



# Universidad Autónoma de Baja California

Instituto de Investigación y Desarrollo Educativo

Evidencias de validez de constructo de la Escala del Rol del  
Estudiante en el Ciclo de la Violencia Escolar (ERECVE)

**TESIS**

Que para obtener el grado de

***MAESTRO EN CIENCIAS EDUCATIVAS***

Presenta

Luis Jorge Sánchez Ariza

Director de Tesis

Juan Carlos Pérez Morán

Comité Tutorial

Joaquín Caso Niebla

Juan Carlos Rodríguez Macías

Ensenada, B. C., México, 24 de abril de 2019.



**Universidad Autónoma de Baja California**  
Instituto de Investigación y Desarrollo Educativo  
Maestría en Ciencias Educativas



**“Evidencias de validez de constructo de la Escala  
del Rol del Estudiante en el Ciclo de la Violencia  
Escolar (ERECVE)”**

**TESIS**

Que para obtener el grado de  
**MAESTRA EN CIENCIAS EDUCATIVAS**

Presenta

**Luis Jorge Sánchez Ariza**

APROBADO POR:

**Dr. Juan Carlos Pérez Morán**  
Director(a) de tesis

---

**Dr. Joaquín Caso Niebla**  
Sinodal

---

**Dr. Juan Carlos Rodríguez macías**  
Sinodal



Ensenada, B.C. a 19 de febrero de 2019

**ASUNTO:** Voto aprobatorio sobre trabajo de tesis de grado de Maestría.

**Dr. José Alfonso Jiménez Moreno**  
**Coordinador de la Maestría en Ciencias Educativas**  
**Presente.**

Después de haber efectuado una revisión minuciosa sobre el trabajo de tesis presentado por el **C. Luis Jorge Sánchez Ariza** para poder presentar la defensa de su examen y obtener el grado de Maestría en Ciencias Educativas, me permito comunicarle que he dado mi VOTO APROBATORIO, sobre su trabajo intitulado:

**“Evidencias de validez de constructo de la Escala del Rol del Estudiante en el Ciclo de la Violencia Escolar (ERECVE)”.**

Esperando reciba el presente de conformidad, quedo de Usted.

Atentamente

Una firma manuscrita en tinta azul, que parece ser la del Dr. Juan Carlos Pérez Morán.

---

Dr. Juan Carlos Pérez Morán



Ensenada, B.C. a 19 de febrero de 2019

**ASUNTO:** Voto aprobatorio sobre trabajo de tesis de grado de Maestría.

**Dr. José Alfonso Jiménez Moreno**  
**Coordinador de la Maestría en Ciencias Educativas**  
**Presente.**

Después de haber efectuado una revisión minuciosa sobre el trabajo de tesis presentado por el **C. Luis Jorge Sánchez Ariza** para poder presentar la defensa de su examen y obtener el grado de Maestría en Ciencias Educativas, me permito comunicarle que he dado mi **VOTO APROBATORIO**, sobre su trabajo intitulado:

**“Evidencias de validez de constructo de la Escala del Rol del Estudiante en el Ciclo de la Violencia Escolar (ERECVE)”.**

Esperando reciba el presente de conformidad, quedo de Usted.

Atentamente

Dr. Joaquín Caso Niebla



Ensenada, B.C. a 19 de febrero de 2019

**ASUNTO:** Voto aprobatorio sobre trabajo de tesis de grado de Maestría.

**Dr. José Alfonso Jiménez Moreno**  
**Coordinador de la Maestría en Ciencias Educativas**  
**Presente.**

Después de haber efectuado una revisión minuciosa sobre el trabajo de tesis presentado por el **C. Luis Jorge Sánchez Ariza** para poder presentar la defensa de su examen y obtener el grado de Maestría en Ciencias Educativas, me permito comunicarle que he dado mi VOTO APROBATORIO, sobre su trabajo intitulado:

**“Evidencias de validez de constructo de la Escala del Rol del Estudiante en el Ciclo de la Violencia Escolar (ERECVE)”.**

Esperando reciba el presente de conformidad, quedo de Usted.

Atentamente

Dr. Juan Carlos Rodríguez Macías

## DEDICATORIA

*A mi papá Luis Nemesio Ariza Pérez, por ser siempre una*

*inspiración en mi vida, tanto personal como académica.*

*A mamá Noemi, mamá Sary, Paco, Kiyoko, Asami y*

*Luis Enrique, familia con la que la vida me bendijo.*



## AGRADECIMIENTOS

Agradezco a la vida por la existencia y a mis abuelos Luis y Noemí, a mi madre Sara, mi hermano Paco, mi prima Asami y mis tíos Kiyoko y Luis, por apoyarme en cada momento de mi vida, inculcándome que con trabajo arduo se alcanzan los sueños.

A mi director de tesis, tutor y mentor, el Dr. Juan Carlos Pérez Morán, a quien admiro, quiero y respeto mucho. Agradezco su amistad, paciencia y todo su apoyo incondicional en cada una de las etapas de mi formación, a lo largo de mis estudios de posgrado.

A quienes han fungido como sinodales, profesores, mentores y amigos incondicionales a través de estos nueve años colaborando en el IIDE: el Dr. Juan Carlos Rodríguez Macías y el Dr. Joaquín Caso Niebla, gracias infinitas por su cariño, comprensión, apoyo y acompañamiento desde el inicio de esta aventura.

A cada uno de mis profesores de la maestría, por su invaluable apoyo, especialmente a la Dra. Edna Luna, a la Dra. Graciela Cordero, al Dr. Jiménez, a la Dra. Maricela López Ornelas, a la Dra. Guadalupe Tinajero, al Dr. Javier Organista y a la Dra. Alicia Chaparro, quienes compartieron conmigo su amplio conocimiento durante mis estudios de posgrado.

A mis queridos amigos de la UEE: Zulma, Rafael, Carlitos, Neftali y Diana por sus innumerables muestras de apoyo y apuestas en este proceso formativo. Por cierto, ¡LES GANÉ!

A mis amigos y personas que amo con todo mi corazón, que siempre están conmigo acompañándome, y compartiendo sus experiencias y ayuda; principalmente por mostrarme que el amor es siempre un motor para vivir y salir adelante.

A todas las personas del equipo administrativo y de intendencia del IIDE, en especial al Mtro. Julio Cano, al Mtro. Iván Contreras, a la Mtra. Estrella Velasco, a Denisse, a Rosalva, a Yesica y a Alejandra, quienes, con su trabajo profesional y entrega a la mejora de la calidad educativa, facilitaron los procesos administrativos en el posgrado.

A todos los que estuvieron conmigo siempre, a mis amigos y compañeros de la Stella Pizza y Breve Café, al Valedor, y a mis compañeros de maestría y doctorado, gracias, gracias totales.

A CONACYT por becarme y permitirme estudiar en la Maestría en Ciencias Educativas del Instituto de Investigación y Desarrollo Educativo.

*Luis Jorge Sánchez Aríza*

# TABLA DE CONTENIDOS

DEDICATORIA .....	I
AGRADECIMIENTOS .....	II
ÍNDICE DE TABLAS .....	V
ÍNDICE DE FIGURAS .....	VI
RESUMEN .....	VII
<b>I. INTRODUCCIÓN .....</b>	<b>1</b>
1.1. ANTECEDENTES .....	3
1.2. PLANTEAMIENTO DEL PROBLEMA .....	10
1.3. OBJETIVOS DE INVESTIGACIÓN .....	12
1.4. JUSTIFICACIÓN .....	13
1.5. ESTRUCTURA DE LA TESIS .....	14
<b>II. MARCO TEÓRICO .....</b>	<b>16</b>
2.1. TEORÍA Y MEDICIÓN DE LOS ROLES EN EL CICLO DE LA VIOLENCIA ESCOLAR .....	16
2.1.1. <i>Teoría de los roles en el ciclo de la violencia escolar</i> .....	18
2.1.2. <i>Instrumentos de medición de los roles en el ciclo de la violencia escolar</i> .....	21
2.2. EVIDENCIAS DE VALIDEZ DE CONSTRUCTO .....	26
2.2.1. <i>Conceptos y enfoques de la validez de constructo</i> .....	28
2.2.2. <i>Evidencias de validez de constructo de estructura interna y relación con otras variables</i> .....	31
2.3. MÉTODOS ESTADÍSTICOS PARA EL ANÁLISIS DE LA ESTRUCTURA INTERNA DE INSTRUMENTOS DE MEDICIÓN .....	36
2.3.1. <i>Análisis Factorial Exploratorio (AFE)</i> .....	39
2.3.2. <i>Análisis Factorial Confirmatorio (AFC)</i> .....	47
2.3.3. <i>Software para el Análisis Factorial</i> .....	51
<b>III. MÉTODO .....</b>	<b>57</b>
3.1. TIPO DE ESTUDIO .....	57
3.2. INSTRUMENTO .....	58
3.3. BASE DE DATOS .....	60
3.4. PROCEDIMIENTO .....	60
3.4.1. <i>Etapa I. Análisis preliminares</i> .....	62
3.4.2. <i>Etapa II. Análisis Factorial Exploratorio</i> .....	65
3.4.3. <i>Etapa III. Análisis Factorial Confirmatorio</i> .....	69
3.4.4. <i>Etapa IV. Análisis de correlación convergente y discriminante</i> .....	74
<b>IV. RESULTADOS .....</b>	<b>75</b>



4.1. RESULTADOS DEL ANÁLISIS PRELIMINAR .....	75
4.1.1. <i>Resultados del análisis descriptivo</i> .....	75
4.1.2. <i>Resultados del análisis de normalidad, linealidad y multicolinealidad</i> .....	79
4.1.3. <i>Análisis de confiabilidad y discriminación de la ERECVE</i> .....	82
4.2. RESULTADOS DEL ANÁLISIS FACTORIAL EXPLORATORIO (AFE) .....	84
4.3. RESULTADOS DEL ANÁLISIS FACTORIAL CONFIRMATORIO (AFC) .....	87
4.4. RESULTADO DEL ANÁLISIS DE CORRELACIÓN CONCURRENTES Y DISCRIMINANTE.....	89
<b>V. DISCUSIÓN Y CONCLUSIONES.....</b>	<b>91</b>
5.1. LOGROS DEL ESTUDIO .....	91
5.2. LIMITACIONES DEL ESTUDIO .....	93
5.3. RECOMENDACIONES PARA FUTUROS ESTUDIOS .....	94
<b>VI. REFERENCIAS .....</b>	<b>96</b>
APÉNDICE 1. DESCRIPTIVOS DE LAS PREGUNTAS DE LA ESCALA DE LOS ROLES DE LOS ESTUDIANTES EN EL CICLO DE LA VIOLENCIA ESCOLAR .....	108
APÉNDICE 2. ESTIMACIONES DE LOS COEFICIENTES DE CORRELACIÓN DE SPEARMAN ENTRE LOS ÍTEMS DE LA ERECVE .....	111

## ÍNDICE DE TABLAS

Tabla 1. Clasificación de los índices de bondad de ajuste en el AFC .....	51
Tabla 2. Ítems y sub-escalas que conforman la Escala del Rol de los Estudiantes en el Ciclo de la Violencia Escolar (ERECVE) .....	59
Tabla 3. Modelo de etapas y actividades .....	61
Tabla 4. Descriptivos de las 19 preguntas de la Escala de los Roles de los Estudiantes en el Ciclo de la Violencia Escolar (ERECVE) .....	76
Tabla 5. Descriptivos de la base de datos sin casos atípicos .....	81
Tabla 6. Confiabilidad del Alpha de Combrach si se elimina el ítem y correlación ítem-total corregida de la ERECVE .....	83
Tabla 7. Varianza explicada resultado de la aplicación del método de extracción de componentes principales.....	85
Tabla 8. Matriz factorial de la ERECVE.....	86
Tabla 9. Índices de bondad de Ajuste del modelo factorial confirmatorio estimado .....	87
Tabla 10. Estimaciones del modelo factorial tanto exploratorio como confirmatorio reespecificados .....	88
Tabla 11. Correlaciones entre las subescalas y la ERECVE en su conjunto, así como con otras variables relacionadas con los temas la violencia intrafamiliar y el clima y la violencia escolar.....	90

## ÍNDICE DE FIGURAS

<i>Figura 1.</i> Definición gráfica del modelo factorial confirmatorio para el análisis de la estructura interna de la ERECVE .....	70
<i>Figura 2.</i> Código producido por el EQS para la formalización matemática del modelo factorial confirmatorio para el análisis de la estructura interna de la ERECVE .....	71
<i>Figura 3.</i> Distribución del puntaje total de la Escala de los Roles de los Estudiantes en el Ciclo de la Violencia Escolar (ERECVE) .....	77
<i>Figura 4.</i> Frecuencias de cada opción de respuesta de las preguntas S144.1, S144.6, S143.1, S144.9 y S143.8 .....	78
<i>Figura 5.</i> Identificación de casos extremos menores a 24 puntos y mayores a 44 .....	79
<i>Figura 6.</i> Gráfico Q-Q de linealidad normal de las puntuaciones de la ERECVE .....	81
<i>Figura 7.</i> Gráfica de sedimentación.....	85

## RESUMEN

Durante las últimas cuatro décadas, en el campo educativo el estudio y la medición de la violencia tomaron importancia debido a la problemática severa al interior de las aulas y los centros escolares, convirtiéndose en una prioridad para los sistemas educativos en el mundo (Olweus, 1983, 1991, 2001; Craig y Harel, 2004; UNISEF, 2005, 2009, 2011; UNESCO, 2017). En Baja California, investigadores de la Unidad de Evaluación Educativa (UEE) de la Universidad Autónoma de Baja California (UABC), interesados en aportar al estudio y la medición de la violencia y el acoso escolar desarrollaron la *Escala de roles del estudiante en el ciclo de la violencia escolar* (ERECVE). El propósito del presente estudio fue analizar las evidencias de validez de constructo en el aspecto de la estructura interna, y de la relación con otras variables de dicha escala. Para ello, se realizó un estudio psicométrico con la aplicación del Análisis factorial exploratorio (AFE), el Análisis factorial confirmatorio (AFC), así como el análisis de la correlación entre los ítems y factores de la escala, y las variables de violencia familiar, ambiente escolar violento, rol de participación del docente en el ciclo de la violencia escolar, y la aportación del rol del director al clima escolar. Los resultados del AFE y AFC del modelo explicativo de la estructura de la escala fueron satisfactorios. Con ello se confirma el modelo estructural de cinco factores, organizados en las categorías del enfoque de estudio del acoso escolar como proceso grupal, propuesto por Olweus (1993) y ampliado por Salmivalli (1999, 2010): víctima, agresor solitario, agresor social, observador, defensor.

**Palabras clave:** Validez de constructo, violencia, acoso escolar, educación secundaria.

---

## I. INTRODUCCIÓN

Desde la década de los setenta, el estudio y la medición de la violencia en el campo educativo adquirieron importancia, debido a la problemática severa que hay al interior de las aulas y de los centros escolares. Por tanto, se ha convertido en una prioridad de atención para los sistemas educativos en el mundo (Olweus, 1983, 1991, 2001; Rigby y Slee, 1991; Boulton y Underwood, 1992; Smith y Sharp, 1994; Craig, Peters y Konarski, 1998; Craig y Harel, 2004; UNICEF, 2005, 2009, 2011; Pepler, Jiang, Craig y Connolly, 2008; D'Angelo y Fernández, 2010; UNESCO, 2017). La Organización Panamericana de la Salud (OPS), la Oficina Sanitaria Panamericana (OSP) y la Oficina Regional de la Organización Mundial de la Salud (OMS), en su informe mundial sobre la violencia y la salud de 2003 mencionan que, en 27 países de distintas regiones del mundo, entre 30% y 40% de los niños en edad escolar reportan haber participado en riñas o intimidado a otros al menos por algún tiempo (OPS, OSP y OMS, 2003). Por su parte, la UNESCO, en su Reporte del estatus global de la violencia y el acoso escolar de 2017, informa que aproximadamente 246 millones de niños y adolescentes experimentan violencia y acoso escolar, de alguna manera u otra, cada año (UNESCO, 2017).

Recientemente UNICEF, en colaboración con el Representante Especial del Secretario General de las Naciones Unidas sobre la Violencia contra los Niños (SRSG-VAC, por sus siglas en inglés), reportaron datos preocupantes en cuanto a la violencia y el acoso escolar experimentado por niños y jóvenes alrededor del mundo. En un estudio realizado en 2016 por dichos organismos, en 18 países, se encontró que, de 100 000 jóvenes encuestados, dos tercios señalaron que fueron víctimas de acoso escolar (UNICEF, 2016; SRSG-VAC, 2016).

En 2018, investigadores de UNICEF Innocenti Research Centre (IRC), preocupados por el alza de la violencia y el acoso escolar en el mundo, desarrollaron y analizaron un

---

indicador global de acoso escolar, para el cual utilizaron encuestas escolares de países de distintas regiones. En su informe *Bullying: A global challenge requires a global measure* reportaron los resultados de encuestas combinadas llevadas a cabo en 10 países. El resultado fue que la prevalencia del acoso escolar varía enormemente entre una nación y otra, desde 7.1% en Tayikistán, hasta 81% en Botswana. Dicho estudio es un referente para reconocer que la violencia y el acoso escolar son un fenómeno complejo que adopta múltiples formas y niveles, según el contexto y la región (UNICEF-IRC, 2018).

En México, la violencia y el acoso escolar son temas que preocupan a las distintas autoridades educativas del país. En 2009 la SEP y UNICEF México reportaron en su *Informe nacional sobre violencia de género en educación básica en México* que 43.2% de los docentes señala casos de acoso en su escuela. También afirmaron que aproximadamente la quinta parte de los docentes de primaria y secundaria señalan la existencia de grupos de alumnos que intimidan a otros compañeros en el salón de clases (SEP y UNICEF México, 2009).

En Baja California, investigadores de la Unidad de Evaluación Educativa (UEE) de la Universidad Autónoma de Baja California (UABC), con apoyo de las autoridades educativas de la Secretaría de Educación del Estado de Baja California (SEEBC) realizó la llamada *Estrategia evaluativa integral UEE 2016-2*, que se aplicó en los cinco municipios del Estado de Baja California (Rodríguez-Macías, Pérez-Morán y Sarabia-Ocampo, 2017). Para dicha estrategia se diseñó y aplicó una encuesta con la que se exploraron distintos indicadores de la violencia y el acoso escolar en educación primaria, secundaria y media superior. Entre los hallazgos más representativos se encontró que en educación primaria 38% de alumnos, al menos en alguna ocasión, observó violencia entre sus compañeros de clase; en educación secundaria 48%, y en educación media superior 40%.

---

Dado lo anterior, y con el interés de aportar a los avances del estudio y la medición de la violencia y el acoso escolar, en la presente tesis se muestran los resultados del desarrollo y la validación de un indicador de violencia y acoso escolar, con base en el estudio del acoso escolar como proceso grupal, propuesto por Salmivalli (1999, 2010). Para ello, se tomó como instrumento base la *Escala de roles del estudiante en el ciclo de la violencia escolar*, diseñada por investigadores de la UEE durante la estrategia evaluativa integral UEE 2016-2 (Rodríguez-Macías, Pérez-Morán y Sarabia-Ocampo, 2017).

## 1.1. Antecedentes

Los primeros estudios sistematizados respecto a la violencia escolar datan de los años 70 del siglo pasado, de autores como Heinemann (1972) y Olweus (1973a, 1973b) en Suecia. Dicha iniciativa se extendió a otros países principalmente escandinavos, pero sin la suficiente fuerza para tener un impacto real en la reducción de la tasa de la violencia escolar. Según Olweus, las autoridades educativas no tomaron en serio el problema de la violencia y el acoso en las escuelas hasta que ocurrió un evento trágico en Noruega en 1982. Un periódico noruego publicó que tres niños, de entre 10 y 14 años de edad, se suicidaron por el acoso que sufrían por sus compañeros de escuela. Dicho evento detonó una serie de reacciones sociales y políticas, que concluyeron en una campaña nacional coordinada por el ministerio de educación de dicho país, en contra de los problemas de violencia y acoso en las escuelas de primaria y secundaria (Olweus, 1993).

Para 1996, la 49ª Asamblea Mundial de la Salud adoptó la resolución WHA49.25, en la que se declara que la violencia es un problema de salud pública elemental y en crecimiento en todas las regiones del mundo. En 2003 la Organización Panamericana de la Salud (OPS),

---

la Oficina Sanitaria Panamericana (OSP) y la Oficina Regional de la Organización Mundial de la Salud (OMS) elaboraron el primer *Informe mundial sobre la violencia y la salud*, como parte importante de la respuesta a dicha resolución. En ese documento se resalta la necesidad de explorar los diferentes tipos de violencia, así como la interacción entre ellos y otras variables asociadas. Uno de los temas importantes en el informe es la propuesta de dividir la violencia en tres grandes categorías generales, a partir de quién comete el acto: *violencia autoinfligida*, *violencia interpersonal* y *violencia colectiva*. La primera categoría corresponde a las autolesiones y el comportamiento suicida; la segunda, a la violencia que ocurre en ambientes sociales (por ejemplo, la familia, la pareja o la comunidad), y la tercera categoría se refiere a situaciones de violencia de corte político y económico (OPS, OSP y OMS, 2003).

Otro de los temas importantes en el informe de la OPS, OSP y OMS fue el hallazgo respecto a la participación de los niños en edad escolar en actos violentos. En 27 países de distintas regiones del mundo, entre 30% y 40% de los niños en edad escolar reportan haber participado en riñas o haber intimidado a otros, al menos por algún tiempo. Con base en los resultados, la OMS recomendó para futuros estudios hacer hincapié en la necesidad de medir la violencia y sus efectos. En especial recomendó recopilar información acerca de la magnitud y las repercusiones de la violencia; de los factores que aumentan el riesgo de ser víctima y de que se cometan actos violentos; así como de la eficacia de los programas de prevención. Aunado a ello, dichos organismos sugirieron utilizar como fuentes de información el testimonio de las personas; los registros de organismos gubernamentales, instituciones asistenciales y de programas locales; los registros comunitarios y gubernamentales; las encuestas poblacionales y estudios especiales.



---

Entre los años 2005 y 2006 la OMS, en colaboración con investigadores en el campo de la salud infantil, se realizó el estudio *Health behavior in school-aged children* (HBSC, por sus siglas en inglés), en el que se analizaron los temas de acoso escolar e intimidación en una gran muestra multinacional de niños en edad escolar. En total, se aplicaron encuestas a más de 200,000 escolares de 40 países diferentes de Europa, América del Norte e Israel (Currie, Roberts, Morgan, Smith, Settertobulte, Samdal y Barnekow, 2008). Uno de los logros más importantes de dicho estudio fue la comparación de las estimaciones de la prevalencia del acoso y la victimización entre los adolescentes en todos los países participantes en un mismo momento, utilizando instrumentos estandarizados. Entre los hallazgos principales de dicho estudio mostró es que alrededor de 35% de los jóvenes participantes informaron que estuvieron involucrados en el acoso escolar de otras personas, en al menos una ocasión durante los dos últimos meses anteriores a la encuesta. Dicha estimación varía sustancialmente entre países y regiones: de 9% a 54% para niños de 11 años; de 17% a 71% para niños de 13 años, y de 19% a 73% para niños de 15 años. Los promedios porcentuales de la participación en el acoso escolar a otras personas para los tres grupos de edad son 30%, 38% y 36%, respectivamente. También se reporta que 35% de los estudiantes participantes señalaron no presentar un comportamiento agresivo, 10% fueron víctimas de acoso escolar, 14% participaron en peleas, 8% acosaron a otros compañeros, 9% participaron en peleas y acoso a otros compañeros, y 24% participaron en peleas o son acosadores y víctimas (Craig y Harel 2004).

En años recientes, algunos países de América Latina y el Caribe, entre los que destacan Brasil, México, Ecuador, Costa Rica y Chile, en el contexto de la *Década internacional para la cultura de paz y la no-violencia para los niños del mundo* (2001-2010),

---

manifestaron su interés en fortalecer una educación inclusiva de calidad, orientada a la promoción de la paz y la no violencia. Como resultado, la OREALC/UNESCO Santiago inició en 2007, con apoyo de la Agencia Española de Cooperación Internacional para el Desarrollo, el proyecto *Educación para la Convivencia y la Cultura de la Paz en América Latina y El Caribe* (UNESCO y Ministerio de Educación de España, 2012). En sus inicios se intentaba, con dicho proyecto, impulsar un programa iberoamericano de educación para la paz, con la finalidad de fortalecer las acciones gubernamentales a favor del diseño, adopción e implementación de políticas educativas. Las dos primeras etapas del proyecto avanzaron con el diseño de contenidos e implementación de acciones para ese propósito. A la par se buscó que los países participantes generaran y sistematizaran aprendizajes, conocimientos, información de estudios e investigaciones transferibles, para que todos los países de la región los pudieran utilizar. Sin embargo, a pesar de los avances presentados por los sistemas educativos, la violencia sigue creciendo en las escuelas, así como su impacto negativo en los aprendizajes de los niños.

