el autor concluye que 75% de la población estudiada tiene características diferentes a su grupo que les harían propensos a ser victimizados y sin embargo no es el 75% de la población la que es víctima de bullying y por lo tanto no son las características personales las que provocan la agresión. En este sentido la clave del fenómeno está en el desbalance de poder y por lo tanto explican que la selección de la víctima se fundamenta en la ausencia de respaldo social que ésta tiene; el agresor en su afán de manifestar su conducta agresiva, busca dentro del grupo al estudiante a quien nadie ayudará y al decir nadie, se refiere a los compañeros, docentes, directivos e incluso padres de familia y para justificar la expresión de agresión, buscan en la víctima un atributo que explique su comportamiento. Los autores concluyen diciendo que el Bullying es un fenómeno basado en oportunidad Felix et. al (2011).

Por otra parte, consecuencias del Bullying tienen un impacto considerable tanto en la víctima como en el agresor así como en los sujetos que rodean a ambos afectando a nivel individual, grupal, a nivel de escuela, familiar y en los ámbitos sociales en los que se desenvuelven estos actores. Hoy en día es preocupación de la política que se orienta a mejorar el clima escolar. El impacto es complejo ya que según Martínez (2011), los sujetos implicados en una situación de maltrato son los espectadores, frente a una minoría que lo sufre, de tal manera que el

posicionamiento puede ser a favor o en contra. Sin embargo, los efectos que pueden desencadenarse son: Desensibilización, incompreensión, indefensión, ansiedad y culpabilidad. (Martínez, 2011). Los menores que son víctimas de Bullying pueden manifestar conductas depresivas, ansiedad, síntomas psicósomáticos e ideación suicida (McCombs, 2004; Kaltiala-Heino et al, 2000; Rigby, 2000; Craig, 1998; Boivin & Hodges, 2001; Rigby & Slee, 1999; Konishi et al, 2010).

Joffre-Velásquez, et al (2011), analizaron la prevalencia del Bullying y las variables asociadas con el riesgo de esta conducta. Para ello se aplicó un cuestionario que identificó la presencia o ausencia del rol de víctima, agresor o víctima-agresor. Los resultados de un análisis estadístico develaron que de un grupo de 688 alumnos con una media de 13.62 ± 0.96 años se identificaron 20.5% víctimas, 13.1% agresores y 27.4% víctimas - agresores. Los factores de riesgo tienen que ver con algún defecto físico, los padres consideran normal el problema, en los agresores la preferencia por programas televisivos violentos, tener amigos que pertenezcan a pandillas y para las víctimas-agresores destacó la combinación de factores inherentes a ambos grupos por separado Joffre-Velásquez, et al (2011).

El impacto trasciende del evento singular que los involucrados viven en ese momento, a través del tiempo con incidencia en el rendimiento educativo. Una investigación en Canadá con participación de 27,217 estudiantes de 15 años y 1,087 directores de establecimientos, encontró derivado de un análisis multinivel que el Logro en Matemáticas y Lectura está relacionado negativamente con escuelas con presencia de Bullying (Konishi et al, 2010).

El Bullying no es un fenómeno exclusivo de algunos países, culturas o nivel socioeconómico. Debido a que como una manifestación específica de la agresión y violencia, tiene lugar en todo momento y lugar. Después de 30 años, los países continúan investigando el fenómeno para tomar mejores decisiones en las políticas públicas que permitan disminuir y erradicar el fenómeno. En Guatemala se desarrolló la escala Bullying GT que se aplicó a estudiantes de 6to. primaria del municipio de Guatemala. Los resultados del análisis develaron un Alfa de Cronbach de 0.878 por lo que se concluyó que la construcción de la escala fue correcta con buena consistencia interna para la etapa inicial de la misma. Se observó capacidad de discriminación de los ítems (en 18 ítems el punto biserial superó 0.400; en 4 ítems 0.300 y 3 ítems debían modificarse). Como apreciación general se observó que cada uno de los ítems utilizados contribuyó a la confiabilidad de la prueba y que los cambios requeridos son mínimos (Gálvez-Sobral, 2008).