Recientemente, UNICEF y el SRSG-VAC en su informe *Ending the torment: Tackling bullying from the schoolyard to cyberspace* reportaron que, en 18 países participantes en un estudio de opinión en 2016, de 100,000 jóvenes que contestaron las encuestas, dos tercios señalaron que fueron víctimas de acoso escolar (SRSG-VAC, 2016; UNESCO, 2017). Con este trabajo se pudieron analizar los costos sociales y económicos que tiene la violencia contra los niños y adolescentes, dado su impacto en la salud física, mental y emocional, y en los resultados del aprendizaje (UNICEF, 2016).

Por su parte, la UNESCO (2017) reporta que la violencia y el acoso escolar afecta a una proporción considerable de niños y adolescentes alrededor del mundo. De acuerdo con

---

este organismo, aproximadamente 246 millones de niños y adolescentes experimentan violencia y acoso escolar de alguna manera u otra cada año. Para el cálculo de dicha estimación, la UNESCO tomó como referencia una población de 1.23 billones de niños en educación primaria y secundaria en todo el mundo, documentada en el reporte *2011 Global education digest* (UNESCO Institute for Statistic, 2011), y lo contrastó con las estimaciones de los porcentajes de niños y jóvenes afectados reportadas en el *2006 UN study on violence against children*, que varían entre 10% y 65%, entre países y estudios analizados.

Debido a que el estudio y la prevención de la violencia y el acoso escolar son prioridades para países, sistemas educativos y organizaciones humanitarias internacionales, investigadores alrededor del mundo han dedicado esfuerzos para desarrollar y validar mediciones e indicadores. Entre los antecedentes más notables en la medición del acoso escolar se encuentran: 1) el *Olweus bully/victim questionnaire* (BVQ, por sus siglas en inglés, Olweus, 1996, 2002); 2) el *Participant-role questionnaire*, desarrollado por Salmivalli, Lagerspetz, Bjorkqvist, Osterman y Kaukialnen (1996), y publicado en el artículo “Bullying as a group process: Participant roles and their relations to social status within the group”; y (3) la versión reducida y adaptada de dicho instrumento, por Sutton y Smith (1999).

Hasta la actualidad los tres instrumentos mencionados se utilizan, dada su validez y consolidación tanto para diagnosticar el nivel de violencia y acoso escolar, como para comprobar la eficacia de programas de prevención y tratamiento. Posteriormente se han desarrollado una variedad de indicadores basados en el enfoque de estudio del acoso escolar como proceso grupal propuesto Olweus (1993) y extendido por Salmivalli (1999, 2010). Sin embargo, la mayoría de ellos no cumplen con los criterios de calidad técnica y validez establecidos por organismos reconocidos en el desarrollo de instrumentos de medición como

---

la American Educational Research Association (AERA), la American Psychological Association (APA), el National Council on Measurement in Education (NCME, 1999), y la International Test Comisión (ITC, 2018).

Actualmente, son pocos los ejemplos de instrumentos e indicadores que cumplen con las características de calidad y validez que soporten interpretaciones y análisis entre países, en temas tan importantes como la violencia y el acoso escolar. Uno de los pocos ejemplos dignos de reconocer a nivel internacional, es el realizado por UNICEF Innocenti Research Centre (IRC), que ha desarrollado un indicador global de acoso escolar entre los niños, utilizando encuestas escolares existentes en todo el mundo. De forma especial, el IRC enfatiza que la tasa de acoso escolar entre los niños es un indicador clave del bienestar, tanto de las víctimas como de los agresores, así como un marcador importante para comparar el desarrollo social global de los niños, tanto por su impacto negativo en diversas dimensiones de su desarrollo personal, social, educativo y de salud, como en su persistencia en la edad adulta. Hasta el momento, con el uso del IRC en estudios internacionales comparativos se ha logrado determinar la complejidad y adopción de las múltiples formas y niveles que presenta el acoso escolar en distintas regiones y contextos de todo el mundo (UNICEF-IRC, 2018).

En México, la violencia y el acoso escolar son temas que preocupan a las distintas autoridades educativas del país (SEP y UNICEF México, 2009; INEE, 2018). En 2009 la SEP y UNICEF México reportaron en su *Informe nacional sobre violencia de género en educación básica en México* que 43.2% de los docentes señala casos de acoso en su escuela. También afirmaron que aproximadamente la quinta parte de los docentes de primaria y secundaria señalan la existencia de grupos de alumnos que intimidan a otros compañeros en el salón de clases (SEP y UNICEF México, 2009).

---

El Instituto Nacional para la Evaluación de la Educación (INEE), en su último informe sobre el estado que guarda el Sistema Educativo Nacional (SEN), presentado al Congreso de la Unión en 2018, menciona una serie de estudios relacionados con los temas de la violencia y la convivencia escolar, realizados desde 2007. Dicha institución señala que las manifestaciones de violencia reportadas por los alumnos de educación básica en 2007 tenían poca ocurrencia (8.7 puntos en promedio para primaria, y 5.7 puntos para secundaria en una escala de 0 a 100). Sin embargo, para 2015, con base en el análisis de la información que proporcionaron los estudiantes de sexto de primaria y tercero de secundaria, en los cuestionarios que acompañan las pruebas del Plan Nacional para la Evaluación de los Aprendizajes (PLANEA), se encontró un aumento importante en actos violentos en los centros escolares de todo el país. Según el INEE, entre 15% y 47% de los alumnos encuestados reportó haber observado actos violentos asociados a: insultos entre compañeros, robo de objetos o dinero, peleas a golpes, y daño a las instalaciones escolares. En particular la acción violenta observada con mayor frecuencia por los alumnos de educación primaria y secundaria fue el insulto entre compañeros (38.5% y 46.5% respectivamente) (INEE, 2017h, 2018).

Por su parte, en educación secundaria, en el Estudio Internacional de Educación Cívica y Ciudadana (ICCS, por sus siglas en inglés) 2016, se encontró que el comportamiento violento que los alumnos mexicanos de segundo grado reportaron con mayor recurrencia fue el *insulto* (64%), seguido por *haber sido objeto de burlas y recibir apodosos ofensivos por parte de compañeros* (63%); el *daño a objetos personales* (28%); *haber sido atacado físicamente por algún compañero* (20%); *recibir amenazas de un compañero* (19%), y *publicar en internet fotografías personales o textos ofensivos* (11%) (INEE, 2017c, 2017g).

---

En Baja California, investigadores de la Unidad de Evaluación Educativa (UEE) de la Universidad Autónoma de Baja California (UABC), con apoyo de las autoridades educativas de la Secretaría de Educación del Estado de Baja California (SEEBC) realizó la llamada *Estrategia evaluativa integral UEE 2016-2*, que se aplicó en los cinco municipios del Estado de Baja California (Rodríguez-Macías, Pérez-Morán y Sarabia-Ocampo, 2017). Para dicha estrategia se diseñó y aplicó una encuesta con la que se exploraron distintos indicadores de la violencia y el acoso escolar en educación primaria, secundaria y media superior. Entre los hallazgos más representativos se encontró que en educación primaria 38% de alumnos, al menos en alguna ocasión, observó violencia entre sus compañeros de clase; en educación secundaria 48%, y en educación media superior 40%.

Por su parte, 20% de alumnos de educación primaria y secundaria, y 14% de educación media superior reportan que al menos en alguna ocasión han sufrido violencia por parte de sus compañeros de clase. Asimismo, 5% de estudiantes de educación primaria, 7% de educación secundaria y 5% de educación media superior señalan que al menos en alguna ocasión sufrieron violencia por parte de alguno de los docentes de su escuela. Otro dato importante es el relacionado con el rol de defensor por parte de los alumnos. En educación primaria 79% de los alumnos señalan que al menos en alguna ocasión ayudaron a compañeros que se encontraban en problemas, a mejorar su situación; 78% en educación secundaria y 82% en educación media superior.

## **1.2. Planteamiento del problema**

Con base en la revisión de los antecedentes del estudio y la medición de la violencia y el acoso escolar, se puede decir que contar con instrumentos e indicadores con las

---

características de calidad técnica y validez establecidos por organismos como la AERA, APA y el NCME (1999), y la ITC (2018), es importante para fundamentar las interpretaciones que ayuden a las autoridades de los sistemas educativos del mundo a diagnosticar el nivel de violencia y acoso en las escuelas, así como a comprobar la eficacia de programas de prevención y tratamiento en los que se incursione. Para ello, es necesario contar con instrumentos de medición que permitan una aplicación rápida y sencilla, sin demeritar su calidad técnica y validez. Al mismo tiempo, que no representen un alto costo económico para los aplicadores ni un alto costo en tiempo para los participantes.

Sin embargo, los instrumentos que actualmente se utilizan en México para el estudio y la medición de la violencia y el acoso escolar no siempre cumplen con los estándares y criterios necesarios para asegurar su calidad técnica (SEP y UNICEF México, 2009; INEE, 2018). De hecho, gran parte de los principales informes diagnósticos publicados por organismos públicos en México a nivel nacional, utilizan cuestionarios extensos y muy completos, pero no informan acerca de sus propiedades psicométricas asociadas a un modelo teórico fuerte o consolidado.

En Baja California, el trabajo realizado por investigadores de la UEE es destacado en el campo de la medición de la violencia y el acoso escolar, pero se requiere continuar afinando los instrumentos diseñados, con base en los estándares más rigurosos de organismos como la AERA, APA y el NCME (1999), y la ITC (2018). Es por ello que, resulta significativo emprender estudios encaminados al análisis de las propiedades psicométricas y de las evidencias de la validez de constructo de la Escala del Rol del Estudiante en el Ciclo de la Violencia Escolar (ERECVE), utilizada para su aplicación en la Estrategia Evaluativa Integral UEE 2016-2. La principal inquietud de los investigadores de la UEE en lo referente al diseño

---

de dicha escala no es explicar algún evento aislado, comportamiento o respuesta individual, sino dar cuenta de la consistencia y validez de los comportamientos o las respuestas de los ítems, con base en factores sustantivos, determinantes y distinguibles (Messick, 1987). A partir del contexto descrito surgió la pregunta que motivó el presente trabajo:

¿Cuáles son las evidencias de validez de constructo en el aspecto de la estructura interna y de la relación con otras variables, de la Escala del Rol del Estudiante en el Ciclo de la Violencia Escolar (ERECVE), para educación secundaria?

### **1.3. Objetivos de investigación**

Para dar respuesta a la pregunta planteada se propuso como objetivo general del presente estudio analizar las propiedades psicométricas relacionadas con las evidencias de validez de constructo en el aspecto de la estructura interna y de la relación con otras variables de la Escala del Rol del Estudiante en el Ciclo de la Violencia Escolar (ERECVE), para educación secundaria, diseñada por Juan Carlos Pérez Morán y Juan Carlos Rodríguez Macías (Rodríguez-Macías, Pérez-Morán y Sarabia-Ocampo, 2017). Para el logro de dicho propósito se establecieron siete objetivos específicos:

- 1) Revisar los antecedentes y fundamentos teóricos de la medición de la violencia escolar, vista como un proceso grupal; así como los métodos y técnicas estadísticas para la obtención de evidencias de validez de constructo, en el aspecto de estructura interna y de relación con otras variables.
- 2) Analizar los supuestos requeridos para aplicar el Análisis factorial exploratorio y el Análisis factorial confirmatorio.
- 3) Caracterizar las puntuaciones por ítem y subescala de la ERECVE; así como verificar la normalidad, linealidad y colinealidad de las puntuaciones.



- 4) Analizar la discriminación y confiabilidad de la ERECVE con base en los supuestos de la Teoría clásica de los test (TCT).
- 5) Analizar la estructura interna del modelo de cinco factores de la ERECVE, mediante la aplicación del Análisis factorial exploratorio (AFE).
- 6) Analizar la estructura interna del modelo de cinco factores de la ERECVE, mediante la aplicación del Análisis factorial confirmatorio (AFC).
- 7) Analizar las correlaciones concurrentes y discriminantes entre las subescalas de la ERECVE y con otras variables.

#### **1.4. Justificación**

La violencia y el acoso escolar son un problema al alza, que preocupa a las autoridades de los distintos niveles del sistema educativo nacional (SEN) (SEP y UNICEF México, 2009; Rodríguez-Macías, Pérez-Morán y Sarabia-Ocampo, 2017; INEE, 2018). Ante la crisis de violencia que vive el país, diagnosticar el nivel de violencia y acoso escolar, y comprobar la eficacia de programas de prevención y tratamiento son una tarea que no puede esperar. Es por ello que diseñar, desarrollar e implementar estudios y mediciones de la violencia y el acoso escolar son una prioridad para autoridades educativas, docentes e investigadores en el campo de la educación en México y el mundo.

La escuela se ha convertido en el sustituto de la familia. Es un espacio en el que la sociedad forma a los futuros ciudadanos para dar vida a la democracia. Es el organismo donde las personas transitan de la infancia a la vida adulta. Es la institución que, con sus actividades, contribuye a edificar la cultura nacional (INEE, 2015). Por ello, es relevante conjuntar esfuerzos que ayuden a las escuelas a ser un espacio seguro, para que los estudiantes

---

puedan desarrollarse en una *cultura de paz y de no-violencia* (UNESCO y Ministerio de Educación de España, 2012).

El principal beneficio que aporta el análisis de las propiedades psicométricas relacionadas con las evidencias de validez de constructo, en el aspecto de la estructura interna y de la relación con otras variables de la ERECVE para educación secundaria, es que los investigadores y profesionales en el campo del estudio y la medición de la violencia y el acoso escolar contarán con información especializada acerca de la calidad técnica de un instrumento de bajo costo de aplicación y con un referente teórico fuerte, que cumple con estándares y criterios básicos para su aplicación en estudios de gran escala.

Aunado a lo anterior, presentar un estudio organizado a partir de un modelo metodológico fundamentado en las recomendaciones, criterios y estándares de la AERA, APA y el NCME (1999), así como del ITC (2018), aporta a estudiantes, investigadores y profesionales en el campo de la medición educativa un ejemplo o guía clara para el desarrollo y la validación de instrumentos similares. Asimismo, revisar los antecedentes y los fundamentos teóricos de la medición de la violencia escolar, vista como un proceso grupal, así como los métodos y técnicas estadísticas para la obtención de evidencias de validez de constructo, en el aspecto de la estructura interna y de la relación con otras variables, presenta a investigadores y profesionales del campo educativo una sistematización que facilita el acceso a información valiosa para fundamentar, proponer y efectuar futuras investigaciones.

## **1.5. Estructura de la tesis**

La tesis consta de cinco capítulos y dos secciones al final (referencias y apéndices). En el primer capítulo (Introducción) se describen los antecedentes del estudio, el planteamiento del

problema, los objetivos de investigación y la justificación del estudio. En el segundo capítulo se presenta el marco teórico que enmarca la presente investigación, el cual consiste en el análisis factorial (exploratorio y confirmatorio) como técnica psicométrica que permite obtener evidenciad de validez de constructo. En el tercer capítulo, se presenta la metodología que siguió el curso de esta investigación. En el cuarto capítulo se exponen los resultados y, por último, en el quinto capítulo, se presentan la discusión y las conclusiones del presente estudio.

---

## II. MARCO TEÓRICO

Este capítulo se organiza en tres subapartados en los que se describen los fundamentos teóricos de la presente tesis. En el primer apartado se abordan los fundamentos teóricos de la medición de los roles en el ciclo de la violencia escolar, vista como un proceso grupal. En el segundo apartado se mencionan los antecedentes conceptuales y enfoques de la Teoría de la validez de constructo; así como los estándares relacionados con las evidencias de validez en el aspecto de la estructura interna y de la relación con otras variables de pruebas psicológicas y educativas. En el tercer subapartado se presentan los métodos y técnicas estadísticas, así como los criterios e índices de ajuste de los modelos para el Análisis factorial exploratorio y confirmatorio, utilizados en el estudio.

### 2.1. Teoría y medición de los roles en el ciclo de la violencia escolar

La violencia y el acoso escolar presentes en múltiples contextos se han convertido en un problema frecuente en las escuelas alrededor del mundo, independientemente de sus características: su calidad, su financiamiento público o privado, su perfil laico o religioso, por mencionar algunas (Olweus, 1983, 1991, 2001; Rigby y Slee, 1991; Boulton y Underwood, 1992; Smith y Sharp, 1994; Craig, Peters y Konarski, 1998; Craig y Harel, 2004; UNISEF, 2005, 2009, 2011; Pepler, Jiang, Craig y Connolly, 2008; D'Angelo y Fernández, 2010; UNESCO, 2017). Además de la prevalencia alta de la violencia y el acoso reportado en las escuelas, investigaciones recientes ponen de manifiesto los efectos nocivos de ambas problemáticas que produce en la salud mental, en el desarrollo infantil, así como en la calidad de vida de los estudiantes en general. Del mismo modo, dichos estudios destacan el impacto

de la violencia en el clima escolar de las instituciones educativas en las que se origina (Springer, Cuevas, Ortiz, Keis y Wilkinson, 2015).

Para entender el fenómeno de la violencia escolar, investigadores en el campo han examinado sus características y dinámicas, con el fin de delinear las múltiples formas que adopta: física, verbal, social o relacional, o a través de las tecnologías de la información. Asimismo, para tratar de identificar a los actores principales que participan en ella y los roles diversos que adoptan (de agresor, víctima y observador), con dinámicas particulares que reproducen estos esquemas relacionales. Hasta el momento, una cantidad grande de las investigaciones se han concentrado en la prevalencia y en la caracterización de los agresores y de las víctimas, lo que coloca en un plano secundario el rol de los observadores de la dinámica de la violencia. Entre las investigaciones más importantes de la dinámica de la violencia escolar se pueden mencionar los aportes de Craig y Pepler (2007); Ortega, Sánchez y Menesini (2002); Salmivalli (2010); Salmivalli, Lagerspetz, Bjorkqvist, Osterman y Kaukialnen (1996); Thornberg y Jungert (2014), y Twemlow, Fonagy y Sacco (2004). Todos ellos resaltan la importancia de los observadores.

Trautmann (2008) sostiene que los problemas del clima y la disciplina escolar no solo involucran a la víctima y al agresor, como antes se pensaba, sino también participan los testigos directos (los que se encuentran durante el acto) y los indirectos (como el personal del centro educativo y la familia). En este sentido, la escuela –como institución social– es la encargada de promover prácticas orientadas a erradicar expresiones como la exclusión, la intolerancia y la discriminación (Fierro, 2013). Por ello, es necesario dotar a autoridades y actores educativos en la comprensión del fenómeno de la violencia y el acoso escolar, al igual que en técnicas de prevención y de su diagnóstico, para saber cómo gestionar las diversas

situaciones relacionadas a esta problemática. Como señala Ortega (2010, p.18), “no debemos intentar evitar los conflictos en la escuela, sino aprender de ellos, aprender a resolverlos de forma asertiva, porque a la escuela se va a aprender, a ser una persona equilibrada, sensata, solidaria y segura”.

En relación con los protagonistas de los roles de la violencia escolar, destacan tres principales: el agresor o acosador, la víctima y los espectadores o testigos. Gran parte de la investigación en el tema de la violencia escolar se ha enfocado en conocer tanto las características de las víctimas y de sus agresores (Nansel et al., 2004; Olweus, 1993; Schwartz, 2000; Veenstra et al., 2005), como de los efectos negativos de este tipo de situaciones en los niños involucrados (Boivin, Hymel y Hodges, 2001; Card y Little, 2006; Hawker y Boulton, 2000). Sin embargo, autores como Salmivalli et al. (1996) han dirigido su atención en el estudio de otros roles: ayudante del agresor, reforzador del agresor y defensor de la víctima.

### **2.1.1. Teoría de los roles en el ciclo de la violencia escolar**

La violencia en el contexto juvenil no es un tema nuevo en la sociedad. Las áreas donde los jóvenes experimentan violencia están directamente relacionadas con el entorno donde desarrollan sus actividades cotidianas, como su hogar, grupo de amigos y escuela, por mencionar algunos (Pulido, 2006). Sin embargo, la violencia en las escuelas es un problema grave en todo el mundo que se ha investigado ampliamente a través de los años (Due, Holstein, Lynch, Diderichsen, Nic-Gabhein, Scheidt y The Health Behaviour in School-Aged Children Bullying Working Group, 2005; Lagerspetz, Björkqvist, Berts y King, 1982).

Diversos estudios han clasificado a los alumnos de los centros escolares en tres roles: agresores o víctimas, ambos o ninguno (Sierra, 2010; Oliva y Calderón, 2014; Valdés, Yañez y Carlos, 2013). La agresión y la victimización representan riesgos significativos para el desarrollo de los jóvenes estudiantes, debido a los vínculos claros entre la agresión y la victimización. Los esfuerzos en investigación e intervención se han enfocado más en cómo estos comportamientos se desarrollan de forma independiente y conjunta. Se sabe que una minoría considerable de estudiantes principalmente en escuelas secundarias está involucrada en la intimidación de igual a igual, ya sea como agresores o víctimas, o como ambos, siendo acosados y acosando a otros. Sin embargo, dado que el acoso escolar recientemente se considera un proceso grupal (Salmivalli, 2010), es importante contemplar más que solo los roles de acosador y víctima para las investigaciones en esta área.

El término *bullying* hace referencia al fenómeno del maltrato, acoso o abuso entre iguales en el contexto escolar, así como la realización repetida de acciones negativas por parte de una o más personas contra un individuo o individuos en una situación en la que existe un desequilibrio de poder físico y/o psicológico (Olweus, 1993). En la década de los setenta Olweus (1978) reveló que la prevalencia de la violencia entre iguales en los centros escolares es comúnmente subestimada, y que la presencia de diferentes factores asociados a este fenómeno como el sexo, la edad, la escuela y el país influyen en su comportamiento (Due et al., 2005; Nansel, Craig, Overpeck, Saluja y Ruan, 2004; UNICEF, 2012).

En los últimos años se han realizado distintas investigaciones relacionados con el estudio de las dinámicas y ciclos de la violencia escolar, principalmente en el ámbito internacional (Olweus, 1978; 1993; Ortega, 2000; Díaz-Aguado, Martínez Arias y Martín Seoane, 2004; Jimerson y Furlong, 2006). Estas investigaciones han aportado información

valiosa de los distintos roles que los estudiantes adoptan cuando hay presencia de violencia en los centros escolares.

El papel del centro escolar en el tema de la violencia se entiende de manera diferente a la del sujeto o grupo de individuos en otros contextos que agreden activamente a una persona. Se considera que los miembros de un grupo tienen roles particulares, impulsados por diversas emociones, actitudes y motivaciones propias del contexto escolar. Sus características individuales interactúan con factores ambientales, como las reglas del aula, lo que contribuye a un proceso que puede tener resultados perjudiciales para la persona o grupo de personas involucradas. Poner la violencia en el contexto escolar ayuda a comprender mejor la motivación de los individuos para agredir, así como a ver la falta de apoyo brindado a las víctimas, la persistencia del acoso y el ajuste de las víctimas en diversos contextos (Salmivalli, 2010). Por ejemplo, según Atlas y Pepler, 1998; Hawkins, Pepler y Craig (2001), los compañeros del aula están presentes en 85% a 88% de los episodios del ciclo de la violencia escolar y parecen desempeñar roles específicos, a los que se denomina *roles de participantes* (Goossens, Olthof y Dekker, 2006; Salmivalli, Lagerspetz, Björkqvist, Österman, y Kaukiainen, 1996; Sutton y Smith, 1999).