Para dar seguimiento a la investigación, en 2010 Gálvez-Sobral desarrolló una investigación cualitativa con grupos focales, conformados por estudiantes de los últimos años de magisterio en Guatemala. El objetivo fue conocer la perspectiva los docentes

en formación sobre el fenómeno Bullying desde dos puntos de vista: estudiantes y próximos docentes. Dentro de los hallazgos se encontró que de los cinco grupos, tres conocían sobre el Bullying y que únicamente uno lo había aprendido como parte de la malla curricular. Se evidenció el mito más común dentro del fenómeno: “las víctimas son escogidas por sus atributos” y con un diagrama y matriz de doble ingreso, se evidenció la complejidad del proceso que produce el Bullying dentro de un aula (Gálvez-Sobral, 2011).

En 2011, se desarrolló un estudio como producto de la aplicación de una escala de conductas agresivas y Bullying en Guatemala aplicado a una muestra nacional con representatividad a nivel departamental, evaluándose 18,780 estudiantes de 933 establecimientos educativos. Los resultados de la investigación demostraron que 34% de los estudiantes se reportaron como víctimas de Bullying, con mayor incidencia en niños que en niñas, con mayor incidencia en el área rural, en escuelas pequeñas y en estudiantes de sobre edad. Entre otros estudios respecto al tema, Sacayón (2003) destaca la agresión psicológica como la más común en Guatemala. El MINEDUC y la Universidad del Valle de Guatemala han realizado estudios sobre Bullying en institutos normales utilizando la escala de diversificado que develó que 18% de estudiantes se consideran víctimas de Bullying (Gálvez-Sobral, 2011).

Validez psicométrica

Los orígenes de la validez como tema importante en la medición psicológica se pueden observar en 1957 cuando la American Psychological Association [APA] publica un documento sobre las recomendaciones técnicas para los tests psicológicos. Luego, en 1966 el término “Estándares” para los tests aparece en la edición de una revista de 3 entidades American Educational Research Association [AERA], American Psychological Association [APA] y National Council on Measurement in Education [NCME]. A partir de los estándares y su divulgación se marca una nueva etapa en la preocupación por las consecuencias del uso de los tests (Padilla et al, 2006). La definición de validez de los Estándares de 1985, planteaba que “se refiere a la adecuación, significado y utilidad de las inferencias específicas hechas desde las puntuaciones en el test”. Luego en la edición de los Estándares de 1999, se reconoce las ventajas de hacer equivalente la justicia en los tests a una validez comparable para todas las personas y grupos y se resaltan las aportaciones de Samuel Messick en el debate sobre la incorporación al análisis de la validez de las consecuencias sociales del uso de los tests como una fuente más de evidencia dentro del proceso de validación.

Por su parte, autores como Cronbach y Meehl en 1955 hablaban de estudiar diferencias entre grupos, correlaciones entre subtests, estudios sobre la estructura interna del test, estudios factoriales, correlaciones con criterios externos, estudios longitudinales, estudios experimentales y estudios sobre el análisis del contenido. Luego Campbell y Fiske (1959) introducen los conceptos de

validez discriminante y convergente, y la matriz multimétodo-multirrasgo para evaluarlas. Para complementarlo con lo que Messick (1975) afirma sobre que toda la evidencia de validez debe ser de constructo y contempla un aspecto más referido a los valores sociales y las consecuencias éticas (Gómez & Hidalgo, 2002).

La definición de validez citada en la edición de los Estándares (edición de 1999 de la AERA, APA y NCME) se refiere a “un juicio integrado del grado en el que la evidencia empírica y la racionalidad teórica apoyan la suficiencia y conveniencia de las inferencias y acciones basadas en las puntuaciones en los tests u otros modos de evaluación” (Padilla et al, 2006).

La validez es vista como un argumento cuidadosamente estructurado, que reúne evidencia desde una variedad de recursos para apoyar o refutar las interpretaciones propuestas de las puntuaciones de un instrumento (Gómez & Hidalgo, 2002). Los resultados de cualquier medición psicométrica tienen significado o validez solamente en el contexto del constructo que se propone medir. Por lo tanto, la validez no es una propiedad del instrumento, sino una interpretación de los datos obtenidos por un determinado procedimiento, por lo que la validez de las puntuaciones del test debe ser establecida en cada uso que se haga del mismo. Por ello se puede decir que la validez no puede ser nunca comprobada, ya que se convierte en un ciclo de continua evaluación y revisión (Cook & Beckman, 2006).

Al respecto Messick (Gómez & Hidalgo, 2002) afirmaba que la validez no es del test o de la observación; lo que se valida son las inferencias derivadas de las puntuaciones del test o de otros indicadores, inferencias sobre el significado de las puntuaciones o la interpretación para propósitos aplicados y sobre las implicaciones para la acción, es decir, las consecuencias sociales y éticas.

Asimismo Messick (Cook & Beckman, 2006) señala como fuentes de evidencia a considerar en la validez:

- **Contenido:** relevancia y representatividad del “constructo” del test.
- **Sustantivo:** razones teóricas de la consistencia observada de las respuestas (acciones y procesos de pensamiento). Aquí se puede preguntar a un grupo de estudiantes que “piensen en voz alta” mientras responden un conjunto de preguntas.
- **Estructural:** configuración interna del test (confiabilidad) y dimensionalidad (análisis factorial); funcionamiento diferencial de los ítems.
- **Generalización:** grado en que las inferencias hechas a partir del test se pueden generalizar a otras poblaciones, situaciones o tareas. Este aspecto tiene especial importancia en la adaptación y/o traducción de escalas y tests de una cultura a otra.
- **Externo:** relaciones del test con otros tests y constructos (concurrente). Análisis de la utilidad de la medida.
- **Consecuencial:** consecuencias éticas y sociales del test. Evaluación del sesgo del test.

Para Kerlinger (2002), en su exposición sobre la validez, menciona que la validez de tipo concurrente se diferencia de la validez predictiva en la dimensión tiempo ya que la validez predictiva se orienta a futuro en tanto que la validez concurrente mide el criterio casi al mismo tiempo, el estatus presente de la persona que responde al instrumento. Para llevar a cabo una validez concurrente es necesario aplicar el instrumento y un criterio externo a la misma población. Para cada sujeto se toman por lo menos dos medidas concurrentes. La validez concurrente es el resultado de la correlación de los dos conjuntos de calificaciones. Tanto la validez predictiva o validez concurrente, conocidas como relacionadas con el criterio, están caracterizadas por la predicción sobre un criterio externo y por la verificación de un instrumento de medición, en el presente o futuro, contra un resultado o medida. La dificultad de esta validación está relacionada con el criterio mismo ya que es necesario que el criterio que se utilice para validar posea de antemano validez (Kerlinger, 2002).

Según Kerlinger, la definición de confiabilidad se enfoca en que si se mide el mismo conjunto de objetos una y otra vez, con el mismo instrumento de medición o uno comparable, se obtendrán similares resultados; si las medidas obtenidas a partir de un instrumento de medición son las verdaderas entonces hay falta de distorsión y un tercer enfoque investigar que tanto error de medición existe en un instrumento de medición (varianza sistemática o por azar), ya que dependiendo del grado de error en un instrumento de medición, el instrumento es poco confiable. La confiabilidad no es una característica del instrumento sino de los resultados, de las puntuaciones obtenidas de una muestra determinada, se calcula con cada muestra. De tal manera que decimos que el instrumento es confiable porque es posible medir a sujetos diferentes (Morales, 2007).

En este sentido y conociendo la importancia de la validez de los instrumentos, el propósito de este estudio fue iniciar el proceso de análisis de validez para el instrumento Bullying GT utilizando el análisis de validez concurrente para medir y utilizando criterio externo la prueba Bullying AVE de TEA Ediciones con baremos de Latinoamérica para responder a la pregunta ¿existe relación entre los resultados de la prueba Bullying GT y la prueba de Bullying AVE cuando se aplican estas pruebas a los mismos sujetos?

Material y método

Objetivo: Medir la validez concurrente de la prueba de Bullying GT.

Hipótesis Nula: No existe correlación entre los resultados de las pruebas de Bullying GT y Bullying AVE.