Además del agresor y la víctima se han identificado otros cuatro roles en distintos contextos (Salmivalli et al., 1996; Salmivalli, Lappalainen, y Lagerspetz, 1998). Salmivalli et al. utilizaron un procedimiento de nominación de pares para identificar cuatro roles de participantes que los alumnos pueden adoptar en la dinámica del acoso escolar, además de ser agresores o víctimas: asistente de agresor, refuerzo de agresor, observador y defensores de la víctima. Los asistentes son jóvenes que se unen con entusiasmo a la intimidación y apoyan activamente al agresor, para que ataque a la víctima. Por su parte los reforzadores



proporcionan retroalimentación positiva a los agresores al reír, animar, incitar, aclamar y/o proporcionar una audiencia. Los observadores se mantienen alejados de la situación de acoso y no toman partido, aunque no actúan, pueden alentar la violencia escolar enviando un mensaje silencioso de que el comportamiento del agresor es aceptable. Por último, los defensores toman partido por las víctimas, los consuelan, apoyan y tratan activamente de detener la intimidación. Los grupos de pares son contextos centrales para que suceda la agresión y la victimización, los pares están presentes en la mayoría de los encuentros de victimización y desempeñan un papel crucial para reforzar o inhibir un ataque (Salmivalli, 2001).

La principal asociación que se forma al referirse a violencia es el proceder agresivamente de un sujeto hacia otro, sin embargo, la intimidación es un subtipo de comportamiento agresivo, en el que un individuo o un grupo de individuos atacan, humillan y/o excluyen repetidamente a una persona relativamente indefensa (Salmivalli, 2010). Los agresores tienen razones para creer que su comportamiento ayuda a mejorar su estatus de pares, pueden ser percibidos como geniales, poderosos y populares, de tal modo que la intimidación puede ser útil para ganar prestigio (Caravita et al., 2009; Juvonen, Graham y Schuster, 2003; Rodkin, Farmer, Pearl y Van Acker, 2006; Vaillancourt, Hymel y McDougall, 2003).

### **2.1.2. Instrumentos de medición de los roles en el ciclo de la violencia escolar**

Entre los instrumentos más notables en la medición de los roles en el ciclo de la violencia escolar, se pueden mencionar el de *Participant-Role Questionnaire* de Salmivalli, et al. (1996) publicado en el artículo *Bullying as a group process: Participant Roles and Their Relations to*

---

Social Status Within the Group; y el de Sutton y Smith (1999) quienes realizaron una adaptación reducida del enfoque de la *Participant-Role Questionnaire* de Salmivalli, et al. (1996).

El *Participant-Role Questionnaire* (Salmivalli, et al., 1996) es un instrumento de 49 ítems que se aplicó a 573 niños finlandeses de sexto grado, entre 12 y 13 años de edad para investigar los autoreportes de los estudiantes y de sus compañeros relacionados a situaciones en donde se presenta comportamiento violento (o roles de la violencia escolar de los participantes). Con dicho instrumento se examina la relación entre los roles de los participantes y el estado sociométrico en la clase, además de tomar en consideración los aspectos relacionados con el ambiente escolar, la aceptación, el rechazo social, y la pertenencia a uno de los cinco grupos de estatus sociométrico (popular, rechazado, descuidado, conflictivo y promedio). Salmivalli et al. (1996), reportan una alta confiabilidad para estas subescalas con un alfa de Cronbach de .81 a .93, lo que respalda la existencia factores que representan a los distintos roles dentro del proceso de la violencia escolar.

Además de los estatus sociométricos, con el *Participant-Role Questionnaire* se puede detectar la existencia de seis roles: el agresor, la víctima, el reforzador del agresor (participa en las agresiones), el asistente del agresor (no ejerce la agresión directa hacia la víctima), el defensor de la víctima y el observador (se mantiene alejado de las peleas y mostraría una actitud pasiva). Para ello los alumnos evalúan qué tan bien cada compañero en su clase, incluyéndose, se adaptan a las descripciones de comportamientos de situaciones de violencia escolar. En la escala de agresor, se describe el comportamiento activo, de toma de iniciativa y de violencia. En la escala de asistente, también presenta el comportamiento de agresión activa pero más de acompañante que de líder. En la escala del reforzador se incluyen

elementos que reflejan tendencias a actuar de maneras que se fortalezca el comportamiento de violencia, como reírse, proporcionar audiencia e incitar al acosador. En la escala del defensor, figuran muestras de apoyo, comportamientos de consuelo a la víctima y esfuerzos activos para hacer que otros detengan la intimidación. En la escala del observador se presenta el no hacer nada y el mantenerse fuera de las situaciones de violencia. Finalmente, en la escala de la víctima se mantiene un elemento que en el cuestionario se describe como “es violentada”.

Después de las respuestas de los estudiantes los puntajes estimados por pares obtenidos se estandarizan por clase, de ese modo cada alumno identifica un rol participante en particular únicamente si su puntaje en ese rol fue el más alto de la clase o el más alto en esa escala que en cualquier otra. Un alumno fue víctima si el 30% o más de sus compañeros de clase los nominaron para el ítem *recibe agresión*. Salmivalli et al. (1996), hallaron que el 87% de los alumnos podrían clasificarse de esta manera, con 8.2% de agresores, 19.5% de reforzadores, 6.8% de ayudantes, 17.3% de defensores, 23.7% de observadores, 11.7% de víctimas y 12.7% sin rol. Hubo diferencias significativas en la variable sexo en la distribución de los roles de los participantes.

Los hombres estaban mayormente posicionados en los papeles del agresor, reforzador y ayudante, mientras que los papeles de las mujeres se situaban en los roles de defensor y observador. Al comparar los informes propios y de los compañeros, los alumnos eran moderadamente conscientes de su rol de participante, aunque subestimaban su participación en el comportamiento activo del acoso y enfatizaron que actuaban como defensores y observadores.

La investigación de (Sutton y Smith, 1999) tuvo dos objetivos principales. El primero fue adaptar la Escala de roles de los participantes para usarla con una muestra diferente, es decir, alumnos más pequeños en escuelas británicas. Se pensó que este era un ejercicio valioso, ya que la estabilidad de la conducta de acoso a lo largo de la vida exige comprensión e intervención a una edad temprana (Farrington, 1992).

Por su parte el segundo objetivo fue discutir nuevas formas de usar la Escala de roles del participante para distinguir los grupos dentro del proceso de acoso escolar, como la estandarización de los puntajes por clase de la escuela al igual que en el estudio anterior de Salmivalli et al. (1996) de tal forma que podría resultar en que los niños sean categorizados como no tener ningún rol debido a un puntaje extremo de un compañero de clase.

El estudio utilizó ocho grupos de cuatro escuelas del sudeste de Londres, lo que dio un total de 206. El número de alumnos que se negaron a dar su consentimiento o que el padre rechazó el consentimiento al inicio del estudio no se incluyó fueron 13. La muestra final fue de 193 niños de los cuales 102 fueron mujeres y 91 hombres con edades comprendidas entre siete años siete meses y 10 años con ocho meses donde la media fue de nueve años y cero meses.

En la investigación no se tomaron medidas socioeconómicas, sin embargo, las áreas de captación de las escuelas fueron más bajas que la clase media. La mezcla étnica fue 38.3% blanca, 45.4% afro-caribeña, 10.2% asiática, 1.5% mediterránea y 4.6% de origen étnico mixto. El tamaño de los grupos varió de 20 a 29 alumnos, con un medio de clase de 24.9 (Sutton y Smith, 1999).

Un estudio piloto indicó que muchos niños dentro del rango de 7 a 10 años de edad tenían varios problemas con el cuestionario presentado por Salmivalli et al. (1996). Los

maestros informaron un alto grado de variabilidad en la edad de lectura y las puntuaciones en la Escala de vocabulario de imágenes británica (Dunn, Dunn, Whetton, y Pintilie, 1982), que tiene una validez concurrente que osciló entre 40 y 147, con una media de 94 para la muestra final. En consecuencia, muchos niños tuvieron problemas para leer y comprender las descripciones de comportamiento, por ejemplo, hace comentarios irónicos sobre la víctima, y en combinación con el número total de preguntas que eran 49 fue una evaluación complicada.

Por lo tanto, se hicieron varios cambios al procedimiento utilizado por Salmivalli et al. (1996) que reflejaba estas preocupaciones. Primero, el procedimiento se llevó a cabo como una entrevista en lugar de un cuestionario. Esto permitió al entrevistador aclarar los puntos de los que el niño no estaba seguro y tener más confianza en la precisión de las respuestas. En segundo lugar, el número de descripciones de comportamiento se redujo de 49 a 21, reteniendo el elemento de víctima y posteriormente se les pidió a cinco evaluadores que seleccionaran 20 elementos más de la lista original. Los artículos se guardaron si al menos cuatro de los cinco evaluadores los seleccionaron.

Algunos elementos revisados incorporaron más de un elemento de la lista original, por ejemplo, tres ítems separados en la escala original del asistente: *ayudan al agresor*, *atrapa a la víctima* y *retiene a la víctima cuando es acosado*, se convirtieron en *ayuda al agresor*, *tal vez atrapando o sujetando a la víctima*. El resultado fue la reducción de 10 ítems en la escala de agresor a cuatro, siete ítems de reforzador a cinco, cuatro ítems de asistente a dos, 20 ítems de defensor a cinco y siete ítems externos a cuatro.

Por último, los alumnos solo debían nominar a otros de sus compañeros del mismo sexo. Se pensó que esto reduciría el esfuerzo mental y el tiempo necesario y se justificó por el hecho de que entre las edades de 7 y 10 años los grupos de pares son predominantemente

del mismo sexo (Hartup, 1996). Los participantes fueron vistos en una sala tranquila y privada lejos de la clase. Se sentaron a la mesa con el experimentador y se les mostró una serie de los nombres de sus compañeros del mismo sexo impresos en una tarjeta, roja o blanca. Se les pidió que indicaran cada nombre y los leyeran en voz alta, y se corrigió cualquier error u omisión.

Las 21 descripciones de comportamiento que conforman los seis roles de agresor, reforzador, asistente, defensor, observador y víctima se leyeron en un orden pseudoaleatorio y las respuestas de los participantes se anotaron en una hoja de puntuación. Se les aseguró a los alumnos que nadie más sabría lo que decían para que pudieran ser completamente honestos y nadie se metería en ningún problema.

Las investigaciones de Salmivalli et al. (1996) y Sutton et al. (1999), aportan aspectos significativos de conocimiento en el tema de los roles en el ciclo de la violencia escolar. La relevancia de estas dos investigaciones rebasa las dos décadas que tienen de haber sido publicadas, por lo cual debemos reconocer estos estudios como bases teóricas sólidas para futuras exploraciones en el tema de los roles del acoso en el contexto educativo.

## **2.2. Evidencias de validez de constructo**

Cronbach y Meehl (1955) establecen que la validez de constructo de los instrumentos consiste en un análisis de la significación de las puntuaciones de una medida expresada en términos de los conceptos psicológicos asumidos en su medición. Lo anterior involucra siempre que una prueba se interpreta como una medida de algún atributo o rasgo que no está operacionalmente definido. En el modelo clásico de validez propuesto por Cronbach y Meehl, la validez de constructo es uno de los tres tipos de evidencias de validez, junto a las evidencias

de validez de contenido y validez de criterio (Guion, 1980; Brown, 1996). Sin embargo, la validez de constructo es una de las más importantes de los distintos tipos de validez, ya que es un concepto unificador que integra las evidencias de criterio y contenido en un marco común para evaluar hipótesis racionales teóricamente relevantes (Messick, 1980). Con ello, lo sustantivo para la validez de constructo, desde el enfoque clásico, son las ideas teóricas detrás del rasgo en cuestión, es decir, los conceptos que definen cómo son vistos los aspectos de interés a medir (Pennington, 2003).

Asimismo, la validez de constructo se concibe como la existencia de evidencias relacionadas a la congruencia entre el referente y la prueba; el énfasis está dado en sustentar el grado en que los puntajes de la prueba representan la medida del atributo psicológico que se presume medir (Leyva, 2011). Para ello, es necesario realizar el análisis de estructura interna y del análisis de la relación de la prueba con otras variables. Cronbach (1984) señala que la meta final de la validación es la explicación y comprensión de un fenómeno, por lo tanto, esto lleva a conceptualizar que toda validación es de constructo.

Por su parte, Messick (1989b) define la validez de constructo desde un enfoque unitario como un juicio evaluativo integral del grado en que la evidencia empírica y los fundamentos teóricos apoyan la capacidad, la adecuación de las inferencias y las consecuencias basadas en los resultados de las pruebas. La validez de constructo es la adecuación de las inferencias basadas en observaciones o mediciones, principalmente de los resultados de las pruebas. De tal forma que los constructos subyacentes a las pruebas son abstracciones creadas deliberadamente por los investigadores con el fin de representar la variable latente, que es la causa de las puntuaciones en una determinada medida que no es directamente observable.

Las evidencias de validez de constructo de una interpretación dada de los puntajes de las pruebas para un uso específico, es una condición necesaria para el uso justificable, ya que la medición de un constructo teórico más general implica una red teórica más amplia y comprensiva (Leyva, 2011). La validez de constructo de una prueba debe ser demostrada por una acumulación de evidencias. Por ejemplo, tomando la definición unificada de validez de constructo propuesta por Messick, podríamos aportar evidencias utilizando análisis de contenido, coeficientes de correlación, análisis factorial, estudios ANOVA que demuestran divergencias entre grupos diferenciales o estudios de intervención pretest-postest, estudios multirrasgo-multimétodo, entre otros.

Con todo ello, evaluar la validez de constructo es equivalente a construir una teoría (Cronbach y Meehl, 1955; Messick, 1989b). Por lo tanto, debe haber un método claro para comparar teorías alternativas de los atributos subyacentes a las pruebas en la fase de representación y la validación. Un método efectivo para comparar teorías es operacionalizar las teorías que compiten en modelos cuantitativos y luego comparar sus ajustes (Embretson, 1983).

### **2.2.1. Conceptos y enfoques de la validez de constructo**

El término validez surgió a finales del siglo XIX con la aparición en el campo de las matemáticas aplicadas del coeficiente de correlación de Pearson (Alvarado y Santisteban, 2006). Posteriormente, la introducción de la estadística a las puntuaciones de una prueba, convierten la validez en un coeficiente de correlación nombrado por Clark Hull en 1928 como coeficiente de validez (Geisinger, 1992). Garrett (1937) mencionó en el libro de *Estadística en Psicología y Educación* que la validez depende de la fidelidad con que una prueba mide lo



que pretende medir. Cabe mencionar que la anterior definición se ha utilizado a través de los años erróneamente por estudiantes e inexpertos en el área de la psicometría. Por su parte, Guilford (1946) señaló que una prueba es válida para algo con lo que se intercorrelaciona, y Cureton (1950) indicó que la puntuación observada de una prueba con un criterio, sólo representa su carácter predictivo.

Para los años cincuenta inicia la discusión más acalorada al rededor del concepto de validez, siendo un comité técnico a cargo de la American Psychological Association (APA), quien publicara el artículo de *Technical recommendations for psychological tests and diagnostic techniques: A preliminary proposal* (APA, 1952), y fue entonces el surgimiento de la primera clasificación de cuatro categorías de validez: predictiva, criterio, contenido y concurrente. Dos años después, hubo una nueva versión donde se clasificaron los tipos de validez y se definieron como el grado en que una prueba es capaz de lograr ciertos objetivos (APA, 1954) y, en el siguiente año, la misma APA publicó un artículo donde destaca la validez de constructo relacionando este tipo de validez con la técnica estadística de análisis factorial (Cronbach y Meehl, 1955).

Para 1966, la APA publicó otro artículo titulado *Technical Recommendations for Achievement Tests* (APA, 1966) en colaboración con la American Educational Research Association (AERA) y la National Council on Measurements Used in Education (ahora llamada National Council on Measurement in Education [NCME]), en el cual se menciona un modelo tripartito de la validez: criterio, contenido y constructo. Para la versión de 1974, se definió la validez como la propiedad de las interpretaciones de puntuaciones y no del test; de esta manera, se relaciona con el atributo (constructo) y con el uso del test (predictiva) (APA, AERA, y NCME, 1974).

En los ochentas, la validez se definió como la exactitud con que pueden hacerse medidas significativas y adecuadas, es decir, que midan realmente los rasgos que pretenden (Magnusson, 1982). Messick (1993) informó que su postura sobre la validez es de un juicio evaluativo integral del grado en el que la evidencia empírica y la racionalidad teórica dan apoyo a qué tan apropiadas y adecuadas son las inferencias y acciones basadas en los puntajes de una prueba o cualquier fue compuesto bajo un propósito (APA, AERA, y NCME, 1999), también señaló tiene cinco fuentes de evidencia: basada en el contenido, procesos de respuesta, estructura interna, relación con otras variables y consecuencias del test.

Entrados al siglo XXI, Borsboom y Mellenbergh (2004) afirmaron que la validez no es compleja, cuenta con facetas dependientes de redes nomológicas o consecuencias sociales del uso de las pruebas, además que presenta de un concepto elemental que se relaciona con el hecho de que una prueba es válida si mide lo que se supone debe medir. Moss, Girard y Haniford (2006), hicieron referencia a que no se habla de la validez de un instrumento sino de la validez de las puntuaciones, debido a que las propiedades de este dependen del contexto de evaluación y de la muestra. La última versión del Standards for Educational and Psychological Testing se publicó en el 2014 en el cual se plantea un modelo de medida triple donde se incluye: la validez, confiabilidad y equidad como una nueva propiedad psicométrica necesaria si se pretende contar con un instrumento de medición adecuado.

La importancia de la validez en el área de la medición de instrumentos educativos y psicológicos es fundamental debido a que depende de este concepto que el constructo a evaluar esté apegado a la realidad evaluada. Un test con propiedades psicométricas estables es aquel que ha sido examinado rigurosamente para asegurar la medición precisa y verdadera del constructo subyacente para cuya medición fue diseñado, y el cual produce resultados

---

consistentes en cada aplicación a distintas muestras de la misma población (APA, 2013). Los Estándares definen la validez como:

“El grado en que la evidencia y la teoría respaldan las interpretaciones de los puntajes de las pruebas para los usos propuestos de las pruebas. La validez es, por lo tanto, la consideración más fundamental en el desarrollo de pruebas y pruebas de evaluación. El proceso de validación implica la acumulación de evidencia relevante para proporcionar una base científica sólida para las interpretaciones de puntuación propuestas. Son las interpretaciones de los puntajes de los exámenes para los usos propuestos los que se evalúan, no el criterio de la prueba. Cuando los puntajes de los exámenes se interpretan de más de una manera, cada interpretación prevista debe ser validada” (AERA, APA y NCME, 2014, p. 11).

Los estándares hacen mención de diversos aspectos que se deben contemplar en la construcción, el desarrollo, la evaluación, administración y documentación de una prueba, deja en claro que la validez es la consideración más importante que se debe tomar en cuenta al desarrollar y evaluar test. La Comisión Internacional de los Test [ITC, por sus siglas en inglés] (2010), sostuvo que los test deben respaldarse por evidencias de confiabilidad y validez, las cuales permitan dar fundamento a las inferencias que se puedan realizar a partir de las puntuaciones obtenidas. Asimismo, si el instrumento es aplicado en una nueva población, se requiere demostrar que este continúa evaluando las mismas cualidades con el mismo grado de precisión y exactitud (Geisinger, 1994).

### **2.2.2. Evidencias de validez de constructo de la estructura interna y de la relación con otras variables**

Uno de los documentos de consulta más importantes para el desarrollo y la evaluación de instrumentos de medición psicológica y educativa es el que se publicó en 1985, 1999 y el más reciente en 2014, por un comité compuesto por la *American Educational Research Association*

(AERA), la *American Psychological Association* (APA) y el *National Council on Measurement in Education* (NCME). Se trata de los *Standards for Educational and Psychological Testing* (2014). En ese documento se establece un marco de referencia que contiene lineamientos que todo organismo dedicado a elaborar instrumentos de evaluación debe seguir para el desarrollo, uso y evaluación.

Sin embargo, en 1895 se nombró un comité de la APA para investigar la viabilidad de estandarizar la recopilación de mediciones mentales y físicas (Haney y Madaus, 1991). Las organizaciones profesionales en el área de la Educación y Psicología en los Estados Unidos, unieron esfuerzos para realizar el desarrollo de estándares de calidad para el desempeño profesional y los instrumentos de medición.

Fue en 1954 que la APA en conjunto con la AERA y NCME publicaron -las primeras recomendaciones para la construcción de pruebas psicológicas en un documento titulado *Technical recommendations for psychological tests and diagnostic techniques* (1954). Estas organizaciones han desarrollado cinco versiones de los estándares para pruebas educativas y psicológicas publicados en 1966, 1974, 1985, 1999 y 2014.

Durante el siglo XX, predominaron distintas conceptualizaciones sobre la validez en el campo de la medición psicológica y educativa (Markus y Borsboom, 2013). Por los años 60 y 80, se consideraba que había diferentes tipos de validez, lo que provocó a los desarrolladores de pruebas que eligieran indiscriminadamente el tipo de validez que mejor les pareciera para la creación de un instrumento psicométrico (Kane, 2013) en relación, principalmente, de la disponibilidad y tipos de datos. En la actualidad, se promueve un concepto unitario de la validez (Elosua, 2003; Montero, 2013) en el cual existen distintos tipos de evidencias de

validez que se recopilan en diferentes facetas de la validación de una prueba (AERA, APA y NCME, 2014).

Los Estándares para Pruebas Educativas y Psicológicas (AERA, APA y NCME, 2014) enumeran cinco fuentes de evidencia para respaldar las interpretaciones y usos propuestos de los puntajes de las pruebas: de contenido, de procesos de respuesta, de estructura interna, de relaciones con otras variables y de consecuencias. Por su parte, la evidencia basada en la estructura interna de las pruebas es “el grado en que las relaciones entre los elementos de la prueba y los componentes de la prueba se ajustan al constructo en el que se basan las interpretaciones de puntaje de prueba propuestas” (AERA, APA y NCME, 2014, p 13). Dicha definición, podría asimilarse al aspecto interno de la validez de constructo definida por Loevinger (1957), o a la representación del constructo apuntada por Embretson (1983). Actualmente, las evidencias de validez de constructo de la estructura interna de una prueba se centran en la evaluación de la dimensionalidad de la prueba y el análisis del funcionamiento diferencial de los ítems (Elosua, 2003).

La estructura interna considera tres aspectos básicos, dimensionalidad, invariancia de medición y confiabilidad. Al evaluar la dimensionalidad, está principalmente interesado el investigador en determinar si las interrelaciones entre los ítems apoyan los puntajes de los exámenes que se utilizarán para extraer inferencias. Es decir, una prueba que pretende informar una puntuación compuesta debe ser predominantemente unidimensional. Para la invariancia de la medición, es útil proporcionar evidencia de que las características del elemento como la discriminación y la dificultad del elemento, son comparables entre los grupos manifiestos, como el sexo o la raza. Por último, los índices de confiabilidad

proporcionan evidencia de que las puntuaciones de las pruebas reportadas son consistentes en las administraciones de pruebas repetidas (AERA, APA y NCME, 2014).

Los análisis de la estructura interna de una prueba pueden indicar el grado en que las relaciones entre los elementos de la prueba y los componentes de la prueba se ajustan al constructo en el que se basan las interpretaciones de la puntuación de la prueba propuesta. El marco conceptual para una prueba puede implicar una única dimensión de comportamiento, o puede postular varios componentes que se espera que sean homogéneos, pero que también sean distintos entre sí.

La medida en que las interrelaciones de los elementos confirman las presunciones del marco sería relevante para la validez. Los tipos específicos de análisis y su interpretación dependen de cómo se usará la prueba. De modo que, si una aplicación particular postula una serie de componentes de prueba cada vez más difíciles, se proporcionará evidencia empírica de la medida en que los patrones de respuesta se ajustaron a esta expectativa.

Por otra parte, la interpretación prevista para un uso dado, implica que el constructo debe estar relacionado con algunas otras variables y como resultado, los análisis de la relación de las puntuaciones de las pruebas con las variables externas a la prueba proporcionan otra fuente importante de evidencia de validez. Las variables externas pueden incluir medidas de algunos criterios que la prueba espera predecir, así como relaciones con otras pruebas hipotetizadas para medir las mismas construcciones y pruebas que miden construcciones relacionadas o diferentes. Las medidas distintas de las puntuaciones de las pruebas, como los criterios de rendimiento, se utilizan a menudo en entornos de empleo.

Las variables categóricas, incluidas las variables de pertenencia a grupos, se vuelven relevantes cuando la teoría que subyace en el uso de una prueba propuesta sugiere que las

diferencias de grupo deberían estar presentes o ausentes si se debe respaldar una interpretación de puntuación de prueba propuesta. La evidencia basada en las relaciones con otras variables proporciona evidencia sobre el grado en que estas relaciones son consistentes con la construcción subyacente a las interpretaciones de puntaje de prueba propuestas y puede involucrar evidencia tanto experimental como correlacional.