Hipótesis Alternativa: Existe correlación entre los resultados de las pruebas de Bullying GT y Bullying AVE.

Metodología y diseño: La validación de una prueba es un proceso continuo que garantiza que una prueba mide lo que debe medir,

el propósito para lo que fue creada. Aiken (1996) la define como "...el grado en el cual ésta mide aquello para lo que se diseñó. ...". Anastasi & Urbina (1998), al respecto señalan: "*La validez tiene que ver con lo que miden y con qué tan bien lo hacen; nos indican qué se puede inferir a partir de sus resultados.*" Agrega, "*La validación concurrente es adecuada para las pruebas que se emplean para diagnosticar el estado actual. ...*" Por su parte Martínez, Hernández, & Hernández (2006), hacen énfasis en "...la validación es un <proceso>, no una acción puntual y que la validez no es de un test, sino de las interpretaciones y decisiones que se toman con las puntuaciones de los test.". De esta cuenta, la investigación toma como propósito validar las decisiones sobre Bullying que se tomen con la prueba de Bullying GT.

Como se indicó en secciones anteriores una forma de validar una prueba es mediante la correlación. Según Kerlinger (2002), es posible hacer la correlación entre la variable dependiente con otra relacionada, empleando los mismos participantes en cada uno de los grupos experimentales (Kerlinger, 2002). En esta investigación se ejecutó un análisis cuantitativo para determinar la validez concurrente de la prueba Bullying GT. El criterio se aplicó 1 semana después, tomando en consideración llevarlo a cabo: "...aproximadamente al mismo tiempo que los resultados de la prueba o después de un intervalo establecido." (Anastasi, Urbina, 1998). Otros autores como Lynn, J. & Rupp, A. (2010) señalan lo siguiente: "Validity refers to the degree to which a measure accurately taps the specific construct that it claims to be tapping. Criterion-related validity is concerned with the relationship between individuals' performance on two measures tapping the same construct. It typically is estimated by correlating scores on a new measure with scores from an accepted criterion measure. There are two forms of criterion-related validity: predictive validity and concurrent validity. Concurrent validity focuses on the extent to which scores on a new measure are related to scores from a criterion measure administered at the same time, whereas predictive validity uses the scores from the new measure to predict performance on a criterion measure administered at a later time" (Lynn, Rupp, 2010). Se utilizó un diseño correlacional. Eisman et al (1998), se refieren al uso de esta técnica: "...la correlación se usa precisamente para establecer la intensidad de la relación entre dos variables en momentos distintos y así poder inferir cierto grado de causalidad entre dichas variables." Para ello, se mide a un mismo grupo de sujetos en dos momentos diferentes (Eisman, Colás, Hernández, 1998). Lo indicado permitió establecer la relación temporal entre el criterio externo y la prueba para establecer la validación concurrente. El criterio utilizado fue la prueba Bullying AVE.

Muestra: Se utilizó un muestreo no probabilístico intencional. El muestreo probabilístico: "...signifies that one sees sampling as a series of strategic choices about with whom, where, and how one does one's research. This statement implies that the way that researchers sample must be tied to their objectives." Avila, (2006) define el muestreo intencional como "...un procedimiento que permite seleccionar los casos característicos de la población limitando la muestra a estos casos. Se utiliza en situaciones en las que la población es muy variable y consecuentemente la muestra es muy pequeña. La elección de este método se elige

cuando la pregunta del investigador es más importante que el método de selección de los sujetos (Dolores, Tongco, 2007).

Población: Se evaluó a un total de 248 estudiantes de 5to y 6to grado de primaria de cinco establecimientos elegidos de forma intencional del departamento de Guatemala y Quetzaltenango. El 65.73% de género femenino y 34.27% masculino. La moda poblacional de edad se observó en 12 años (45.53%), seguida por estudiantes de 13 años (33.74%), de 11 años (4.07%) y 14 años (1.22%). El 15.45% de estudiantes no indicó su edad.