Las relaciones entre los puntajes de las pruebas y otras medidas destinadas a evaluar el mismo o constructos similares proporcionan evidencia convergente, mientras que las relaciones entre los puntajes de las pruebas y las medidas supuestamente de diferentes constructos proporcionan evidencia discriminante. Por ejemplo, dentro de algunos marcos teóricos, se espera que las puntuaciones en una prueba de selección múltiple de comprensión de lectura se relacionen estrechamente (evidencia convergente) con otras medidas de comprensión de lectura basadas en otros métodos, como las respuestas de ensayo. A la inversa, puede esperarse que los puntajes de las pruebas se relacionen menos estrechamente (evidencia discriminante) con las medidas de otras habilidades, como el razonamiento lógico. Las relaciones entre los diferentes métodos para medir el constructo pueden ser especialmente útiles para afinar y elaborar el significado y la interpretación de las puntuaciones.

La evidencia de la relación de las puntuaciones de las pruebas con un criterio relevante puede expresarse de varias maneras. El grado de precisión y el rango de puntaje dentro del cual se necesita la precisión depende del propósito para el cual se usa la prueba. Un tema importante en los entornos educativos es el grado en que la evidencia de validez basada en relaciones de criterio de prueba puede generalizarse a una nueva situación sin un estudio adicional de la validez en esa nueva situación. Cuando se usa una prueba para predecir los

mismos criterios o criterios similares (en diferentes momentos o en diferentes lugares, generalmente se encuentra que las correlaciones observadas en los criterios de prueba varían sustancialmente.

En el pasado, esto se ha tomado para implicar que siempre se requieren estudios de validación locales. Más recientemente, se ha desarrollado una variedad de enfoques para generalizar la evidencia de otros entornos, siendo el meta-análisis el más utilizado en la literatura publicada. En particular, los meta-análisis han demostrado que, en algunos dominios, gran parte de esta variabilidad puede deberse a artefactos estadísticos como las fluctuaciones de muestreo y las variaciones en los estudios de validación en los rangos de puntajes de prueba y en la confiabilidad de las medidas de criterio.

Cuando se toman en cuenta estas influencias, se puede encontrar que la variabilidad restante en los coeficientes de validez es relativamente pequeña. Por lo tanto, los resúmenes estadísticos de estudios de validación basados en situaciones similares pueden ser útiles para estimar las relaciones de criterio de prueba en una situación nueva. Esta práctica se conoce como el estudio de generalización de la validez.

### **2.3. Métodos estadísticos para el análisis de la estructura interna de instrumentos de medición**

Uno de los conceptos más importantes en el área de la medición es el de validez, el cual hace referencia al conjunto de evidencias teóricas y empíricas que fundamentan las interpretaciones de un test (AERA, APA y NCME, 1999). En particular las evidencias basadas en la estructura interna, hacen referencia a las relaciones entre los ítems y los componentes del constructo de interés. Lo esperado es que los ítems asociados a una dimensión presenten



correlaciones altas entre sí, mayores o superiores a las obtenidas entre los ítems de distintos componentes. Lo anterior indica que una evidencia de validez de constructo de la estructura interna de una prueba, se obtiene al comprobar que los ítems presentan patrones de intercorrelaciones que representan fielmente la estructura teórica del constructo a evaluar (Rojas 2014). Con lo mencionado anteriormente, es importante acentuar que la aportación de las evidencias de estructura interna es comprobar que las dimensiones teóricas del constructo medido se encuentren presentes en el test (Elosua, 2003). Con ello, una de las principales violaciones a la validez es la subrepresentación del constructo, es decir, que al menos una de los componentes del constructo no esté representada en las dimensiones del test (Messick, 1989b).

Las evidencias de validez de la estructura interna de una prueba son evaluadas mediante la técnica estadística de Análisis Factorial (AF) que es una estrategia matemática para probar la dimensionalidad de una escala, identificar la cantidad de factores que subyacen a los ítems y aproximarse a la validez de constructo (Streiner, 1994), esta puede variar, en algunos casos significativamente, según la población, la pertenencia de un ítem con un factor específico y la carga factorial de cada ítem (Rios y Wells, 2014). El AF se clasifica como una técnica estadística multivariada, la cual, es un modelo de regresión múltiple que relaciona variables latentes con variables observadas y busca esencialmente nuevas variables o factores que expliquen los datos (Ferrando y Anguiano-Carrasco, 2010). Es importante señalar que la dimensionalidad o estructura interna de una prueba se refiere a la cantidad de dominios, factores o dimensiones, en los cuales se agrupan un conjunto de ítems (Gorsuch, 1997). Por lo general, las pruebas se construyen con base en algunos supuestos o dimensiones teóricas que se deben probar empíricamente.

En general, se conocen dos tipos básicos de AF: el Análisis Factorial Exploratorio (AFE) y el Análisis Factorial Confirmatorio (AFC). En especial, el procedimiento del AFE, supone extraer, con criterios estadísticos, los factores, y rotar la solución inicial para así reducir y obtener la estructura factorial más simple y significativa (Pérez-Gil, Chacón y Moreno, 2000). Sin embargo, para Messick (1995) la aplicación del AFE no logra por sí sola aportar suficientes evidencias de validez de constructo de la estructura interna de una prueba considerándose una aproximación débil a la definición y/o validación de un constructo si dicha técnica no se acompaña con otras como el AFC. Ambas técnicas son por excelencia las más utilizadas para analizar el número de dimensiones presentes en un test y para la validación de constructo (Gómez, Chacón y Moreno 2000).

El AFE es una técnica estadística que permite revisar hipótesis sobre la posible estructura y dimensiones de un constructo (Bentler y Weeks, 1980; Bollen, 1989; Jöreskog y Sörbon, 1989). Dicha técnica se basa en analizar conjuntos de variables de las cuales se supone representan la medición de un constructo mediante patrones de asociaciones entre sí a los que se les llaman factores. Como resultado de la identificación de los patrones de asociaciones entre las variables que integran una medición, se puede asignar significados a dichos patrones con base en la teoría que fundamenta la construcción de los ítems. Esta técnica es útil cuando no se cuenta con una teoría consolidada sobre el constructo de interés. En contraparte, el AFC busca comprobar el ajuste de los datos a un modelo teórico previamente establecido. Esta técnica permite comprobar si los ítems se agrupan según las dimensiones que se definieron para su diseño y construcción (Gómez, Chacón y Moreno, 2000).

Las evidencias de validez basadas en el análisis de la estructura interna de una prueba, se complementan comúnmente con análisis psicométricos. Modelos como el de la Teoría Clásica de los Tests (TCT) y el de RASCH, requieren, idealmente, de estudios de los supuestos de unidimensionalidad para el sustento de las fórmulas matemáticas asociadas a sus planteamientos teóricos de medición (Burga, 2005). En ambos modelos psicométricos, se concibe un único rasgo latente o constructo subyacente a todos los ítems de una prueba (Wright y Linacre, 1995). Por su parte, en constructos multidimensionales como la actitud hacia las matemáticas, la cual presenta tres dimensiones básicas: cognitiva, comportamental y afectiva (Bazan y Sotero, 2000), los análisis usualmente se realizan de forma desagregada o por medio del análisis de Teoría de Respuesta.

Por último, las evidencias de validez de constructo de estructura interna de una prueba, tienen repercusiones a lo largo del proceso de validación y en específico en la aplicación de los modelos psicométricos. El no contar con las especificaciones de los componentes del constructo que subyace a una prueba, además de provocar problemas de tipo operativo, trae consigo inconvenientes asociados a otros aspectos de la validez. Un caso típico es el asociado al tema de la variancia irrelevante de constructo, es decir, la variación de los puntajes del test que no está asociada a variaciones de los puntajes del constructo (Montero, 2013).

### **2.3.1. Análisis Factorial Exploratorio (AFE)**

Los inicios del AFE remontan a los trabajos de Galton (1889) quien propuso el concepto de rasgo latente para explicar por qué un grupo de variables se encuentra relacionadas entre sí, fundamentándose en el hecho de que distintas variables alrededor de un fenómeno poseen algo en común y algo que las hace diferentes. Dado lo anterior, la varianza total explicada de

una variable se debe a la comunalidad y especificidad, es decir, a los factores que comparte con las otras variables y a factores propios de la variable (Pérez y Medrano, 2010). La lógica subyacente al AFE es, si un conjunto de variables se encuentra correlacionadas entre sí, estas se deben a que poseen un factor o rasgo latente en común, además, dicho factor explica en parte la varianza de las variables medidas (Blalock, 1966).

Por su parte, Pearson (1901) contribuyó en el desarrollo de los principios básicos del AFE con el desarrollo el coeficiente de correlación y los principios en que se basa el análisis factorial de componentes ejes principales. Por su parte, a Spearman (1904) también se le considera como el creador del AFE, ya que lo empleó con el fin de estudiar las correlaciones entre diferentes test de habilidades, en los cuales se reconocía un factor común de inteligencia. Los estudios de Spearman decantaron en la teoría bifactorial de la inteligencia, en donde se fundamenta un factor general de inteligencia (comunalidad del test), que subyacía de manera parcial con otras habilidades mentales (Pérez y Medrano, 2010).

A través de los años, distintos autores en diversos países retomaron la tarea de desarrollar la técnica del AFE entre los que se destacan los trabajos de Stephenson, Vermon y Eysenck en Inglaterra; Kelley, Hottelling, Cattell y Horn en Norteamérica; Meili en Alemania y Rimoldi en Argentina (Yela, 1996). A pesar de los distintos desarrollos y avances, la lógica operativa básica del AFE continúa siendo la misma, se considera la existencia de una serie de variables subyacentes inobservables, pero verificables a partir del uso de múltiples indicadores observables (Pérez y Medrano, 2010). El principal valor de este método consiste en reducir una gran cantidad de indicadores de escaso significado teórico, por un número menor de variables teóricamente significativas (Blalock, 1966). El AFE es una de las técnicas mayormente utilizadas en estudios relacionados con el desarrollo y validación de pruebas en

las que no se cuenta con información clara o precisa de la dimensionalidad del constructo a medir, debido a que es la técnica por excelencia para explorar el conjunto de variables latentes o factores comunes que explican las respuestas del conjunto de ítems de la prueba (Lloret-Segura, Ferreres-Traver, Hernández-Baeza y Tomás-Marco, 2014).

Previo a realizar un AFE es importante verificar tres supuestos principales: la normalidad, linealidad y multicolinealidad de las puntuaciones; los resultados del AFE y de los análisis estadísticos designados a evaluar los supuestos, pueden verse distorsionados por la presencia de casos con puntuaciones marginales (univariados y multivariados), por lo que se recomienda, en un primer momento, identificar casos atípicos o con valores extremos mediante una exploración inicial de los datos (Martínez Arias, 1999). Los casos atípicos son aquellos que presentan valores extremos, lo que produce un sesgo asimétrico del comportamiento del resto de los datos (Uriel y Aldas, 2005). El método mayormente utilizado para detectar casos atípicos o extremos univariados, consiste en calcular las puntuaciones típicas de cada variable y estimar como potenciales casos atípicos los que muestren puntajes Z fuera del rango  $\pm 3$  (Tabachnick y Fidell, 2001). Pagano (1998) señala que la media es un estadístico sensible a los valores extremos, un valor lejano de la tendencia central produce desplazamientos de la media. Así mismo, si la media aritmética se ve distorsionada también serán afectadas las correlaciones entre las variables, y el AFE.

Un acercamiento alternativo al análisis de los valores extremos es la inspección visual de los gráficos de box plots o diagrama de caja, el cual, es una forma estandarizada de mostrar la distribución de datos basada en el resumen de cinco números: mínimo, primer cuartil, mediana, tercer cuartil y máximo; en conjunto presentan los valores atípicos como puntos aislados en los extremos. En cuanto a la distribución, los índices de asimetría y curtosis

indican variaciones leves de la curva normal y en consecuencia resultan adecuados para realizar el AFE considerando que los valores se mantengan dentro del umbral  $\pm 1.5$  (George y Mallery, 2001).

El supuesto de linealidad de las relaciones entre variables resulta primordial en el AFE. Los valores del coeficiente de correlación utilizados en los AFE sólo pueden interpretarse cuando el patrón de relaciones entre las variables es lineal (Batista Foguet y Gallart, 2000). Para la evaluación estadística de este supuesto, se realiza un Análisis de Regresión Múltiple con una estimación curvilínea, utilizando funciones lineales en el caso de no obtener resultados significativos, y utilizando funciones cuadráticas en el caso de que no se obtengan, para ambos casos se puede llegar a la conclusión de que no hay una relación lineal entre las variables (Pérez y Medrano, 2010).

Del mismo modo, es recomendable realizar un diagnóstico de multicolinealidad entre las variables con el propósito de identificar correlaciones extremas o redundantes. Según Martínez Arias (1999), el AFE requiere de la intercorrelación entre variables, si estas son superiores o iguales a .90 es posible que se debilite el análisis y obtenga una solución factorial poco estable. Para evaluar la multicolinealidad se debe observar la matriz de correlación, atendiendo la presencia de valores iguales o mayores a .90 (Pérez y Medrano, 2010).

Además de los supuestos estadísticos requeridos por el AFE, existen algunos requerimientos adicionales de gran importancia para la realización de este análisis. El AFE se basa en la matriz de intercorrelaciones por lo que es necesario que la muestra sea de gran tamaño para asegurar un menor error de muestreo, ya que si se trabaja con muestras pequeñas aumenta la probabilidad de que las correlaciones se perturben de una muestra a otra, se obtengan factores inestables y los resultados sean incorrectos (Blalock, 1966). El AFE

debe conducirse empleando muestras grandes, de 300 participantes en adelante, para lograr resultados útiles y estables (Tabachnick y Fidell, 2001). Lo ideal sería contar con mínimo 10 participantes por variable (Nunnally y Bernstein, 1995). Para los casos de muestras grandes o de gran escala es recomendable realizar un análisis factorial diferenciado por desagregados como el sexo o grupo étnico (Kline, 2000).

Otro criterio previo al AFC es determinar si los ítems están suficientemente interrelacionados. En el caso del estudio de la estructura interna de una prueba, una vez obtenidos los datos de la muestra seleccionada, se debe determinar si los ítems están suficientemente interrelacionados para que el AFE pueda aplicarse convenientemente (Comrey, 1973). Para ello, las pruebas estadísticas más utilizadas son el test de esfericidad de Bartlett y la medida de adecuación muestral de Kaiser-Mayer-Olkin (KMO). El test de esfericidad de Bartlett evalúa la hipótesis nula que afirma que las variables no están correlacionadas. Para ello, dicho método compara la matriz de correlación de los datos obtenidos en una aplicación con una matriz de identidad, en la que todos los términos de la diagonal son unidades y el resto de los términos son ceros. Si los resultados que se obtuvieron de la comparación resultan significativos a un nivel de probabilidad  $p < .05$ , se rechaza la hipótesis nula y se considera que las variables están lo suficientemente correlacionadas para realizar el AFE (Everitt y Wykes, 2001). Dicha prueba puede revelar resultados significativos a pesar de no haber correlaciones altas entre las variables, por ello se sugiere la utilización a la par de la medida KMO (Pérez y Medrano, 2010).

Kaiser (1970) introdujo una Medida de Adecuación de Muestreo (MSA por sus siglas en inglés) de matrices de datos analíticos de factores. Kaiser y Rice (1974) luego lo modificaron a KMO dejándolo como un promedio de los términos de la diagonal de la matriz

de correlación anti-imagen, la cual contiene los valores negativos de los coeficientes de correlación parcial de todas las variables en estudio. La lógica del índice KMO es que, si las variables comparten factores comunes, los coeficientes de correlación parcial deben ser pequeños y los valores de la diagonal de la matriz deben ser altos, es decir, si el porcentaje de coeficientes altos es grande en la matriz anti-imagen, mayor es la interrelación entre las variables (Sierra Bravo, 1981). La interpretación del KMO es similar a los coeficientes de confiabilidad, con un rango de 0 a 1 y considerando como adecuado un valor igual o superior a .70, el cual indica una correlación satisfactoria entre las variables (Hair, Anderson, Tatham, y Black, 1999).

Posterior a la verificación del cumplimiento de los supuestos previos para la aplicación del AFE se debe seleccionar un método de extracción de factores donde se trata de reducir la información contenida en las variables superficiales a un número pequeño de variables latentes. Los más utilizados son el método de componentes principales, mínimos cuadrados no ponderados, mínimos cuadrados generalizados, máxima verosimilitud y ejes principales, alfa (Khan, 2006). En especial, el método de componentes principales es el que explica la mayor cantidad de varianza posible en los datos observados. Este método analiza la varianza total asociada a las variables, incluyendo la varianza específica y la varianza de error.

Por su parte, el método de ejes principales, solo toma en cuenta la varianza que las variables tienen en común, también llamada covarianza, excluyendo a la varianza específica y la varianza del error de medición (Tabachnick y Fidell, 2001). Tiene la ventaja de recuperar factores débiles y es recomendable para soluciones factoriales con pocas variables. Además, con tamaños muestrales pequeños y correlaciones moderadas entre las variables, ofrece



soluciones más estables. Sin embargo, una de sus limitaciones es que no ofrece índices de ajuste (Beauducel, 2001).

El método de extracción mínimos cuadrados no ponderados, es una técnica para la estimación de los parámetros desconocidos en un modelo de regresión lineal, se aplica cuando las varianzas de las observaciones son desiguales o presenta heterocedasticidad y cuando existe un cierto grado de correlación entre las observaciones (Greene, 2000). Otro método de extracción muy utilizado es el de máxima verosimilitud, el cual, proporciona las estimaciones de los parámetros que con mayor probabilidad ha producido la matriz de correlaciones observada, asumiendo que la muestra procede de una distribución normal multivariada. El método de ejes principales es un método iterativo de estimación basado en la extracción sucesiva de aquellos factores que explican la mayor parte de la varianza común (Winter y Dodou, 2012; Tabachnick y Fidell, 2001).

El resultado inicial de los métodos de extracción en el AFE es una matriz factorial no rotada. Esta matriz factorial sin rotar presenta una serie de limitaciones de interpretación y en la mayoría de los casos en donde se extrae más de un factor, es indispensable una matriz adicional de factores rotados. El término rotación de factores proviene de la representación geométrica del análisis factorial; en donde los factores son representados en los ejes de referencia y los pesos factoriales de las correlaciones de cada variable se ubican como coordenadas entre los ejes correspondientes (Carroll, 1953; Pérez y Medrano, 2010). La rotación se lleva a cabo para que la solución factorial se aproxime a lo que comúnmente se denomina estructura simple, en donde cada variable con una correlación cercana a 1 se acerca con uno de los factores y se aleja de los factores restantes con correlaciones cercanas a 0. La rotación se utiliza entonces para eliminar las correlaciones negativas altas y reducir el

número de correlaciones de cada variable en los diversos factores. Sin embargo, hay que entender que las estructuras simples son modelos teóricos de una solución aproximada (Pérez y Medrano, 2010).

Ahora bien, las rotaciones factoriales pueden ser oblicuas u ortogonales. Para ello, los métodos más utilizados son el Varimax (Kaiser, 1958) y el Promax (Gorsuch, 1983). Las soluciones resultantes de los métodos de rotación oblicua son más congruentes con la estructura de las variables que se encuentran correlacionadas. Por su parte, la rotación ortogonal absoluta es sólo teórica y se interpreta como una solución en donde todas las correlaciones entre los factores son menores a .32. Tabachnick y Fidell (2001) proponen realizar una rotación oblicua inicial con un método Promax como filtro y subsecuentemente obtener la matriz de correlación entre los factores. Consecutivamente, si se observan correlaciones superiores a .32 entre los factores se escoge una rotación oblicua, y menores, una ortogonal.

Con las rotaciones se colocan las variables más cerca de los factores obtenidos ayudando a su explicación, para lo cual se concentra la varianza de las variables en menos factores posibilitando la interpretación de la solución factorial (Kaiser, 1958). En la actualidad hay varios algoritmos en los paquetes estadísticos que pueden ayudar a generar la matriz rotada sin recurrir a procedimientos gráficos de rotación (Thompson, 2004).

Como último procedimiento del AFE se interpretan y denominan los factores, examinando en la matriz rotada siempre que se trate de más de un factor, el patrón de correlaciones bajas y altas de cada variable con los distintos factores, utilizando el conocimiento teórico que se posea acerca de las variables incluidas en el análisis. Es recomendable que cada factor se configure con al menos tres o cuatro ítems con correlaciones

iguales o superiores a .45 para ser interpretado, además de focalizar las correlaciones ítem-factor más altas y con ello tener la posibilidad de definir cada factor de la solución (Glutting, 2002). El AFE, aunque se considera una técnica válida con fines exploratorios, lleva a resultados ciertamente azarosos y en muchas de las ocasiones poco estables especialmente en los casos en donde se desconoce el constructo que se pretende validar (Pérez y Medrano, 2010).

### **2.3.2. Análisis Factorial Confirmatorio (AFC)**

El AFC se ha convertido en uno de los procedimientos estadísticos más utilizados en la investigación aplicada, Brown (2006) lo definió como un tipo de modelo de ecuación estructural (SEM por sus siglas en inglés) que trata específicamente los modelos de medición, es decir, las relaciones entre medidas o indicadores observados (p. Ej., Ítems de prueba, puntajes de prueba, calificaciones de observación conductual) y variables latentes o factores. Una característica fundamental del AFC es su naturaleza impulsada por hipótesis.

El método más apropiado y estadísticamente más riguroso para obtener evidencias de validez de constructo de la estructura interna de una prueba es el AFC. Para dicha técnica se considera una estructura factorial hipotética en términos de bondad de ajuste que se contrasta con la estructura observada en la muestra de estudio (puntuaciones de las respuestas a los ítems), donde altos índices de bondad de ajuste se pueden interpretar como fuertes evidencias de validez de constructo de la prueba en estudio (APA, 2013).

Brown (2006) señaló que, en el AFC, un constructo es un concepto teórico. En psicología clínica y psiquiatría los trastornos mentales (por ejemplo, depresión mayor, esquizofrenia) son constructos manifestados por varios grupos de síntomas que son

informados por el paciente u observados por otros. En sociología, la delincuencia juvenil podría interpretarse como una construcción multidimensional definida por diversas formas de mala conducta (por ejemplo, delitos contra la propiedad, violencia interpersonal, uso de drogas y mala conducta académica). El AFC es una herramienta analítica indispensable para la validación de construcciones en las ciencias sociales y del comportamiento.

Casi siempre, el AFC se utiliza durante el proceso de desarrollo de escala para examinar la estructura latente de un test, verificar el número de dimensiones o factores y la relación de cargas factoriales. Este análisis también ayuda a determinar cómo se debe calificar una prueba, cuando la estructura latente es multifactorial (dos o más factores), el patrón de cargas de factor respaldado por AFC designará cómo se puede puntuar una prueba usando subescalas; es decir, el número de factores es indicativo del número de subescalas y el patrón de relaciones elemento-factor (qué elementos cargan sobre qué factores) indica cómo deben puntuarse las subescalas. Dependiendo de otros resultados y extensiones del análisis, puede apoyar el uso de puntajes totales (compuestos de todos los ítems), además de los puntajes de subescalas compuestos de subconjuntos de ítems (Brown, 2006).

A diferencia del AFC, el AFE no permite definir qué ítems corresponden con los factores, tampoco las relaciones que posiblemente hay entre los propios factores, más allá de si están o no relacionados entre sí. Es por ello que el AFE es una técnica exploratoria que permite determinar la cantidad de factores esperados para la solución, pero no su composición ni las relaciones que cada uno de los factores puedan tener con el resto. Al contrario, el AFC sí permite definir la cantidad de factores esperados, qué factores están relacionados entre sí, y qué ítems se espera estén relacionados con cada factor (Lloret-Segura, Ferreres-Traver, Hernández-Baeza y Tomás-Marco, 2014).

En el AFC inicialmente se especifica el modelo, se estiman los parámetros, y se evalúa el ajuste del modelo. Consecutivamente, si el modelo propuesto de la estructura de la prueba presenta un ajuste adecuado, se entiende que se pueden evaluar e interpretar los parámetros. En caso contrario, si el modelo propuesto no se ajusta adecuadamente, se puede reconfigurar el modelo y de nuevo iniciar el proceso (Batista y Coenders, 2000; Barbero, Vila y Holgado, 2011; Catena, Ramos y Trujillo, 2003; Fadlelmula, 2011). Esto último, es importante se tome en cuenta siempre los presupuestos teóricos con los que se fundamentó la estructura original de la prueba.

La falta de coherencia entre las características del instrumento de medición y el cumplimiento de los requisitos para el uso del AFC cobran importancia en la estimación de los parámetros. El grado de congruencia y de fidelidad estructural de una prueba con la realidad a medir depende en gran medida de los métodos empleados para las estimaciones de los parámetros. De la elección adecuada del método de estimación dependerá que, al evaluar el ajuste del modelo especificado, tenga lugar una mayor o menor aproximación a la realidad empírica mediante el logro de los criterios establecidos para los índices de bondad de ajuste (Morata-Ramírez, Holgado-Tello, Barbero-García y Méndez, 2015).

Los resultados del AFC pueden proporcionar evidencia de la validez convergente y discriminante de constructos teóricos. La validez convergente está indicada por la evidencia de que los diferentes indicadores de constructos teóricamente similares o superpuestos están fuertemente interrelacionados; por ejemplo, síntomas que supuestamente son manifestaciones de una sola carga de trastorno mental sobre el mismo factor. Por otra parte, la validez discriminante está indicada por los resultados que muestran que los indicadores de constructos teóricamente distintos no están altamente correlacionados; por ejemplo, los

comportamientos que se supone son manifestaciones de diferentes tipos de carga de delincuencia en factores separados, y los factores no están tan altamente correlacionados como para indicar que una construcción más amplia se ha separado erróneamente en dos o más factores (Brown, 2006). En la **Tabla 1** la clasificación de los índices de bondad de ajuste en el AFC divididos de acuerdo a su categoría inferencial o descriptivo, absoluto o comparativo, su índice de ajuste y el criterio de interpretación.

**Tabla1.** Clasificación de los índices de bondad de ajuste en el AFC

Alcance	Tipo de ajuste	Índice de ajuste	Criterios de interpretación
Inferencial	Absoluto	Estadístico $\chi^2$	Se mantiene el modelo si $p \geq 0.05$ ; es sensible a grandes muestras
	Comparativo	Comparación de modelos anidados ( $\chi^2$ )	Los modelos no difieren si $p \geq 0.05$
Descriptivo	Absoluto	$\chi^2/gl$	Ratios menores que 2 ó 3 suelen considerarse indicadores de ajuste aceptable. Es sensible a grandes muestras
		RMR	Cuanto más pequeño, mejor ajuste; 0 indica ajuste perfecto; no se recomienda porque es difícil de interpretar.
		SRMR	$\leq 0.08$ , el modelo se ajusta.
	Comparativo	RMSEA	$\leq 0.06$ , el modelo se ajusta
		GFI/AGFI	$\geq 0.95$ , el modelo se ajusta; no se recomienda
		NFI	$\geq 0.95$ : el modelo ajusta. En muestras pequeñas tiende a sobrestimarse.
		NNFI o TLI	$\geq 0.95$ (para variables continuas o $\geq 0.096$ (para variables categóricas): el modelo se ajusta.
	CFI	$\geq 0.95$ (para variables continuas o variables categóricas): el modelo se ajusta.	

Nota: Adaptado de "Medición de ciencias sociales y de la salud" por F. Abad, J. Olea, V. Ponsoda y C. García, 2011, Síntesis, Madrid, España.

### 2.3.3. Software para el Análisis Factorial

Para la obtención de evidencias de validez de estructura interna y de asociación con otras variables hay una diversidad de paqueterías estadísticas computarizadas que pueden ser utilizadas para facilitar los análisis y estimaciones de distintos modelos matemáticos. Los más recomendados para la aplicación de las técnicas estadísticas para el análisis de la estructura interna de pruebas se pueden dividir en tres categorías: (1) de uso exclusivo para análisis exploratorios, (2) de uso exclusivo para análisis confirmatorios, y (3) aquellos que son de uso general pero que permiten llevar a cabo ambos tipos de análisis.

Entre los programas especializados para realizar análisis factoriales exploratorios se puede mencionar el FACTOR. Por su parte, entre los programas especializados para realizar

análisis factoriales confirmatorios entre los más usados se encuentran el EQS, el LISREL y el paquete LAVAAN de R Statistics. En cuanto a los programas estadísticos de uso general que permiten la aplicación de ambos análisis factoriales encontramos se encuentran entre los más usados: el IBM SPSS STATISTICS, el IBM SPSS AMOS, Mplus, SYSTAT 13.1 y MINITAB.

El software Factor es utilizado para realizar diversos cálculos implicados en el análisis factorial, permite efectuar y ajustar el modelo de análisis factorial exploratorio utilizando el análisis factorial exploratorio de máxima probabilidad. Por su parte, para los análisis factoriales confirmatorios, entre los más utilizados se encuentra el EQS Implementa aproximaciones matemáticas y estadísticas al análisis de sistemas de ecuaciones de estructura lineal. Fue diseñado para hacer a los métodos de análisis avanzado multivariables accesibles para ser aplicados en análisis de datos estadísticos y prácticos. El EQS hace posible quitar complicaciones del modelado estructural al proveer una aproximación simple, consistente, técnica, avanzada y precisa al modelado estructural. El EQS provee a los investigadores y estadísticos de un método simple para conducir todo el rango de modelos de ecuaciones estructurales, incluyendo las regresiones múltiples, las regresiones multivariadas, en análisis factorial confirmatorio, análisis de medios estructurados, análisis de trayectoria, y comparaciones de población múltiple.

El LISREL es una aplicación programa proveída por Windows para llevar a cabo modelado de ecuaciones estructurales (SEM), y otros modelados de estructura lineal relacionados (ej., modelado de estructuras multinivel, modelado multinivel lineal y no lineal, etc.). LISREL para Windows es útil importando los datos externos en varios formatos como SPSS, SAS, MS Excel, etc. Como un archivo de sistema PRELIS (PSF). El software usa archivos gráficos con la extensión predeterminada llamada PTH de manera que pueda



capturar el diagrama de trayectoria. Es también útil en ajustar el modelo de medición a los datos. LISREL estima parámetros usando ecuaciones simultáneamente. Esto hace posible estimar muchos parámetros en estructuras de estimación complejas. Quizás la mayor ventaja de estos es que uno puede distinguir entre variables latentes y variables observables.

El IBM SPSS STATISTICS es un sistema global para el análisis de datos. Dicho software puede analizar datos de casi cualquier tipo de archivo y utilizarlos para generar informes estadísticos de análisis descriptivos y complejos, así como representarlos en formato tabular y gráfico. En especial el IBM SPSS AMOS (Analysis of Moment Structures) es un módulo del programa fácil de usar para visualizar el modelo de ecuaciones estructurales (SEM). Con dicho modulo AMOS, se puede rápidamente especificar, ver, y modificar un modelo factorial gráficamente usando sencillas herramientas de trazado. Después se puede evaluar el ajuste del modelo, hacer cualquier modificación, e imprimir un gráfico del modelo final con calidad de publicación.

El Amos, permite utilizar fácilmente el SEM para probar hipótesis complejas sobre relaciones de un gran grupo de variables y obtener nuevos puntos de vista de los datos. Es un software que da soporte a investigación y teorías ampliando los métodos de análisis multivariado estándar que incluye regresión, el análisis de factores, la correlación y el análisis de varianza. Con SPSS AMOS se pueden crear modelos de comportamiento y actitudes que reflejen con precisión las relaciones complejas mediante una interfaz de usuario programática o grafica intuitiva.

Implementa la aproximación general al análisis de datos conocido como modelado de ecuaciones estructurales, también conocido como análisis de estructuras de covarianza, o modelado causal. Esta aproximación incluye, casos especiales, y muchas técnicas

convencionales bien conocidas, incluyendo el modelo lineal general y el análisis factorial común.

El R statistical es un lenguaje y entorno para estadísticas computacionales y gráficos. Es un proyecto de GNU que es similar al lenguaje S y entorno que fue desarrollado en los laboratorios Bell (formalmente en ATyT, ahora Lucent technologies) por John Chambers y colegas. Este programa provee una amplia variedad de técnicas estadísticas (modelado lineal y no lineal, pruebas estadísticas clásicas, análisis de tiempo-series, clasificación, agrupamiento) y gráficas, y es altamente extenso. El lenguaje S es comúnmente el vehículo de elección de investigadores y metodólogos estadísticos, y R provee una fuente abierta para participar en esa actividad. El paquete LAVAAN fue desarrollado para proveer a los usuarios de R, investigadores y profesores una fuente abierta gratuita, es un paquete de calidad comercial para el modelado de variables latentes. Se puede usar LAVAAN para estimar una variedad amplia de modelos estadísticos multivariados, incluyendo el análisis de trayectoria, análisis factorial confirmatorio, modelado de ecuaciones estructurales y modelos de crecimiento de curva.

Una de las fortalezas del software es la facilidad con la que tramas bien diseñadas con calidad de publicación pueden ser producidas, incluyendo símbolos matemáticos, y formulas donde sean necesarias. Gran cuidado ha sido tomado en cuenta hasta en el más mínimo error en la elección del diseño de los gráficos, pero los usuarios mantienen el control total. El programa está disponible como un software gratuito bajo los términos de la Free Software Foundation's GNU General public license, en forma de código fuente. Compila y corre en una amplia variedad de plataformas UNIX, y sistemas similares (incluyendo FREEBSD y Linux), Windows y MacOS.

El MPLUS es un programa de modelado estadístico que provee a investigadores con una herramienta flexible para analizar sus datos. Ofrece a los investigadores con una amplia opción de elección de modelos, estimadores y algoritmos en un programa que tiene una interfaz de fácil uso y muestras graficas de los datos y los resultados de análisis. También permite el análisis tanto de sección-transversal y de datos longitudinales, datos de nivel singular y multinivel, datos que vienen de diferentes poblaciones tanto con heterogeneidad observada como no observada, y datos que contienen valores faltantes.

Los análisis pueden ser llevados a cabo para variables observadas que sean continuas, censuradas, o combinaciones de este tipo de variables. Este software tiene capacidades extensas para estudios de simulación de Monte Carlo, donde los datos pueden ser generados y analizados de acuerdo a cualquiera de los modelos incluidos en el programa.

El SAS/STAT permite realizar análisis de varianza, regresión lineal, hasta inferencias Bayesianas, con herramientas de modelado de alto-desempeño para datos masivos. A su vez provee herramientas tanto para necesidades estadísticas especializadas y de toda la empresa. Analiza cualquier tipo y tamaño de datos, incluye técnicas exactas para conjuntos pequeños de datos, herramientas de modelado estadístico de alto-desempeño para tareas de datos amplios y métodos modernos para analizar datos con valores faltantes. Entre los softwares más reconocidos para la aplicación del AFE se encuentra el SYSTAT 13.1 (SYSTAT Software, 2010). Dicho software permite importar matrices en diversos formatos incluyendo las diseñadas en el SPSS. Este programa posee la opción de calcular una multiplicidad de matrices de correlaciones como las matrices tetracóricas. Pero no puede aún calcular las matrices policóricas.

El MITITAB es un programa estadístico muy potente, versátil y de fácil uso, que proporciona un amplio rango de aplicaciones estadísticas, capacidad de gráfico. Capacidad estadística comprensiva y completa, que incluye análisis de datos exploratorios, cálculos básicos, regresión de análisis de datos exploratorios, análisis de varianza, tamaño de muestra, análisis multivariante, distribuciones no-paramétricas, series de tiempo, tabulación cruzada y simulación.

---

### III. MÉTODO

El presente capítulo se organiza en tres sub-apartados en donde se describe el método desarrollado para el logro de los objetivos de investigación. En el primer sub-apartado se describe el tipo de estudio realizado para analizar las evidencias de validez de constructo, en el aspecto de la estructura interna de la ERECVE. En el segundo y tercero sub-apartados se describen el instrumento en estudio y la base de datos producto de la estrategia evaluativa integral 2016-2 realizada por investigadores de la Unidad de Evaluación Educativa (UEE) en colaboración con autoridades educativas de la Secretaría de Educación del Estado de Baja California (SEEBC). En el cuarto y último sub-apartado del capítulo, se describen las etapas, actividades y distintos análisis realizados a lo largo del estudio.

#### 3.1. Tipo de estudio

Para alcanzar los objetivos de investigación se realizó un estudio psicométrico con el fin de obtener evidencias de validez de constructo, en el aspecto de la estructura interna de la Escala del Rol de los Estudiantes en el Ciclo de la Violencia Escolar (ERECVE). Para ello se siguieron los *Estándares para las pruebas educativas y psicológicas* de la AERA, APA y NCME (1985 y 1999), así como las recomendaciones de autores reconocidos en el campo de la teoría de la validez de constructo (Messick, 1989b; Kane, 2016; Mislevy, 2009) y de los métodos para el análisis de la estructura interna de las pruebas con base en Análisis Factoriales (AF) (Yela, 1966; Bentler y Weeks, 1980; Long, 1983; Bollen, 1989; Jöreskog y Sörbon, 1989). En especial se decidió aplicar el Análisis Factorial Exploratorio (AFE), el Análisis Factorial Confirmatorio (AFC), así como análisis preliminares de normalidad, linealidad y multicolinealidad de las

puntuaciones, y de confiabilidad, discriminación y correlación entre sub-escalas y otras variables.

### 3.2. Instrumento

El instrumento del cual se analizaron las evidencias de validez de constructo es la Escala del Rol de los Estudiantes en el Ciclo de la Violencia Escolar (ERECVE). Dicho instrumento fue diseñado por Pérez-Morán y Rodríguez-Macías para su aplicación en la estrategia evaluativa integral UEE 2016-2 en los cuatro municipios del Estado de Baja California (Rodríguez-Macías, Pérez-Morán, y Sarabia-Ocampo, 2017). La escala se diseñó a partir de las categorías basadas en el enfoque de estudio del acoso escolar como proceso grupal propuesto Olweus (1993) y extendido por Salmivalli (1999, 2010) y que pueden observarse en el *Participant-Role Questionnaire* (Salmivalli, Lagerspetz, Björkqvist, Österman, y Kaukiainen, 1996). En total la ERECVE consta de 19 ítems, con cinco opciones de respuesta. El alumno de secundaria debe señalar la frecuencia con que ocurren una serie de eventos relacionados con distintos tipos de roles en el ciclo de la violencia escolar. Las opciones de respuesta son: a) Nunca, b) Raramente, c) Mensualmente, d) Semanalmente y e) Diariamente.

Por su parte, la estructura de la ERECVE se compone de cinco sub-escalas. la primera sub-escala, referente al rol de la *víctima* en el ciclo de la violencia escolar, la integran cinco ítems. La segunda sub-escala, correspondiente al rol del *agresor solitario*, la conforman tres ítems. La tercera sub-escala, relacionada con el rol del *agresor social* la componen cuatro ítems. La cuarta sub-escala, conformada por tres ítems, corresponde al rol del *observador*. Por último, la quinta sub-escala, relacionada con el rol de *pacificador* o *defensor*, la integran

cuatro ítems. En la **Tabla 2** se pueden observar los ítems agrupados por cada una de las sub-escalas de la ERECVE, así como la posición del ítem en la escala y su clave de respuesta.

**Tabla 2.** Ítems y sub-escalas que conforman la Escala del Rol de los Estudiantes en el Ciclo de la Violencia Escolar (ERECVE)

Sub-escala	No. de ítem	Ítem	Clave
<b>Sub-escala del Rol de Víctima en el Ciclo de la Violencia Escolar (SRVCVE)</b>	S144.1	Sufro violencia por parte de mis compañeros(as) de clase.	+
	S144.2	Sufro violencia por parte de compañeros(as) de otros salones de clase.	+
	S144.3	Sufro violencia por parte de alguno de los docentes de mi escuela.	+
	S144.4	Sufro violencia por parte de algún miembro del personal administrativo de la escuela.	+
	S144.5	Sufro violencia por parte de compañeros(as) fuera de mi escuela o en la comunidad en donde vivo.	+
<b>Sub-escala del Rol del Agresor Solitario en el Ciclo de la Violencia Escolar (SRAS1CVE)</b>	S144.6	Practico violencia sobre mis compañeros(as) de clase.	+
	S144.7	Practico violencia sobre compañeros(as) de otros salones de clase.	+
	S144.8	Practico violencia sobre compañeros(as) fuera de mi escuela o en la comunidad en donde vivo.	+
<b>Sub-escala del Rol del Agresor Social en el Ciclo de la Violencia Escolar (SRAS2CVE)</b>	S143.1	Me reúno con un grupo de amigos(as) para practicar violencia sobre compañeros(as) de mi escuela.	+
	S143.3	Me reúno fuera de mi escuela con un grupo de amigos(as) para realizar actos vandálicos.	+
	S143.5	Realizo actos de vandalismo para recibir la admiración de mis amigos(as).	+
	S143.7	Practico violencia sobre otros compañeros(as) para sentirme superior a ellos.	+
<b>Sub-escala del Rol de Observador en el Ciclo de la Violencia Escolar (SROCVE)</b>	S144.9	Veo violencia entre mis compañeros(as) de clase.	+
	S144.10	Veo violencia entre compañeros(as) de otros salones de clase.	+
	S144.11	Veo violencia entre compañeros(as) de mi escuela y de otras escuelas o de la comunidad en donde vivo.	+
<b>Sub-escala del Rol de Pacificador en el Ciclo de la Violencia Escolar (SRPCVE)</b>	S143.2	Me reúno con un grupo de compañero(as) para realizar actividades extraescolares.	-
	S143.4	Me reúno fuera de mi escuela con un grupo de amigos(as) para realizar actividades deportivas o culturales.	-
	S143.6	Realizo actividades de voluntariado con el fin de ayudar a personas en condiciones de vulnerabilidad.	-
	S143.8	Ayudo a otros compañeros(as) que se encuentran en problemas para mejorar su situación.	-

### 3.3. Base de datos

Se utilizó para el análisis de la estructura interna de la ERECVE una base de datos conformada por 203 variables. Dicha base de datos se integra por las respuestas de alumnos de secundaria que participaron en la aplicación de la estrategia evaluativa integral 2016-2 coordinada por los investigadores de la UEE en colaboración con autoridades educativas de la SEEBEC en los cinco municipios del Estado de Baja California (Rodríguez-Macías, et al., 2017). La base de datos se conforma por 15,861 casos referidos a alumnos de Educación Secundaria Obligatoria (ESO), de los cuales el 49.7% son mujeres y el 50.3% hombres. La edad de los alumnos encuestados oscilaba entre 12 y 18 años, siendo el 23.7% de 1º, 24.1% de 2º, 26.0% de 3º y el 26.2% de 4º de ESO.

Por su parte, del 100% de los casos registrados en la base de datos, el 65.5% corresponde a alumnos del turno matutino, el 0.1% al turno intermedio, el 5.5% al turno completo, y el 28.4% y 5% a alumnos de los turnos vespertino y nocturno respectivamente. En cuanto a la variable del tipo de modalidad de educación secundaria, el 50.2% de los alumnos encuestados pertenecen a la modalidad general, el 28.8% a la técnica, el 7.4% a la telesecundaria, el 9.0% a la privada, y el 4.2% y 0.3% a las modalidades indígena y de trabajadores respectivamente.

### 3.4. Procedimiento

El procedimiento para llevar a cabo el estudio psicométrico se organizó en cuatro etapas: (I) Análisis preliminar, (II) Análisis Factorial Exploratorio (AFE), (III) Análisis Factorial Confirmatorio (AFC), y (IV) Análisis de correlación convergente y discriminante. En la **Tabla 3**



se señalan las actividades del estudio organizadas en las cinco etapas mencionadas, así como los análisis y algunos estadísticos reportados a lo largo de las etapas del estudio. Cabe señalar que para la realización de los análisis estadísticos se utilizaron diversos softwares entre los que se encuentran el paquete estadístico IBM SPSS versión 21, el Structural Equation Modeling Software (EQS 6.1) (Bentler, 2012) y el R Statistical X64 V 3.5.2 junto con las paqueterías estadísticas *psych* (Revelle, 2008) y *lavaan* (Rosseel, 2012).

**Tabla 3. Modelo de etapas y actividades**

Etapas	Actividades
<b>I. Análisis preliminares</b>	1.1. Eliminación de valores perdidos y datos erróneos.
	1.2. Análisis descriptivo de las puntuaciones totales de la escala y por ítem
	1.3. Exploración de la normalidad y eliminación de casos con puntuaciones atípicas.
	1.4. Verificación de los supuestos de normalidad, linealidad y multicolinealidad de las puntuaciones.
	1.5. Análisis de confiabilidad y discriminación mediante los supuestos de la TCT.
<b>II. Análisis Factorial Exploratorio (AFE)</b>	2.1. Selección de los métodos de extracción.
	2.2. Aplicación de los métodos de extracción por Componentes Principales y Máxima Verosimilitud (maximum likelihood).
	2.3. Definición teórica del número de factores a extraer.
	2.4. Definición del modelo con base en evidencia empírica utilizando el criterio de Kaiser, el % de varianza explicada y el gráfico de sedimentación (scree test).
	2.5. Selección del método de rotación de la matriz factorial.
	2.6. Aplicación de la rotación ortogonal y oblicua con los métodos VARIMAX y PROMAX respectivamente.
	2.7. Interpretación y denominación teórica de los factores
<b>III. Análisis Factorial Confirmatorio (AFC)</b>	3.1. Definición gráfica del modelo factorial confirmatorio.
	3.2. Formalización matemática e identificación del modelo factorial confirmatorio.
	3.3. Estimación del modelo factorial confirmatorio definido.
	3.4. Análisis del ajuste del modelo factorial confirmatorio estimado.
	3.5. Interpretación y reespecificación, de ser necesario, del modelo factorial confirmatorio de origen.
<b>IV. Análisis de correlación concurrente y discriminante</b>	4.1. Análisis de correlación entre las Subescalas de la ERECVE
	4.2. Análisis de correlación de las subescalas de la ERECVE con las subescalas de Violencia Intrafamiliar (SEVIN), Ambiente Escolar Violento (SAEV) Rol del Docente en el Ciclo de la Violencia Escolar (SRDCVE), y Rol del Director en el Clima Escolar (SRDCE)

Nota: Elaboración propia

### 3.4.1. Etapa I. Análisis preliminares

La primera etapa del análisis preliminar de los datos se organiza en cuatro actividades: (1) eliminación de valores perdidos y datos erróneos, (2) análisis descriptivos del total de la escala, por sub-escala y por cada ítem, (3) análisis de normalidad, linealidad y multicolinealidad de las puntuaciones, y (4) análisis de discriminación y confiabilidad mediante los supuestos de la TCT. Durante la primera actividad se revisó la base de datos minuciosamente mediante un análisis descriptivo de los datos con el paquete estadístico IBM SPSS versión 21 para identificar todos aquellos casos con valores perdidos, incompletos o erróneos. Una vez identificados los casos con datos perdidos y erróneos se eliminaron de la base de datos para realizar los primeros análisis estadísticos descriptivos de la escala y sus ítems. En total se eliminaron 5 494 casos con datos perdidos, incompletos o erróneos quedando una base de datos de 10 367 casos.

Después de realizar los primeros datos descriptivos de la escala se procedió a lograr los supuestos de normalidad, linealidad y multicolinealidad de las puntuaciones requeridos para efectuar el AFE y el AFC según las recomendaciones de los especialistas en el campo (Bentler y Weeks, 1980; Bollen, 1989; Jöreskog y Sörbon, 1989). Para el análisis del supuesto de normalidad de las puntuaciones se consideraron las observaciones de distintos autores que señalan que el incumplimiento del mismo no trae consigo consecuencias estadísticamente significativas debido a que las distribuciones asimétricas de las estimaciones obtenidas no alteran drásticamente los valores del índice de ajuste Root Mean Square Error (RMSEA) (Muthén y Hofacker, 1988). Cabe señalar que el incumplimiento del supuesto de normalidad es un evento que frecuentemente ocurre en el campo de las ciencias sociales (Holgado Tello, Chacón Moscoso, Barbero García, y Vila Abad, 2010).