Instrumentos: Se utilizaron dos instrumentos que miden el grado de Bullying de un estudiante:

- **Bullying GT:** Prueba diseñada por DIGEDUCA/MINEDUC y CIE/UVG para identificar el grado de agresión ejercida de un menor contra otro en el ámbito escolar. Con esta prueba se obtiene una medida de Bullying en 3 escalas de Bullying (Víctima de Bullying, Víctima de Agresión, Víctima de Agresión Leve), 12 Indicadores de Diferencia (con relación a la edad, estatura, peso, dinero, color de piel, idioma, costumbres, gustos, deportes, inteligencia, religión, cultura) y un Índice de diferencia percibida respecto al grupo. La consistencia interna de la prueba fue de 0.878. La correlación punto biserial develó que 18 ítems de la prueba se encuentran por arriba de 0.400, cuatro ítems arriba de 0.300. (Gálvez-Sobral, 2008)
- **Bullying AVE:** Prueba diseñada por TEA Ediciones que mide las conductas de hostigamiento y acoso en el entorno escolar y en las relaciones entre iguales. Con la prueba se obtienen 22 indicadores (2 indicadores globales -Índice Global de Acoso, Intensidad de Acoso-, 8 indicadores de violencia y acoso -hostigamiento, intimidación, amenazas, coacciones, bloqueo social, exclusión social, manipulación social, agresiones-, 4 factores globales de acoso -hostigamiento, intimidación, exclusión, agresiones- y 8 escalas clínicas -ansiedad, estrés postraumático, distimia, disminución de la autoestima, flashbacks, somatización, autoimagen negativa, autodesprecio).

Procedimiento: Una vez elegida la población evaluar, se aplicó la prueba Bullying GT y posteriormente la prueba de Bullying AVE

a los estudiantes de 5to y 6to primaria de establecimientos de Ciudad Capital y Quetzaltenango. Todos los estudiantes de los establecimientos respondieron ambas pruebas con una 1 semana de diferencia entre cada prueba. Los resultados de ambas pruebas fueron digitalizados en una base de datos que fue objeto de un proceso de verificación y limpieza. Las pruebas fueron calificadas según el procedimiento específico de cada una. Se realizó un análisis de correlación de Pearson entre las variables Total de Bullying (Prueba Bullying GT) e Índice Global de Acoso (Prueba Bullying AVE).

Definición conceptual de variables

• **Bullying:** Gálvez-Sobral (2008), se refiere al Bullying como "...un fenómeno propio de niños y niñas en edad escolar, se refiere a una situación que involucra una víctima que por sus propios medios se considera indefensa ante acciones negativas que provocan otros alumnos agresores, la persecución es considerada por la víctima como sistemática y puede ser de carácter físico y psicológico." (Gálvez-Sobral, 2008).

Definición operacional de variables (id el nivel de medición)

• **Bullying GT:** Se definió como los resultados obtenidos en la variable Total de Bullying de la Prueba de Bullying GT.

• **Bullying AVE:** Se definió como los resultados obtenidos en la variable Índice Global de Acoso de la Prueba Bullying AVE.

Resultados

En la Tabla 1 se observa la validez concurrente producto de una correlación de Pearson (bilateral) entre las variables de total de ambas pruebas: Total Bullying GT e Indicador Global de Acoso AVE. La correlación resultante obtuvo un indicador de 0.759 y estadísticamente significativo (0.001 bilateral). La correlación puede considerarse como fuerte entre el instrumento en estudio y el instrumento utilizado como criterio externo.

Tabla 1. Validez concurrente entre el indicador de Bullying Total (Prueba Bullying GT y criterio externo Bullying AVE)

		BULLYING GT: TOTAL
PERFIL AVE: INDICE GLOBAL DE ACOSO (M)	Correlación de Pearson	.759
	Sig. (bilateral)	.000
	N	248

Tabla 2. Validez concurrente entre la variable que indica Bullying Moderado (Prueba Bullying GT y criterio externo Bullying AVE)

		BULLYING GT: VICTIMA DE AGRESIÓN MODERADA
BULLYING AVE: VICTIMA DE AGRESIÓN MODERADA	Correlación de Pearson	.373
	Sig. (bilateral)	.012
	N	45

Tabla 3. Validez concurrente entre la variable que indica Bullying Leve (Prueba Bullying GT y criterio externo Bullying AVE)

		BULLYING GT: VICTIMA DE AGRESIÓN LEVE
BULLYING AVE: VICTIMA DE AGRESIÓN LEVE	Correlación de Pearson	.294
	Sig. (bilateral)	.000
	N	192

En la Tabla 2 se observa la validez concurrente entre las variables indicadoras de agresión leve, entre las cuales se observó un coeficiente de 0.373.