En especial, para que las estimaciones cumplieran los supuestos de normalidad se procedió a eliminar los valores atípicos y extremos con base en el cálculo de las puntuaciones típicas de cada variable considerando como potenciales a eliminarse los casos que presentaron puntajes  $Z$  fuera del rango  $\pm 3$  (Tabachnick y Fidell, 2001). Además del análisis mencionado, se identificaron puntos aislados en los extremos o valores atípicos con la notación  $-O < Q1 - 1.5(Q3-Q1)$  y  $+O > Q3 + 1.5(Q3-Q1)$ , y con la inspección visual de la distribución de las puntuaciones en el histograma y en los gráficos de caja y bigotes (box plots). En total, se eliminaron 824 casos con valores atípicos de la base de 10 367 casos quedando un total de 9 543 casos con puntuaciones mínimas de 24 y máximas de 44.

Una vez eliminados los casos atípicos, se procedió a verificar el supuesto de normalidad de las puntuaciones. Para dicho fin, se realizó la prueba de contrastes de bondad de ajuste con el estadístico Kolmogorov-Smirnov, la cual resulta demasiado sensible a pequeñas desviaciones de la normalidad cuando se trabaja con muestras de gran tamaño, como es el caso del presente estudio. Sin embargo, siguiendo las recomendaciones para el caso, se acompañó el estadístico de bondad de ajuste mencionado con la estimación de los índices de la asimetría y de la curtosis, considerando que los valores en el umbral  $\pm 1,5$  indican variaciones bajas de la normal y en consecuencia adecuados para realizar el AFE (George y Mallery, 2001). Como método alternativo se realizó el análisis visual del gráfico q-q plot, con el que se proporcionó información de la linealidad de la distribución normal y se permitió determinar si los datos recabados se ajustan razonablemente a una distribución normal.

Para el diagnóstico de multicolinealidad se analizaron correlaciones bajas y elevadas o redundantes entre los ítems de la escala. Primero se verificó que todas las intercorrelaciones entre los ítems fueran de moderadas a altas. Consecutivamente, se exploró la existencia de intercorrelaciones redundantes mayores a .90 debido a que aumenta la probabilidad de que

el AFE se debilite y se obtenga una solución factorial con poca estabilidad (Martínez Arias, 1999). Como un tercer análisis de multicolinealidad previo al AFE se aplicaron la prueba de esfericidad de Bartlett y la medida de adecuación muestral de Kaiser-Mayer-Olkin (KMO). Con la prueba de esfericidad de Bartlett se verificó la hipótesis nula de las correlaciones a un nivel de significancia de  $p < .05$ . (Everitt y Wykes, 2001). Dado que la prueba de esfericidad de Bartlett puede mostrar resultados significativos a pesar de no haber correlaciones altas, se aplicó como prueba adicional la medida KMO. Para ello se obtuvo el promedio de los términos de la diagonal de la matriz de correlación anti-imagen y se estableció en un rango de 0 a 1 como adecuado un valor igual o superior a .70, interpretándose como una interrelación alta entre los ítems (Hair, Anderson, Tatham, y Black, 1999).

Por su parte, en la última actividad de la primera etapa del estudio, se realizó el análisis de la confiabilidad y discriminación de la escala mediante los supuestos de la TCT. Para ello, con ayuda de los paquetes estadísticos IBM SPSS versión 21, Structural Equation Modeling Software (EQS 6.1) (Bentler, 2012) y *psych* en R Statistical X64 V 3.5.2 (Revelle, 2008) se obtuvieron el Alpha de Combrach y el Alpha Ordinal. Aunado a ello, se obtuvieron la tanto el estimado de cada como el promedio de la correlación ítem-total corregida y los Alfas de Cronbach y el Ordinal (RHO) si se elimina el ítem. En especial, es más apropiado el uso del Alpha de ordinal, dadas las recomendaciones de especialistas, debido a que para su estimación se utiliza de la matriz de correlaciones policórica (PCC) que es más adecuada que la matriz de Pearson para el análisis de variables ordinales (Richaud, 2005).

### 3.4.2. Etapa II. Análisis Factorial Exploratorio

Para la segunda etapa del estudio referente al AFE, se siguieron los procedimientos recomendados en especial por Bentler y Weeks (1980) y Bollen (1989). Esta segunda etapa se organizó en siete actividades: (1) selección de los métodos de extracción, (2) aplicación de los métodos de extracción, (3) definición teórica del número de factores a extraer, (4) definición empírica del número de factores a extraer, (5) selección del método de rotación de la matriz factorial, (6) aplicación del método de rotación seleccionado, e (7) interpretación y denominación teórica de los factores.

Después de haber verificado los supuestos y prerrequisitos para proceder a la aplicación del AFE se procedió a seleccionar y aplicar los métodos de extracción. Tomando en cuenta las características de los datos y las recomendaciones de Brown (2006), así como de Tabachnick y Fidell (2007) se seleccionaron y aplicaron los métodos de extracción de Componentes Principales (CP) y Máxima Verosimilitud (en inglés, Maximum Likelihood, ML) o también denominado Máxima Probabilidad (MP). Con la aplicación del método de extracción de ML se pudo contar con información para explicar una mayor cantidad de varianza posible de los datos observados fundamentada en la varianza total asociada a las variables, la varianza específica y la varianza de error de medición.

Por su parte, aunque el método de ML es menos utilizado en el AFE, se decidió aplicar siguiendo las recomendaciones de distintos autores que lo recomiendan para el análisis de datos que presentan una distribución normal multivariada (Byrne, 2001; Costello y Osborne, 2005; Khan, 2006). Aunque la escala de medida de los ítems no es continua sino ordinal, es comúnmente utilizado el método de ML para el ajuste de los modelos, siempre teniendo en cuenta la normalidad razonable de las distribuciones de las puntuaciones en los ítems de la

escala y la ausencia de multicolinealidad, así como el tamaño razonable de la muestra. Al respecto, autores como Hu, Bentler y Kano (1992) han encontrado que, en caso de que fuera razonable asumir la normalidad, el método ML funcionaba mejor cuando el tamaño de la muestra en estudio es mayor a 500 casos. En este punto, es importante señalar que una ventaja del presente estudio es que se cuenta con una muestra mayor a 9 000 casos, contando con mucho más de 10 participantes por ítem (Nunnally y Bernstein, 1995) por lo que se pueden esperar *a-priori* resultados útiles y relativamente estables en el AFE (Tabachnick y Fidell, 2001). Por otra parte, en el campo de las ciencias sociales, es frecuente el uso del método ML cuando se inicia con el análisis de la matriz de covarianzas (más que de la de correlación) y el número de categorías es el adecuado. En este sentido, Bentler y Chou (1987) indican que no se debe temer el uso de métodos continuos con variables que tienen cuatro o más categorías.

Para la definición teórica y empírica de los factores a extraer de la ERECVE se decidió tomar como base las categorías del modelo teórico de violencia escolar vista como proceso grupal propuesto por Salmivalli (1999, 2010), y contrastar con distintas técnicas empíricas de extracción de factores. Retomando la **Tabla 3** se puede observar que los ítems de la ERECVE se organizan en cinco categorías o denominaciones *a-priori* de los factores: (1) rol de víctima, (2) rol del agresor solitario, (3) rol del agresor social, (4) rol de observador y (5) rol de pacificador o defensor. En este punto es importante señalar que el ERECVE no aborda todas las categorías propuestas por los autores de referencia. Es importante señalar que en la estrategia integral de evaluación aplicada por Investigadores de la UEE en colaboración con autoridades educativas de la SEBC en 2016 se procuró explorar la mayor cantidad de variables que ayudarán a explicar el clima y la violencia escolar en educación primaria, secundaria y media superior por lo que se requirió economizar en la extensión del cuestionario

sin repercutir con la esencia de los constructos. Es por ello que con base en dichos argumentos se decidió no incorporar los ítems relacionados con los roles de asistente de agresor y refuerzo de agresor.

Ahora bien, es importante señalar que la toma de decisiones para definir la extracción correcta del número de factores es una de las principales problemáticas del análisis factorial señalada desde la década de los años 60 del siglo pasado autores como Cattell (1966). El utilizar un único criterio para definir la extracción correcta del número de factores o no contar con un modelo teórico *a-priori* que guíe dicha decisión puede genera problemas de sobrestimación o subestimación del número real de factores. En el caso del presente estudio se tiene la ventaja de contar con una teoría consolidada con evidencias teóricas y empíricas que guía la definición de factores que subyacen a la ERECVE. Sin embargo, se consideró relevante utilizar un conjunto de criterios empíricos con los cuales contrastar el número de factores definidos con el modelo teórico *a-priori*, además de considerar que la sobre-extracción es menos riesgosa puesto que conduce a menos error en la medición (Reise, Waller y Comrey, 2000). De forma especial, y con base en las recomendaciones mencionadas, se aplicaron los métodos de extracción basadas en el criterio de Kaiser, el porcentaje de varianza explicada y el gráfico de sedimentación (scree test).

Por su parte, para las tres últimas actividades de la etapa II, encaminadas a la rotación e interpretación de la matriz factorial, se decidió aplicar el método de rotación con filtro propuesto por Tabachnick y Fidell (2001), así como contrastar el patrón de correlaciones bajas y altas de cada variable con los distintos factores en la matriz rotada definidos en el modelo teórico de referencia. Es importante señalar que debido a que las soluciones de los métodos de rotación oblicua son más adecuadas para las variables psicológicas, pues por lo general, los estudios arrojan intercorrelaciones altas, se decidió seguir las recomendaciones referidas

a la conveniencia de aplicar el método de rotación oblicua Promax (Gorsuch, 1983) como filtro inicial y como método de contraste la rotación ortogonal Varimax (Kaiser, 1958). Para las actividades finales del AFE relacionadas con la interpretación y denominación de los factores se examinó en la matriz rotada el patrón de correlaciones de cada ítem y se contrastó con la solución de factores y el modelo teórico antes mencionado.



### 3.4.3. Etapa III. Análisis Factorial Confirmatorio

Durante la etapa III del presente estudio, siguiendo las recomendaciones de Byrne (2001), se complementó el AFE con la aplicación del AFC para contar con una aproximación más fuerte de la definición y validación del constructo de interés. Dado que se posee información del modelo teórico subyacente de la ERECVE y los resultados del AFE aplicado llegaron a una solución congruente y satisfactoria se consideró apropiado aplicar el AFC. Para ello, con apoyo del Structural Equation Modeling Software (EQS) 6.1 (Bentler, 2012) se formalizó en términos operativos el modelo factorial confirmatorio, se evaluó su ajuste a los datos recabados, y se analizó la significación estadística de las relaciones entre las variables observables y latentes (Batista Foguet y Gallart, 2000).

Fueron seis las actividades realizadas para el AFC tomando en consideración las recomendaciones de Bentler y Weeks (1980), y Jöreskog y Sörbon (1989): (1) Definición gráfica del modelo factorial confirmatorio, (2) formalización matemática e identificación del modelo factorial confirmatorio, (3) Estimación del modelo factorial confirmatorio definido, (4) análisis del ajuste del modelo factorial confirmatorio estimado, e (5) interpretación y reespecificación, de ser pertinente, del modelo factorial confirmatorio identificado.

Para la definición gráfica del modelo factorial confirmatorio, con base en los resultados satisfactorios del AFE, se decidió mantener la postura inicial de la organización de la estructura de la ERECVE según las categorías del modelo teórico de violencia escolar vista como proceso grupal propuesto por Salmivalli (1999, 2010). En la *Figura 1* puede observarse la definición gráfica del modelo factorial confirmatorio para su estimación.

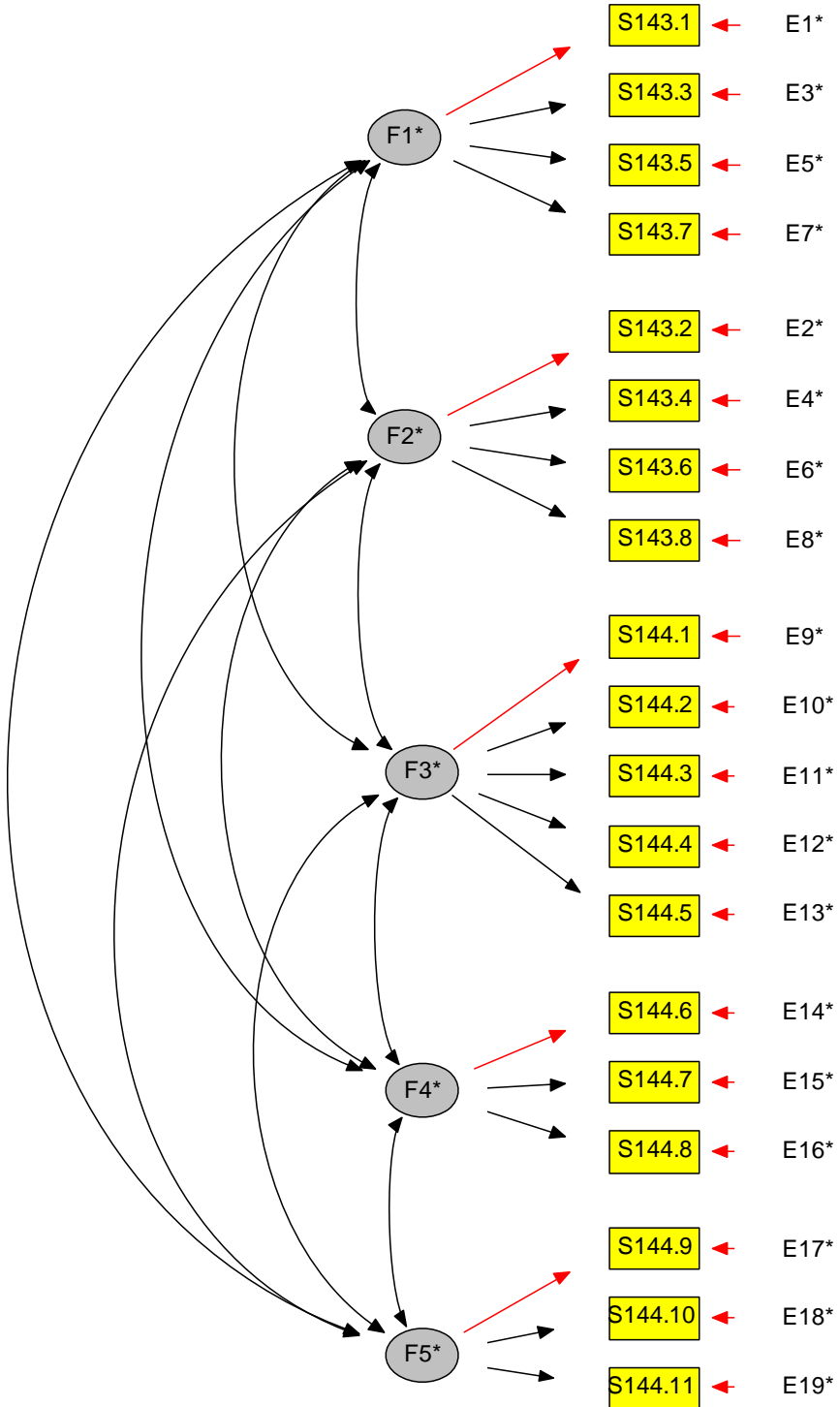


Figura 1. Definición gráfica del modelo factorial confirmatorio para el análisis de la estructura interna de la ERECVE

Después de la definición gráfica y del modelo factorial confirmatorio se procedió a su formalización matemática e identificación. En la *Figura 2* puede observarse el código producido por el EQS para la formalización matemática del modelo factorial confirmatorio estimado. Por su parte, para la identificación del modelo factorial confirmatorio se revisaron las recomendaciones de distintos autores (por ej. Long, 1983; Jöreskog y Sörbom, 1989; Hatcher, 1994; Ullman, 1996). Sin embargo, finalmente se tomaron en cuenta las recomendaciones de Hatcher (1994) y Ullman (1996) que coinciden con la notación del EQS: (1) comparación del número de datos con el número de parámetros a estimar, (2) establecimiento de una escala para los factores comunes, (3) aseguramiento de la identificabilidad de la(s) sección(es) del modelo que contiene la relación entre las variables observadas y los factores, y (4) fijar arbitrariamente el coeficiente de regresión del término de error al valor 1<sup>2</sup>.

$$\begin{aligned}V1 &= 1F1 + E1; \\V2 &= 1F2 + E2; \\V3 &= *F1 + E3; \\V4 &= *F2 + E4; \\V5 &= *F1 + E5; \\V6 &= *F2 + E6; \\V7 &= *F1 + E7; \\V8 &= *F2 + E8; \\V9 &= 1F3 + E9; \\V10 &= *F3 + E10; \\V11 &= *F3 + E11; \\V12 &= *F3 + E12; \\V13 &= *F3 + E13; \\V14 &= 1F4 + E14; \\V15 &= *F4 + E15; \\V16 &= *F4 + E16; \\V17 &= 1F5 + E17; \\V18 &= *F5 + E18; \\V19 &= *F5 + E19;\end{aligned}$$

*Figura 2.* Código producido por el EQS para la formalización matemática del modelo factorial confirmatorio para el análisis de la estructura interna de la ERECVE

Para la identificación del modelo, se puede decir que se cumplió fácilmente con la restricción  $qs+[s(s+1)/2]$  debido a que el modelo se sobre-identifica al contar con más datos que parámetros. Con ello, se hizo posible, al utilizar los grados de libertad, someter a prueba el modelo factorial. Aunado a ello, para el requisito de identificación del modelo se fijó el coeficiente de regresión (carga factorial) de una de las variables observadas que cargan sobre cada factor a 1. En la *Figura 2*, referida al código producido por el EQS para la formalización matemática del modelo factorial confirmatorio se puede observar que las variables observadas (ítems) 1 (S143.1), 2 (S143.2), 9 (S144.1), 14 (S144.6) y 17 (S144.9) se fijaron a los factores 1, 2, 3, 4 y 5 respectivamente. También se procuró que cada factor tuviera asociados al menos tres variables con cargas no nulas y que los errores asociados con dichos indicadores no estuvieran correlacionados entre sí, cargando cada variable sobre un solo factor permitiendo la covariación entre ellos (ver *Figura 1*). Para el cumplimiento del último de los requisitos de identificación recomendados por Hatcher (1994) y Ullman (1996) se fijó arbitrariamente en el EQS el coeficiente de regresión del término de error al valor 1<sup>2</sup>.

Siguiendo las recomendaciones de Bentler y Weeks (1980), y Jöreskog y Sörbon (1989), después de la identificación del modelo se procedió a su estimación. En especial, para la estimación del modelo factorial se siguió la propuesta de dos pasos de Sharma (1996): (1) estimar los parámetros del modelo factorial hipotetizado, y (2) determinar su ajuste. Dado que los 19 ítems corresponden a una escala de nivel de medición ordinal con 5 opciones de respuesta, y que aunado a ello la escala en su conjunto, después de los procedimientos previos para lograr los supuestos para el AFE, solo alcanzó una normalidad multivariable razonable, se decidió aplicar métodos robustos de estimación para el AFC. En especial se aplicó el ML robusto como método de estimación y la chi-cuadrado robusta de SatorraBentler (Satorra y Bentler, 1994).

Por su parte, siguiendo los criterios que señalan Hu y Bentler, 1995 para determinar el ajuste del modelo factorial se utilizaron múltiples índices: estadístico  $\chi^2(df)$ , por sus siglas en inglés), Santorra-Bentler (S-B)  $\chi^2(df)$ , Criterio de Información de Akaike (AIC, por sus siglas en inglés) Índice de Ajuste No Normativo (NNFI, por sus siglas en inglés), Índice de Ajuste Comparativo (CFI, por sus siglas en inglés), Error Cuadrático Medio de Aproximación (RMSEA, por sus siglas en inglés), Raíz Cuadrada Media Residual (RMR), Raíz Cuadrada Media Residual Estandarizada (SRMR), Índice de Bondad de Ajuste Ajustado (AGFI). Los criterios para evaluar el ajuste con los índices relativos fueron: NNFI  $p \geq 0.95$ , CFI  $p \geq 0.95$  y RMSEA  $p < 0.06$ ; y para los absolutos: RMR  $p =$ , AGFI  $p =$ , (SRMR)  $p < 0.08$ ; y la S-B  $\chi^2(df) < 3$  (Ruíz, Pardo y SanMartín, 2010). El IAC permite comparar los modelos por lo que se prefieren valores lo más bajo posible.

Terminado el proceso de estimación se procedió a la interpretación de la solución estandarizada en contraste con el modelo factorial identificado. En términos generales tanto el ajuste como la solución estandarizada fue satisfactoria. Sin embargo, se decidió realizar una ligera reespecificación del modelo sin afectar la congruencia de la estructura de la ERECVE con las categorías del modelo teórico de violencia escolar vista como proceso grupal de Salmivalli (1999, 2010). Para la nueva estimación se excluyeron los ítems S144.4 y S144.5 pertenecientes a la Sub-escala del Rol de Víctima en el Ciclo de la Violencia Escolar (SRVCVE). Con la reespecificación del modelo factorial realizada mejoraron los índices de ajustes llegando finalmente a una solución satisfactoria. Con todo lo anterior, se asumió que el conjunto de 19 ítems representa adecuadamente las 5 dimensiones latentes referidas a las sub-escalas de la ERECVE procediendo al análisis de sus correlaciones con otras variables.

#### **3.4.4. Etapa IV. Análisis de correlación convergente y discriminante**

Para obtener evidencias de validez de constructo de relación con otras variables tanto convergentes como discriminantes se analizaron las correlacionales entre los factores de la ERECVE y variables exógenas asociadas a los temas del clima y la violencia escolar, así como de la violencia intrafamiliar. En especial las variables exógenas fueron la Subescala de Violencia Intrafamiliar Negativa (SEVIN), la Subescala de Violencia Intrafamiliar Positiva (SEVIP), la Subescala del Rol del Docente en el Ciclo de la Violencia Escolar (SRDCVE), la Subescala del Rol del Director en el Clima Escolar (SRDCE) y la Subescala de Ambiente Escolar Violento (SAEV). Como criterio para determinar colinealidad o redundancia entre factores fueron correlaciones muy altas ( $\Rightarrow 0.85$ ) y correlaciones muy bajas o negativas para determinar validez discriminante.

---

## IV. RESULTADOS

En este apartado se presentan los resultados del estudio organizados según las distintas etapas del procedimiento: (I) Análisis preliminares, (II) Análisis Factorial Exploratorio (AFE), (III) Análisis Factorial Confirmatorio (AFC) y (IV) Análisis de correlación concurrente y discriminante. En lo que respecta a los análisis preliminares, se describe en el primer subapartado del presente capítulo los resultados del análisis descriptivo de la ERECVE, por sub-escalas y por cada ítem, así como los resultados del análisis de normalidad, linealidad, multicolinealidad, y de discriminación y confiabilidad de la escala. Como segundo subapartado del capítulo se muestran los resultados del AFE desagregados en el porcentaje de varianza explicada del modelo, los resultados de la aplicación de los métodos Varimax y Promax, y la interpretación resultante del modelo estimado.

Para el tercer subapartado se presentan los resultados del AFC referentes a los índices de ajuste del modelo factorial confirmatorio y a su interpretación y reespecificación. En el último subapartado se describen los resultados del análisis correlacional convergente y discriminante entre los factores de la ERECVE y las variables exógenas asociadas a los temas del clima y violencia escolar.

### 4.1. Resultados del análisis preliminar

#### 4.1.1. Resultados del análisis descriptivo

Como ya se comentó en el apartado referido al procedimiento del presente estudio, después de identificar y eliminar todos aquellos casos con valores perdidos e incompletos o erróneos (5 494) se realizaron con los datos restantes (N=10 367) los análisis estadísticos descriptivos

de la escala y sus ítems. En la **Tabla 4** se describen las medidas de tendencia central, el error típico de la media, la desviación estándar, la varianza, la asimetría y la curtosis de la ERECVE y de cada ítem que la compone. Nótese que en la mayoría de los ítems se obtienen valores cercanos al mínimo.<sup>1</sup>

**Tabla 4.** Descriptivos de las 19 preguntas de la Escala de los Roles de los Estudiantes en el Ciclo de la Violencia Escolar (ERECVE)

Pregunta	Media	Error típ.	Mediana	Moda	Desv. típ.	Varianza	Asimetría	Curtosis
S143.1	1.26	0.008	1	1	0.77	0.60	3.55	12.77
S143.2	2.17	0.012	2	1	1.23	1.52	-0.92	-0.19
S143.3	1.20	0.007	1	1	0.70	0.49	4.12	17.34
S143.4	2.29	0.013	2	1	1.36	1.86	-0.81	-0.64
S143.5	1.14	0.006	1	1	0.57	0.32	4.96	26.57
S143.6	2.01	0.011	2	1	1.14	1.30	-1.16	0.59
S143.7	1.16	0.006	1	1	0.59	0.34	4.63	23.46
S143.8	2.76	0.014	2	2	1.38	1.90	-0.40	-1.14
S144.1	1.29	0.007	1	1	0.72	0.52	3.28	11.92
S144.2	1.16	0.005	1	1	0.54	0.29	4.58	24.69
S144.3	1.08	0.004	1	1	0.43	0.18	6.75	50.72
S144.4	1.06	0.004	1	1	0.36	0.13	8.14	73.72
S144.5	1.10	0.004	1	1	0.46	0.21	6.00	41.32
S144.6	1.13	0.005	1	1	0.50	0.25	5.20	31.37
S144.7	1.10	0.004	1	1	0.45	0.20	6.17	43.63
S144.8	1.12	0.005	1	1	0.50	0.25	5.55	34.63
S144.9	1.89	0.012	1	1	1.21	1.46	1.46	1.13
S144.10	1.85	0.012	1	1	1.18	1.39	1.51	1.36
S145.11	1.73	0.011	1	1	1.12	1.26	1.72	2.12
Puntaje total ERECVE*	34.04	0.06	34	35	6.08	36.92	1.73	7.78

Error típico de la asimetría= 0.24

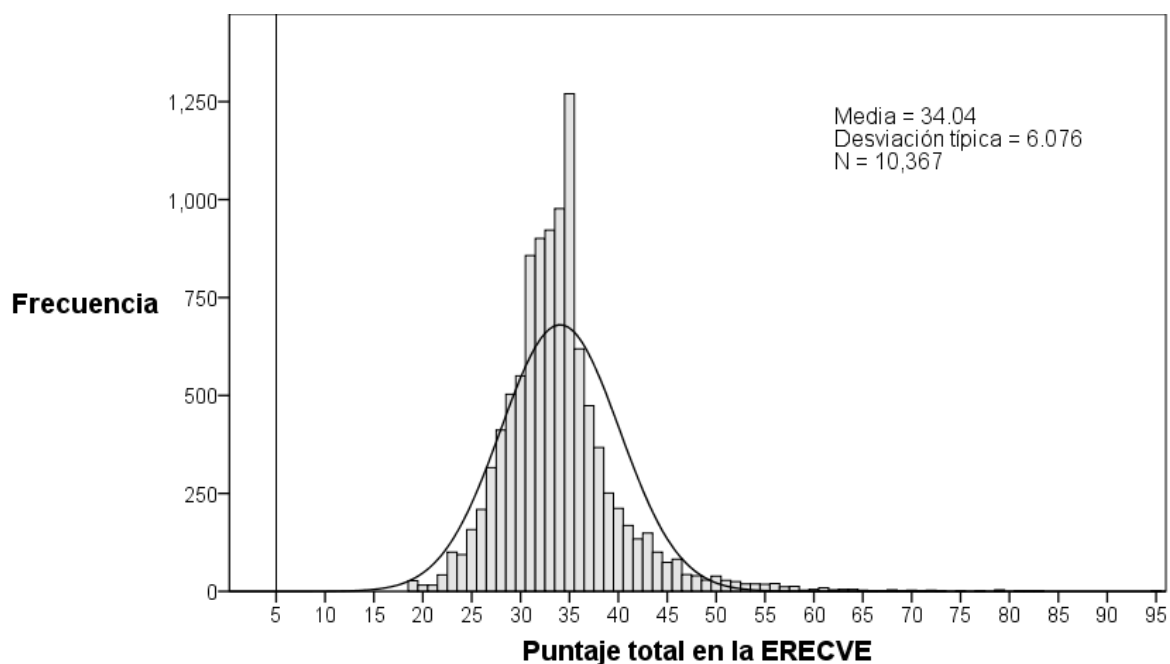
Error típico de la Curtosis= 0.048

\*El puntaje total de la ERECVE es resultado de la suma de la puntuación obtenida por los participantes en cada ítem tomando en cuenta la recodificación según la clave de respuestas.

<sup>1</sup> Los ítems presentan una escala de 5 categorías donde el valor mínimo es 1= Nunca y 5= Diariamente. Por su parte el valor mínimo del puntaje total de la escala es de 19 y el máximo es de 95.



En la *Figura 3* se muestra la distribución del puntaje total de la ERECVE. Puede observarse a simple vista que dicha distribución es asimétrica dado que las puntuaciones se encuentran sesgadas a la cola derecha. Con ello, los supuestos de normalidad previos al AFE no se cumplen, lo que hace necesario la aplicación de procedimientos para identificar y eliminar casos atípicos.



*Figura 3.* Distribución del puntaje total de la Escala de los Roles de los Estudiantes en el Ciclo de la Violencia Escolar (ERECVE)

Por su parte, en la *Figura 4* se pueden observar las frecuencias de cada opción de respuesta en cinco preguntas más representativas (S144.1, S144.6, S143.1, S144.9 y S143.8) de los cinco roles de las subescalas del ERECVE. Nótese que en todos los casos se presenta con mayor frecuencia casos de alumnos que señalan que *Nunca* participan o aportan en un rol del ciclo de la violencia escolar en específico. Sin embargo, en los casos del rol del observador y pacificador aumenta la frecuencia de alumnos que señalan que participan en el

ciclo de la violencia escolar en dichos roles. Los descriptivos de todos los ítems de la escala se pueden revisar en el **Apéndice 1**.

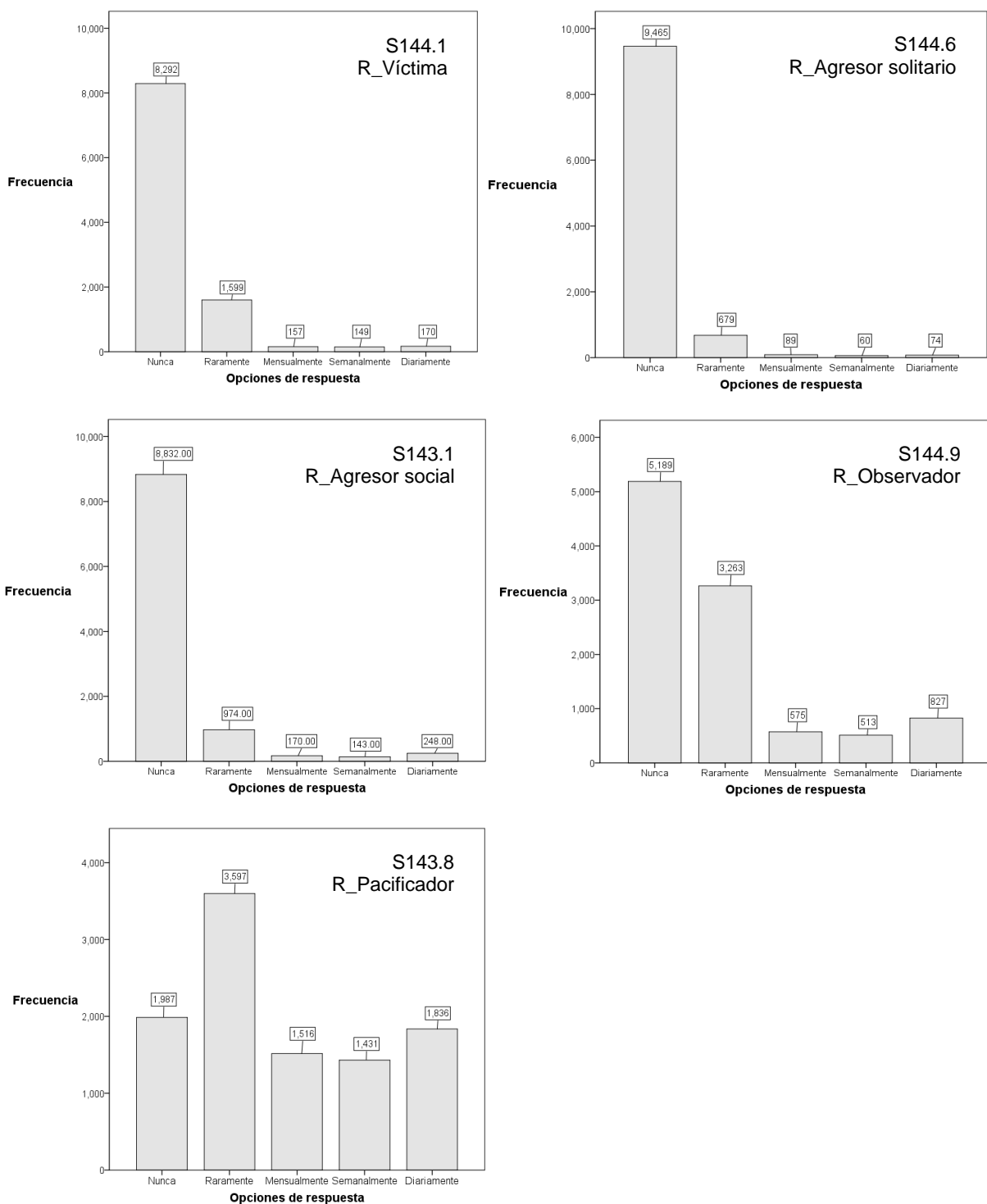
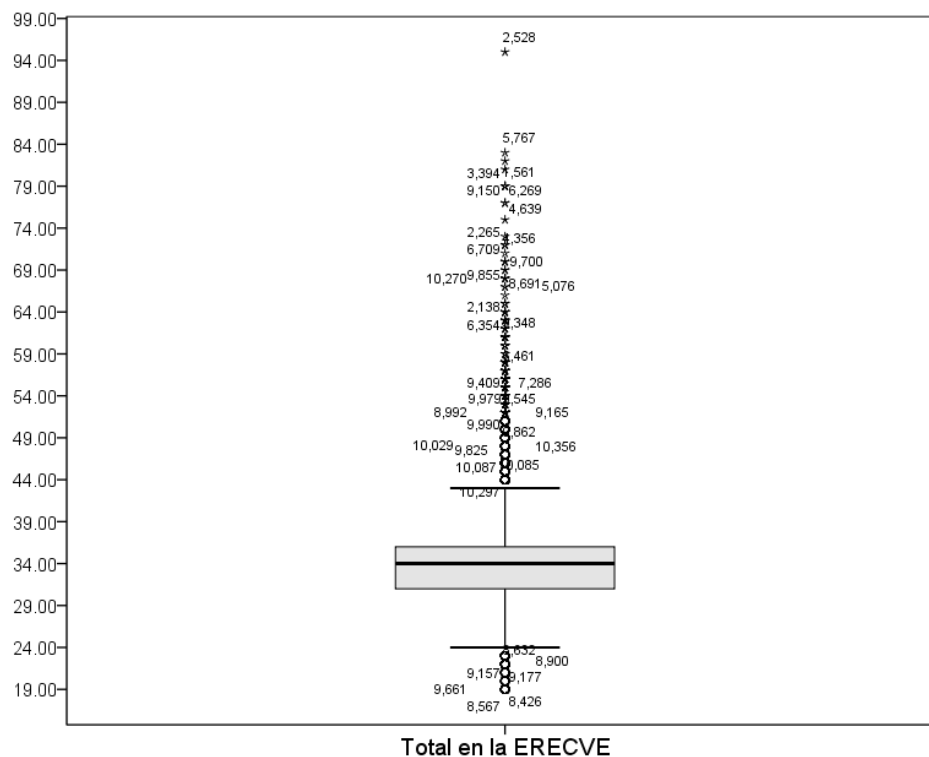


Figura 4. Frecuencias de cada opción de respuesta de las preguntas S144.1, S144.6, S143.1, S144.9 y S143.8

#### 4.1.2. Resultados del análisis de normalidad, linealidad y multicolinealidad

Para que las estimaciones cumplieran los supuestos de normalidad primero se identificaron las puntuaciones atípicas que presentaron puntajes Z fuera del rango  $\pm 3$  (Tabachnick y Fidell, 2001) y resultantes de la notación  $-O < Q1 - 1.5(Q3-Q1)$  y  $+O > Q3 + 1.5(Q3-Q1)$ . En ambos procedimientos se señalaron eliminar puntuaciones menores a 24 y mayores a 44. La identificación de valores atípicos con puntuaciones Z y notación  $Q3-Q1$  se acompañó con la inspección visual del gráfico de caja y bigotes (box plots). Como ya se mencionó, en la *Figura 3* puede observarse a simple vista que dicha distribución es asimétrica dado que las puntuaciones se encuentran sesgadas a la cola derecha. Por su parte, en la *Figura 5* pueden observarse a simple vista una serie de casos extremos menores a 24 y mayores a 44 que corrobora y amplía el sesgo positivo identificado en el histograma.



*Figura 5.* Identificación de casos extremos menores a 24 puntos y mayores a 44

Consecutivamente a la identificación de los casos atípicos se procedió a la eliminación de 824 casos de la base de 10 367 quedando un total de 9 543 casos. Ahora bien, con la eliminación de los casos atípicos se procedió a la verificación del supuesto de normalidad de las puntuaciones. Los resultados de la prueba de contrastes de bondad de ajuste del estadístico Kolmogorov-Smirnov<sup>2</sup> presentó un estadístico de 0.81 significativo al 0.000 con 9542 *gl*. Con ello se acepta la hipótesis nula de normalidad con lo que no se cumple con el supuesto previo para el AFE. Es común en la literatura que la prueba de contrastes de bondad de ajuste del estadístico Kolmogorov-Smirnov resulte demasiado sensible a pequeñas desviaciones de la normalidad cuando se trabaja con muestras de gran tamaño. A modo de contraste se analizó la estimación de los índices de asimetría y curtosis. Aunque los resultados arrojaron valores  $z(G)$  cercanos quedaron fuera del umbral  $\pm 1,5$  rechazando las hipótesis de distribución mesocúrtica y simétrica (ver **Tabla 5**). Sin embargo, dado que los valores son cercanos al umbral representando variaciones leves de la normal puede aceptarse con cautela una normalidad razonable para realizar el AFE (George y Mallery, 2001).

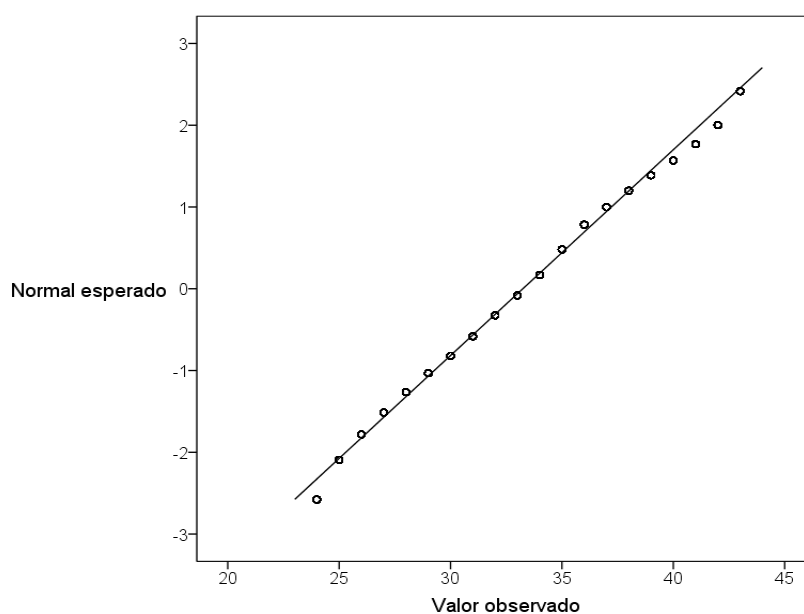
La decisión tomada con el análisis de la curtosis y la simetría de las puntuaciones se corroboró con el análisis visual del gráfico q-q plot (ver *Figura 6*). Nótese que los valores observados se encuentran visualmente alineados con la normal esperada. Con dicha información se puede aceptar con más seguridad el supuesto de normalidad y linealidad de las puntuaciones para efectuar el AFE.

---

<sup>2</sup> Se aplica con la corrección de significación de Lilliefors

**Tabla 5.** Descriptivos de la base de datos sin casos atípicos

Descriptivos	Estadístico	Error típ.
Media	33.24	.04068
Intervalo de confianza para la media al 95%	Límite inferior	33.16
	Límite superior	33.32
Media recortada al 5%	33.21	
Mediana	33.00	
Varianza	15.80	
Desv. típ.	3.97	
Mínimo	24.00	
Máximo	43.00	
Rango	19.00	
Amplitud intercuartil	4.00	
Asimetría	.096	.025
Z(G1)	3.83	
Curtosis	-.133	.050
Z(G2)	-2.64	
$\sum Z^2$	21.68	
Estimador-M de Huber <sup>a</sup>	33.16	
Biponderado de Tukey <sup>b</sup>	33.10	
Estimador-M de Hampel <sup>c</sup>	33.14	
Onda de Andrews <sup>d</sup>	33.10	

*Figura 6.* Gráfico Q-Q de linealidad normal de las puntuaciones de la ERECVE

Para el diagnóstico de multicolinealidad se presentan los resultados de correlaciones bajas y elevadas o redundantes entre los ítems de la escala, de la prueba de esfericidad de Bartlett y de la medida de adecuación muestral de Kaiser-Mayer-Olkin (KMO). Aplicando la prueba de correlación de Spearman se encontraron correlaciones entre  $-.009$  y  $.718$ . Sin embargo, alrededor del 30% de las correlaciones entre los ítems son valores iguales o menores a 0.1 (ver **Apéndice 2**). Siguiendo los resultados de las pruebas para determinar la factibilidad del AFE se calculó el índice de adecuación muestral Kaiser-Meyer-Olkin ( $KMO = 0.745$ ) y la prueba de esfericidad de Bartlett ( $\chi^2 = 34660.806$ ;  $gl = 171$ ;  $p < 0.000$ ). Tomando en cuenta el criterio de que entre más cerca de 1 sea el valor obtenido del test KMO, implica que la relación entre las variables es alta, se consideró que las variables se encuentran lo suficientemente interrelacionadas para proceder con el AFE. Por su parte, los resultados de la prueba de esfericidad de Bartlett fueron convergentes con los de la prueba KMO resultando significativa, por lo que se puede aceptar la hipótesis nula que sugiere la posibilidad de aplicar el AFE.

#### **4.1.3. Análisis de confiabilidad y discriminación de la ERECVE**

La confiabilidad y discriminación promedio de la ERECVE fueron satisfactorias y adecuadas para realizar el AFE. En cuanto a los resultados de confiabilidad, se obtuvo un Alpha de confiabilidad de Cronbach de 0.71, un coeficiente de confiabilidad RHO de 0.79 y un Omega de 0.77. En la **Tabla 6** se muestran los coeficientes de confiabilidad del Alpha de Cronbach si se elimina el ítem, y la correlación ítem-total corregida de cada ítem y del promedio de la ERECVE. Nótese que la media de la correlación ítem-total cumple el criterio de  $\Rightarrow 0.2$ . Solo los ítems S144.3, S144.4 y S1448 no cumplieron con el criterio establecido.

**Tabla 6.** Confiabilidad del Alpha de Combrach si se elimina el ítem y correlación ítem-total corregida de la ERECVE

Ítems	Media de la escala si se elimina el ítem	Varianza de la escala si se elimina el ítem	Correlación ítem-total corregida	Correlación múltiple al cuadrado	Alfa de Cronbach si se elimina el ítem
S143.1	26.16	32.829	.234	.204	.699
S143.2	25.24	29.060	.358	.171	.688
S143.3	26.22	33.347	.203	.253	.701
S143.4	25.12	28.307	.354	.181	.691
S143.5	26.26	33.756	.202	.279	.702
S143.6	25.40	28.725	.439	.249	.675
S143.7	26.25	33.623	.233	.248	.700
S143.8	24.63	27.745	.383	.211	.686
S144.1	26.14	33.051	.249	.230	.698
S144.2	26.26	33.881	.238	.215	.701
S144.3	26.33	34.647	.127*	.171	.706
S144.4	26.35	34.786	.126*	.156	.707
S144.5	26.31	34.314	.211	.124	.703
S144.6	26.30	34.350	.207	.275	.703
S144.7	26.33	34.479	.216	.294	.704
S144.8	26.31	34.309	.198*	.193	.703
S144.9	25.58	28.695	.445	.540	.674
S144.10	25.63	28.666	.470	.604	.671
S144.11	25.73	29.263	.443	.421	.675
<b>Promedia de correlación ítem-total</b>			<b>.281</b>		

\*Ítems que no cumplen el criterio de discriminación => 0.2.

## 4.2. Resultados del Análisis Factorial Exploratorio (AFE)

En este apartado se presentan primero los resultados de la aplicación de los métodos de extracción de Componentes Principales (CP) y Máxima Verosimilitud (en inglés, Maximum Likelihood, ML) empleados para obtener una definición empírica del número de factores con la cual contrastar la definición teórica de origen subyacente a la ERECVE. Consecutivamente se presentan los resultados de la aplicación de los métodos de rotación Varimax y Promax, y los resultados de la interpretación y denominación teórica de los factores.

En la **Tabla 7** se presenta el porcentaje de varianza explicada resultado de la aplicación del método de extracción de CP. Nótese que los cinco factores estimados obtuvieron autovalores superiores a la unidad y una explicación del 52.12% de la varianza. Cabe mencionar que las estimaciones del porcentaje de varianza explicada por el método de extracción de CP concuerda con exactitud con la estimación del método ML (52.12%).

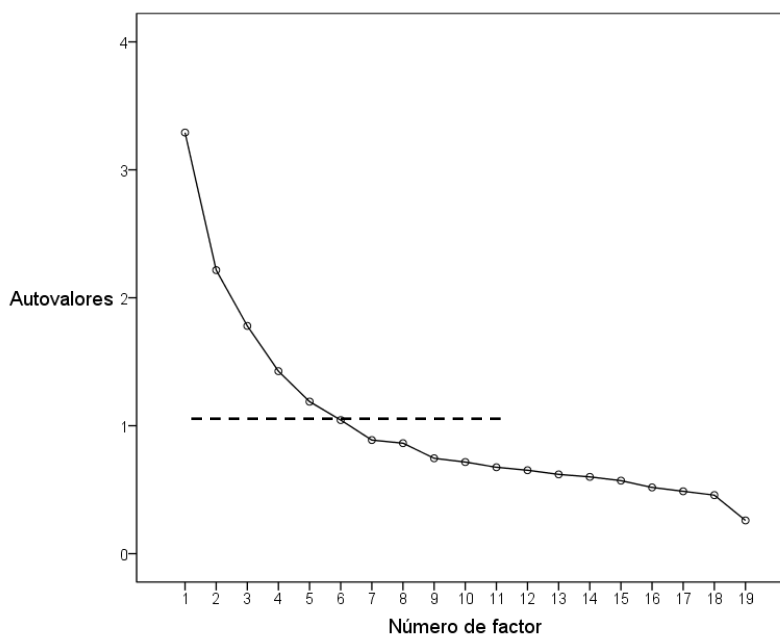
Además de los resultados obtenidos con los métodos de extracción de CP y ML se analizó la gráfica de sedimentación (ver *Figura 7*). Con base en la regla de Kaiser y el Scree test (Cattell, 1966) se analizó el gráfico de sedimentación corroborando los resultados de los métodos CP y ML. Nótese que si aplicamos exclusivamente la prueba de Scree test o de caída se podría explicar hasta seis factores teniendo una postura flexible. Sin embargo, siempre es más aconsejable seguir la teoría subyacente del instrumento en caso de existir.



**Tabla 7.** Varianza explicada resultado de la aplicación del método de extracción de componentes principales

Componente	Autovalores iniciales			Suma de las saturaciones al cuadrado de la rotación <sup>a</sup>
	Total	% de la varianza	% acumulado	Total
1	3.291	17.322	17.322	2.577
2	2.216	11.665	28.987	2.213
3	1.780	9.368	38.355	2.278
4	1.426	7.507	45.861	2.059
5	1.188	6.253	52.115	2.060
6	1.044	5.497	57.612	
7	.888	4.673	62.285	
8	.864	4.546	66.831	
9	.746	3.924	70.755	
10	.716	3.770	74.525	
11	.675	3.554	78.078	
12	.652	3.432	81.510	
13	.620	3.262	84.771	
14	.600	3.158	87.929	
15	.571	3.005	90.934	
16	.518	2.725	93.659	
17	.488	2.566	96.226	
18	.457	2.407	98.633	
19	.260	1.367	100.000	

<sup>a</sup> Cuando los componentes están correlacionados, las sumas de los cuadrados de las saturaciones no se pueden añadir para obtener una varianza total.