En la Tabla 3 se observa la validez concurrente entre las variables indicadoras de agresión leve, entre las cuales se observó un coeficiente de 0.294.

Discusión y conclusiones

Con relación a la muestra seleccionada es oportuno indicar que se trata de una muestra intencional en Guatemala y Quetzaltenango. Los resultados del presente análisis deben acotarse y referirse a esta población que corresponde a 248 estudiantes de 3ro. y 6to. primaria, con una proporción de 6 por cada 10 estudiantes de género femenino y entre 12-13 años de edad.

El análisis de confiabilidad de la prueba de Bullying GT fue mayor que la observada en la medición inicial al momento de su creación. Según lo indicado por Kerlinger (2002), Anastasi, Urbina (1988) y Aiken (1996) la confiabilidad resultante devala consistencia, reproductibilidad, falta de distorsión de la prueba. Lo anterior se debe a que el coeficiente es alto (Mayor a 0.80) y se considera que un coeficiente cercano a 1 y permite generalizar que con este mismo instrumento se obtienen similares resultados, y que no existe varianza sistemática o por azar.

La confiabilidad resultante indica que la prueba GT permite dar significado a los resultados obtenidos por ella y permiten generalizar sobre Bullying aludiendo a la estabilidad temporal

encontrada (Abad, et. al, 2006). Por lo anterior sabemos que la prueba mide con precisión. En esta investigación esto sucedió con la muestra a la que fue aplicada la prueba.

Los análisis de validez de un instrumento son importantes en la medida que se pueda interpretar el uso que se da a los resultados obtenidos con la misma y especialmente a las decisiones que se tomen con ellos. Este ejercicio de validación se enmarca en uno de un proceso permanente que como indican Cook y Beckman (2006), garantizan que la prueba mide lo que debe medir y las conclusiones derivadas de dicha medición son relevantes y significativas para el propósito con el que fue creado. Los resultados de la validez concurrente aplicando el método de correlación de Pearson entre los resultados de los mismos sujetos, confirman una correlación de .749, que se considera moderada entre la prueba Bullying GT y el criterio utilizado, prueba Bullying AVE. Se encontró correlación entre el total de los puntajes obtenidos por los mismos estudiantes, utilizando ambas pruebas con un lapso no mayor a 1 semana.

Dado que la prueba Bullying GT posee indicadores de Bullying según la presencia y magnitud del fenómeno experimentado por cada sujeto, interés establecer la validez concurrente con el indicador de magnitud de la prueba Bullying AVE. Se observó que los estudiantes con puntuaciones que los ubican en una escala moderada de Bullying correlacionan con el nivel moderado de la prueba Bullying AVE en un coeficiente de 0.373 y los estudiantes que por sus resultados se ubicaron en un nivel leve en la prueba de Bullying GT correlacionan con un coeficiente de 0.294 con el criterio externo siendo ambos indicadores significativos.

Los resultados observados en la correlación permiten concluir de acuerdo a Campbel, Fiske (1959) y Cook y Beckman (2006) que la prueba posee validez externa y concurrente debido a que correlaciona positivamente con una medida obtenida por un test independiente que mide el mismo constructo, mostrando así la utilidad de la medida nueva. Estos resultados confirman sobre el continuo de evaluación que la prueba Bullying GT permite hacer inferencias sobre el constructo que pretende evaluar. A pesar de los resultados positivos, es importante tomar en cuenta las limitaciones en este estudio que surgen al comprar la prueba Bullying GT con una prueba elaborada en España para el contexto Español.

Bibliografía

- Aiken L (1996) *Rating Scales and Checklists: Evaluation Behavior, Personality, and Attitudes* United States of America: Wiley.
- Anastasi A, Urbina S (1998) *Test Psicológicos* México: Pearson Educación
- Avila H (2006) *Introducción a la metodología de la investigación* Retrieved 2015, from <http://www.eumed.net/libros-gratis/2006c/203/21.htm>
- Cook DA, Beckman TJ (2006) *Current concepts in validity and reliability for psychometric instruments* American Journal of Medicine **119** (2) e7-16
- Dolores M, Tongco C (2007) *Purposive sampling as a tool for informant selection* Ethnobotany Research Applications **5**:147-158
- Eisman L, Colás M, Hernández F (1998) *Métodos de investigación en psicopedagogía* McGraw-Hill
- Felix E, Sharkey J, Green J, Furlong MJ, Tanigawa D (2011) *Getting Precise and Pragmatic about the Assessment of Bullying: The Development of the California Bullying Victimization Scale* Aggressive Behavior **234**-247
- Gálvez-Sobral J (2008) *Bullying en la Ciudad de Guatemala*. Guatemala: Ministerio de Educación: DIGEDUCA
- Gálvez-Sobral J (2011) *Bullying. El fenómeno del acoso escolar en Guatemala. Resultados de una muestra nacional de sexto grado del Nivel Primario*. Guatemala: Ministerio de Educación: DIGEDUCA
- Gálvez-Sobral J (2011) *Bullying. La percepción de los futuros docentes en Guatemala*. Guatemala: Ministerio de Educación: DIGEDUCA.
- Gómez B, Hidalgo M (2002) La validez de los test, escalas y cuestionarios. *La sociología en los Escenarios* No.8
- Joffre-Velásquez V, García-Maldonado G, Saldívar-González A, Martínez-Perales G, Lin-Ochoa D, Quintanar-Martínez S, Villasana-Guerra A (2011) *Bullying en alumnas de secundaria. Características generales y factores asociados al riesgo* Bol Med Hosp Infant **68** (3): 193-202
- Kerlinger N (2002) *Investigación del Comportamiento* México: Mac Graw Hill
- Konishi Ch, Hymel Sh, Zumbo B, Li Z (2010) *Do School Bullying and Student-Teacher Relationships Matter for Academic Achievement? A Multilevel Analysis* Canadian Journal of School Psychology **25** (1) 19-39
- Lynn J, Rupp A (2010) *SAGE Research methods* Retrieved 2015, from <https://srmo.sagepub.com/view/encyc-of-research-design/n67.xml>
- Martínez A (2011) *Incidencia del acoso escolar en el rendimiento académico* España: Facultad de Ciencias de la Educación
- Martínez M, Hernández M, Hernández V (2006) *Psicometría* España: Alianza Editorial, S.A.
- McCombs (2004) Kalliala-Heino, Rimpela, Rantanen, Rimpela (2000) Rigby (2000) Craig (1998) Boivin, Hodges (2001) Rigby, Slee (1999) In: Konishi Ch, Hymel Sh, Zumbo B, Li Z (2010) *Do School Bullying and Student-Teacher Relationships Matter for Academic Achievement? A Multilevel Analysis* Canadian Journal of School Psychology **25** (1) 19-39
- Morales P (2007) *La fiabilidad de los tests y escalas* España: Madrid: Universidad Pontificia Comillas: Facultad de Ciencias Humanas y Sociales
- Olweus D (1993) *Acoso Escolar "Bullying"* En: Las Escuelas: Hechos e Intervenciones Centro de investigación para la Promoción de la Salud, Universidad de Bergen
- Olweus D (1999) In: PK Smith, Y Morita, J Junger-Tas, D Olweus, R Catalano, P Slee (1999) *The nature of school bullying: A cross-national perspective* Routledge 22-27
- Padilla JP et al (2006) *La evaluación de las consecuencias del uso de los tests en la teoría de validez* Psicothema **18** (2): 307-312
- Palys T In LM Given (Ed) (2008) *Purposive sampling* The Sage Encyclopedia of Qualitative Research Methods 697-8
- Solberg M, Olweus D (2003) *Prevalence estimation of school bullying with the Olweus Bully/Victim Questionnaire* Aggressive Behavior **239**-268