**Figura 7.** Gráfica de sedimentación

Los resultados de la aplicación de los métodos de rotación fueron de igual forma satisfactorios. Si tomamos de referencia la matriz de estructura obtenida con el método de rotación de CP se puede decir que todos los ítems cumplen con el criterio de saturación  $\Rightarrow 0.45$ , pero si tomamos de referencia la matriz de configuración obtenida con el método de rotación de CP y la matriz de factores rotados obtenida con el método de rotación de ML encontramos que el ítem S144.5 no cumple del todo con el criterio establecido (ver **Tabla 8**). Dado que en su gran mayoría fueron satisfactorios los resultados del AFE se decidió proseguir con el AFC. Por su parte, las correlaciones entre los factores fueron bajas en general. La correlación más elevada se dio entre los factores 1 y 3 ( $r = 0.301$ ), y las más bajas entre los factores 1 – 2 ( $r = -0.043$ ), y 2 – 4 ( $r = 0.053$ ).

**Tabla 8.** Matriz factorial de la ERECVE

Ítems	Matriz de configuración					Matriz de estructura				
	Componente					Componente				
	1	2	3	4	5	1	2	3	4	5
S143.1		.611					.640			
S143.2			.610					.621		
S143.3		.729					.730			
S143.4			.724					.694		
S143.5		.775					.759			
S143.6			.691					.713		
S143.7		.694					.704			
S143.8			.673					.677		
S144.1				.561					.596	
S144.2				.637					.659	
S144.3				.739					.693	
S144.4*				.611					.583	
S144.5*				.427					.484	
S144.6					.764					.768
S144.7					.824					.811
S144.8					.680					.688
S144.9	.865					.865				
S144.10	.916					.905				
S144.11	.804					.805				

\*Ítems en donde el método de rotación de ML no concordó con el de CP.

### 4.3. Resultados del Análisis Factorial Confirmatorio (AFC)

Los resultados del AFC se organizan en dos grandes subapartados referidos a la evaluación del ajuste del modelo factorial a los datos y el análisis de la significación estadística de las relaciones entre las variables observables y latentes. Por su parte, en los resultados de la evaluación del ajuste se encontró que el modelo se ajusta excelente a los datos considerando el índice de bondad de ajuste RMSEA ( $0.029 < 0.06$ ). Sin embargo, los índices NFI, NNFI, CFI, MFI, así como la  $S-B\chi^2/(gl)$  no alcanzaron valores aceptables de ajuste por muy poco (ver **Tabla 9**).

**Tabla 9.** Índices de bondad de Ajuste del modelo factorial confirmatorio estimado

Índices	Criterio	Método de ML	Método Robusto
$\chi^2(gl)$		34686.858 (171)	11025.968 (171)
$\chi^2/(gl), p ; S-B \chi^2/(gl), p$	Cociente < 5	3238.328/(142)=22.805, p=.00000	1278.7303/(142)=9.005, p=.00000
AIC	Menor estimación	2954.328	994.730
NFI	p=>.95	.907	.884
NNFI	p=>.95	.892	.874
CFI	p=>.95	.910	.895
MFI	p=>.95	.850	.942
GFI	p=>.95	.965	
AGFI	p=>.95	.953	
RMR	p< .08	.024	
SRMR	p< .08	.038	
<b>RMSEA</b>	p< .06	.048	<b>.029</b>

Tomando en cuenta dichos resultados se decidió realizar una ligera reespecificación del modelo sin afectar la estructura dimensional de la ERECVE. Para la nueva estimación se excluyeron los ítems S144.4 y S144.5 pertenecientes a la SRVCVE. Con la reespecificación del modelo factorial mejoraron la mayoría de los índices de bondad de ajuste llegando a una solución satisfactoria (AIC=803.191, NFI=0.918, NNFI=0.907, CFI=0.926,

IFI=0.926, MFI=0.953, RMSEA=0.03). En la **Tabla 10** se pueden observar las estimaciones del modelo factorial tanto exploratorio como confirmatorio reespecificados.

**Tabla 10.** Estimaciones del modelo factorial tanto exploratorio como confirmatorio reespecificados

Ítems	Matriz de estructura <sup>a</sup> Componentes					R <sup>2</sup>
	1	2	3	4	5	
S143.1		.644				.266
S143.2			.625			.206
S143.3		.733				.369
S143.4			.701			.249
S143.5		.763				.420
S143.6			.714			.405
S143.7		.709				.351
S143.8			.679			.314
S144.1					.754	.397
S144.2					.773	.399
S144.3					.608	.100
S144.6				.779		.426
S144.7				.817		.527
S144.8				.693		.237
S144.9	.863					.608
S144.10	.907					.828
S144.11	.809					.469

<sup>a</sup>Método de extracción: Análisis de componentes principales.  
Método de rotación: Normalización Promax con Kaiser.

#### 4.4. Resultado del análisis de correlación concurrente y discriminante

En la **Tabla 11** se describen las correlaciones entre las subescalas y la escala en su conjunto, así como con otras variables relacionadas con los temas del clima y la violencia escolar e intrafamiliar. Nótese en especial que las subescalas se correlacionan moderadamente entre sí. También nótese que las subescalas con las que se evalúa la aportación de los roles al ciclo de la violencia escolar se asocian negativamente con la subescala del rol del pacificador o defensor. Dicha correlación negativa se espera según la teoría de la violencia escolar vista como un proceso grupal de Salmivalli (1999, 2010). También se esperaba una correlación negativa entre la Subescala del Rol del Pacificador en el Ciclo de Violencia Escolar (SRPCVE) y la Escala de Ambiente Escolar Violento (EAEV) Por otro lado, en todos los casos hay una correlación significativa al nivel 0.01.

En lo que respecta a la correlación de las subescalas y la escala en su conjunto con otras variables se puede ver que en la mayoría de los casos hay una asociación significativa. En especial se esperaban correlaciones negativas de las subescalas de la ERECVE y en su conjunto con la Escala del Rol del Director al Clima Escolar (ERDCE). También nótese que la correlación más alta y negativa (-0.705) es la que presenta la Escala del Rol del Director en el Clima Escolar (ERDCE) y la Escala del Rol del Docente en Ciclo de la Violencia Escolar (ERDCVE). Con ello se puede suponer que, a mayor participación del Director para favorecer el clima escolar, menor es la aportación de los Docentes al ciclo de violencia escolar.

**Tabla 11.** Correlaciones entre las subescalas y la ERECVE en su conjunto, así como con otras variables relacionadas con los temas la violencia intrafamiliar y el clima y la violencia escolar

Escalas	SRVCVE	SRA1CVE	SRA2CVE	SROCVE	SRPCVE	ERECVE	EVI	EAEV	ERDCVE	ERDCE
Subescala del Rol de Víctima en el Ciclo de la Violencia Escolar	1									
Subescala del Rol del Agresor Solitario en el Ciclo de la Violencia Escolar	.547**	1								
Subescala del Rol del Agresor Social en el Ciclo de la Violencia Escolar	.295**	.435**	1							
Subescala del Rol del Observador en el Ciclo de la Violencia Escolar	.349**	.319**	.163**	1						
<b>Subescala del Rol del Pacificador en el Ciclo de la Violencia Escolar</b>	<b>-.115**</b>	<b>-.087**</b>	<b>-.213**</b>	<b>-.196**</b>	1					
Escala del Rol de los Estudiantes en el Ciclo de la Violencia Escolar	.622**	.624**	.462**	.619**	.368**	1				
Escala de Violencia Intrafamiliar	.121**	.134**	.144**	.091**	.173**	.260**	1			
Escala de Ambiente Escolar Violento	.289**	.224**	.194**	.455**	<b>-.099**</b>	.371**	.179**	1		
Escala del Rol del Docente en Ciclo de la Violencia Escolar	.190**	.170**	.169**	.174**	.130**	.313**	.326**	.504**	1	
<b>Escala del Rol del Director en el Clima Escolar</b>	<b>-.120**</b>	<b>-.109**</b>	<b>-.092**</b>	<b>-.141**</b>	<b>-.124**</b>	<b>-.234**</b>	<b>-.275**</b>	<b>-.381**</b>	<b>-.705**</b>	1

\*\* La correlación es significativa al nivel 0,01 (bilateral).

---

## V. DISCUSIÓN Y CONCLUSIONES

En este capítulo se presentan la discusión de resultados y las conclusiones del estudio. Para ello, se organizan tres apartados referidos a los logros, limitaciones del estudio, y recomendaciones para futuras investigaciones. En especial, se discuten los hallazgos del estudio a la luz de los antecedentes y fundamentos teóricos y empíricos del estudio y la medición de la violencia y el acoso escolar en el mundo (Olweus, 1983, 1991, 2001; Smith y Sharp, 1994; Craig y Harel, 2004; UNISEF, 2005, 2009, 2011; UNESCO, 2017). También se discuten los resultados logrados en contraste con los fundamentos y criterios metodológicos de las técnicas estadísticas del AFE y AFC empleadas para la estimación del modelo de cinco factores de la Escala del Rol del Estudiante en el Ciclo de la Violencia Escolar (ERECVE). Finalmente se concluye revisando las limitaciones del estudio y proponiendo recomendaciones con base en los estándares y criterios de calidad técnica y validez establecidos por organismos como la AERA, APA y el NCME (1999), y la ITC (2018).

### 5.1. Logros del estudio

Cabe recordar que el propósito general del estudio consistió en analizar las propiedades psicométricas relacionadas con las evidencias de validez de constructo, en el aspecto de la estructura interna y de la relación con otras variables de la ERECVE para educación secundaria, propósito que en términos generales se considera se logró satisfactoriamente. Como se mencionó en los antecedentes de la tesis, son pocos los instrumentos e indicadores de la medición de la violencia y acoso escolar que reportan avances en el estudio y sistematización de sus evidencias de calidad técnica y validez. En especial, organismos

---

internacionales como la UNICEF (2006), la UNESCO (2017), y la OPS, la OSP y la OMS (2003) han desarrollado instrumentos e indicadores para medir la violencia y el acoso escolar, pero debido a la necesidad estandarizar y economizar en las aplicaciones y el análisis entre una gran cantidad de variables de países de distintas regiones, se sacrifica el aspecto sustantivo y la fidelidad estructural asociada a teorías consolidadas.

En México, es más precario el panorama en contraste con lo que sucede en otras partes del mundo respecto al avance en el estudio y medición de la violencia y el acoso escolar. Son de reconocerse los esfuerzos de la SEP y la UNICEF México (2009), el INEE (2018) y la UEE (Rodríguez-Macías, Pérez-Morán, y Sarabia-Ocampo, 2017) en el campo en cuestión. Sin embargo, ninguna de las mediciones aplicadas hasta el momento presenta evidencias de la calidad técnica del aspecto sustantivo y de la validez de constructo asociadas a una teoría consolidada que aporte mayor información para interpretaciones a mayor profundidad.

Para que las autoridades educativas puedan diseñar políticas e intervenciones que impacten en la mejora del clima escolar es fundamental que cuenten con información granulada del fenómeno a intervenir. Es por ello y lo hasta ahora dicho, que los hallazgos de la presente tesis se consideran una gran aportación al estudio y medición de la violencia y el acoso escolar al ofrecer información de las evidencias de validez de constructo de la ERECVE fundamentada en la teoría reconocida y consolidada de la violencia vista como un proceso grupal propuesta por Olweus (1993) y extendida por Salmivalli (1999, 2010).

Además del logro satisfactorio del objetivo general de la tesis, se considera que cada uno de los objetivos específicos presentan un logro satisfactorio. En lo que respecta a la revisión de los antecedentes y los fundamentos teóricos de los temas centrales de la tesis, se revisaron con la suficiente amplitud los antecedentes del estudio y la medición de la violencia



y el acoso escolar tanto a nivel internacional como nacional. En especial, se profundizó en la revisión de la propuesta teórica de la violencia vista como proceso grupal de Olweus (1993) y extendida por Salmivalli (1999, 2010). Aunado a ello, se revisaron la estructura y propiedades psicométricas del Olweus Bully/Victim Questionnaire (BVQ, por sus siglas en inglés, Olweus, 1996, 2002), el *Participant-Role Questionnaire* desarrollado por Salmivalli, Lagerspetz, Bjorkqvist, Osterman, y Kaukialnen (1996), y la versión reducida y adaptada de dicho instrumento de Sutton y Smith (1999). También es importante mencionar la revisión realizada en los temas de las teorías y estándares de la validez de constructo relacionados a las evidencias de estructura interna y asociación con otras variables (Messick, 1989b; AERA, APA y NCME, 1999; Mislevy, 2009; Kane, 2016; ITC, 2018), así como los métodos estadísticos del AFE y AFC típicamente utilizados para analizar dichas evidencias (Yela, 1966; Bentler y Weeks, 1980; Long, 1983; Bollen, 1989; Jöreskog y Sörbon, 1989).

## 5.2. Limitaciones del estudio

A modo de discusión, es importante plantear, además de los logros alcanzados, las limitaciones que se presentaron a lo largo del estudio para conseguir los objetivos de investigación. Entre las distintas limitaciones presentes en el estudio es necesario discutir aquellas relacionadas con la necesidad de una revisión más detallada de las propiedades psicométricas de otros instrumentos de medición e indicadores de la violencia y el acoso escolar, así como de variables asociadas que ayuden a su explicación. Durante la revisión se focalizó la búsqueda bibliográfica de instrumentos asociados con la teórica de la violencia vista como proceso grupal propuesta por Olweus (1993) dejando de lado aquellos instrumentos e indicadores de constructos próximos o parecidos.

Otra de las limitaciones fue el logro razonable, pero no satisfactorio, del cumplimiento de los supuestos requeridos (normalidad, linealidad, colinealidad de las puntuaciones, ajuste del modelo) para aplicar el AFE. En especial, es importante recordar que la prueba de contrastes de bondad de ajuste del estadístico Kolmogorov-Smirnov presentó un estadístico de 0.81 significativo al 0.000 con 9542 *gl*. Con el resultado de dicha prueba se debe aceptar la hipótesis nula de normalidad y, por consiguiente, el incumplimiento del supuesto para el AFE. Es común en la literatura que la prueba de contrastes de bondad de ajuste del estadístico Kolmogorov-Smirnov resulte demasiado sensible a pequeñas desviaciones de la normalidad cuando se trabaja con muestras de gran tamaño, como en el caso de la presente tesis. Sin embargo, es posible, mediante una detección más fina de datos atípicos, contar con una base de datos que logre los supuestos de normalidad para cumplir con estándares rigurosos para la aplicación del AFE (Byrne, 2001; Costello y Osborne, 2005; Khan, 2006).

### **5.3. Recomendaciones para futuros estudios**

A modo de conclusión, es importante que, con base en las limitaciones del estudio identificadas y la reflexión de posibilidades de implementación para avanzar en el estudio y medición de la violencia y el acoso escolar, se propongan recomendaciones para futuros estudios:

- Revisar con mayor amplitud estudios en donde se reporten las propiedades psicométricas con mayor detalle de otros instrumentos de medición e indicadores de la violencia y el acoso escolar, así como de variables asociadas que ayuden a su explicación.

- Ajustar los procedimientos analíticos para mejorar el logro de los supuestos requeridos para aplicar el AFE y el AFC, en especial a lo referente a la normalidad de las puntuaciones.
- Aplicar a la par de a los AFE y AFC modelos más complejos o sofisticados como el Modelo Multidimensional de Teoría de Respuesta al Ítem con la finalidad de contar con mayor información de la calidad técnica de la ERECVE.
- Elaborar una propuesta de rediseño de la ERECVE con base en los resultados de la presente tesis, así como con la complementación de otras categorías de la teoría de la violencia vista como proceso grupal propuesta por Olweus (1993) y extendida por Salmivalli (1999, 2010).
- Incorporar variables de contexto para complementar las evidencias de asociación con otras variables de la ERECVE.

## VI. REFERENCIAS

- Aguilera, M., Muñoz, G., y Orozco, A. (2007). *Disciplina, violencia y consumo de sustancias nocivas a la salud en escuelas primarias y secundarias de México*. México: INEE. Recuperado de: <http://publicaciones.inee.edu.mx/buscadorPub/P1/D/230/P1D230.pdf>
- Alsaker, F., y Nägele, C. (2008). Bullying in kindergarten and prevention. *An international perspective on understanding and addressing bullying*, 1, 230-252.
- American Educational Research Association, American Psychological Association, y National Council on Measurement in Education. (1985). *Standards for educational and psychological testing*. Washington, D.C.: American Psychological Association.
- American Educational Research Association, American Psychological Association, y National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- American Educational Research Association, American Psychological Association, y National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- American Psychological Association. American Educational Research Association y National Council on Measurement in Education. (1954). *Technical recommendations for psychological tests and diagnostic techniques*. Washington: American Psychological Association.
- Arseneault, L., Bowes, L., y Shakoor, S. (2010). Bullying victimization in youths and mental health problems: "Much ado about nothing"? *Psychological Medicine*, 40, 717–729.
- Atlas, R., y Pepler, D. (1998). Observations of bullying in the classroom. *The Journal of Educational Research*, 92, 86 –99.
- Batista, J., y Coenders, G. (2000). *Modelos de Ecuaciones Estructurales (modelos para el análisis de relaciones causales)*. Madrid, Editorial La Muralla, S.A.
- Beauducel, A. (2001). On the generalizability of factors: The influence of changing contexts of variables on different methods of factor extraction. *Methods of Psychological Research Online*, 6(1), 69-96.
- Bentler, P. (2012). *EQS 6.2 for windows*. Encino, CA: Multivariate Software.
- Bentler, P., y Weeks, D. (1980). Linear structural equations with latent variables. *Psychometrika*, vol. 45, 289-308.
- Blalock, H. (1966). *Estadística Social*. México: Fondo de Cultura Económica.

- Boivin, M., Hymel, S., y Hodges, E. (2001). Toward a process view of peer rejection and harassment. En J. Juvonen y S. Graham (Eds.), *Peer harassment in school: The plight of the vulnerable and victimized*, pp. 265-289. New York: Guilford.
- Bollen, K. (1989): *Structural Equations with Latent Variables*. Nueva York: Wiley.
- Boulton, M., y Smith, P. (1994). Bully/victim problems in middle-school children: Stability, self-perceived competence, peer perceptions and peer acceptance. *British Journal of Developmental Psychology*, 12(3), 315-329.
- Boulton, M., y Underwood, K. (1992). Bully/Victim problems among middle school children. *British Journal of Educational Psychology*, 62:73–87.
- Bowling, N., y Beehr, T. (2006). Workplace harassment from the victim's perspective: A theoretical model and meta-analysis. *Journal of Applied Psychology*, 91 (5), 998-1012.
- Brown, J. (1996). Testing in language programs. Upper Saddle River, NJ: Prentice Hall Regents.
- Brown, T. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.
- Byrne, B. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum
- Caravita, S., Miragoli, S., y Di Blasio, P. (2009). Why should I behave in this way? Rule discrimination within the school context related to bullying. *Social Development*, ed L. R. Elling (New York, NY: Nova Science Publishers), 269–290.
- Card, N., y Little, T. (2006). Proactive and reactive aggression in childhood and adolescence: A meta-analysis of differential relations with psychosocial adjustment. *International Journal of Behavioral Development*, 30, 446-480.
- Carroll, J. (1953). An analytic solution for approximating simple structure in factor analysis. *Psychometrika*, 18,79-87.
- Cattell, R. (1966). The Scree Test for the number of factors. *Multivariate Behavioral Research*, 1, 141-16.
- Comrey, A. (1973). *A first course in factor analysis*. Nueva York: Academic Press.
- Costello, A. y Osborne, J. (2005). Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From your Analysis. *Practical Assessment, Research y Evaluation*, 10 (7), 1-9.
- Craig, M, y Harel, Y. (2001). Bullying, physical fighting and victimization. *Young people's health in context: International report from the HBSC*, 2, 133-144.

- Craig, M., y Pepler, J. (1997). Observations of bullying and victimization in the schoolyard. *Canadian Journal of School Psychology, 13*, 41-60.
- Craig, M., Peters D, y Konarski R. (1998). *Bullying and victimization among Canadian school children*. Ottawa, Applied Research Branch, Strategic Policy, Human Resources Development Canada.
- Craig, W., Schumann, L., Edge, M., y Teske, C. (2012). *Are bullying and victimization on the rise in Canada?* Paper presented at the Canadian Public Health Association Conference, Edmonton, Alberta, Canada.
- Craig, W., y Harel, Y. (2004). Bullying, physical fighting and victimization. En C. Currie, C. Roberts, A. Morgan, R. Smith, W. Settertobulte, O. Samdal y V. Barnekow (Eds.), *Young people's health in context. Health Behaviour in School-aged Children (HSBC) Study*, pp. 133-144. Copenhagen: WHO Publications.
- Cronbach, L. (1984). *Essentials of psychological testing* (4th ed.). New York, NY: Harper y Row.
- Cronbach, L., y Meehl, P. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281-302.
- Currie, C., Roberts, C., Settertobulte, W., Morgan, A., Smith, R., Samdal, O., Barnekow, V., y World Health Organization. (2004). *Young people's health in context: Health Behaviour in School-aged Children (HSBC) study: international report from the 2001/2002 survey* (No. EUR/04/5048327). Copenhagen: WHO Regional Office for Europe.
- D'Angelo, L., y Fernández, L. (2010). Clima, conflictos y violencia en la escuela: un estudio en escuelas secundarias de gestión pública y privada del Área Metropolitana de Buenos Aires. *Buenos Aires: Unicef-Flacso*.
- Del Barrio, C., Gutiérrez, H., Barrios, A., van der Meule, K., y Granizo, L. (2005). Maltrato por abuso de poder entre escolares, ¿de qué estamos hablando? *Pediatría Atención Primaria, 7*(25), 75-100.
- Díaz-Aguado, M., Martínez-Arias, R., y Seoane, G. (2004). *Prevención de la violencia y lucha contra la exclusión desde la adolescencia*. Instituto de la Juventud.
- Due, P., Holstein, B., Lynch, J., Diderichsen, F., Nic Gabhein, S., Scheidt, P., y The Health Behaviour in School-Aged Children Bullying Working Group. (2005). Bullying and symptoms among school-aged children: International comparative cross sectional study in 28 countries. *The European Journal of Public Health, 15*, 128–132.
- Dunn, L., Dunn L., Whetton, C., y Pintilie, D. (1982). *British picture vocabulary scale*. Windsor, Berks: NFER-NELSON. 68 p.
- Eisenbraun, K. (2007). Violence in schools: Prevalence, prediction, and prevention. *Aggression and violent behavior, 12*(4), 459-469.

- Elosua, P. (2003). Sobre la validez de los tests. *Psicothema*, 15(2), 315-321.
- Embretson, S. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin*, 93(1), 179-197.
- Eslea, M., Menesini, E., Morita, Y., O'Moore, M., Moran-Merchan, J., Pereira, B., y Smith, P. (2004). Friendship and loneliness among bullies and victims: Data from seven countries. *Aggressive Behavior*, 30, 71–83.
- Everitt, B. y Wykes, T. (2001). *Diccionario de Estadística para Psicólogos*. España: Ariel.
- Farrington, D. (1992). Understanding and preventing bullying. In: Tonry M, Morris N, editors. *Crime and justice: An annual review of research*. Vol. 17. Chicago: University of Chicago Press. pp 381–458.
- Farrington, D. P. (1993). Understanding and preventing bullying. *Crime and justice*, 17, 381-458.
- Fierro, C. (2013). Convivencia inclusiva y democrática: Una perspectiva para gestionar la seguridad escolar. *Sinéctica*, (40), 01-18.
- Galton, F. (1889). *2ature Inheritance*. Londres: Macmillan.
- George, D., y Mallery, M. (2003). *Using SPSS for Windows step by step: a simple guide and reference*. Boston, MA: Allyn y Bacon.
- Glutting, J. (2002). Some psychometric properties of a system to measure ADHD. *Measurement and Evaluation in Counseling and Development*. 34, 194-209.
- Gómez, J., Chacón, S., y Moreno, R. (2000). Validez de constructo: el uso del análisis factorial exploratorio-confirmatorio para obtener evidencias de validez. *Psicothema*, 12 (2) 442-446.
- Goossens, F., Olthof, T., y Dekker, P. (2006). New participant role scales: Comparison between various criteria for assigning roles and indications for their validity. *Aggressive Behavior*, 32, 343–357.
- Gorsuch, R. (1983). *Factor analysis*. Hillsdale, NJ: Lawrence Erlbaum.
- Greene, W. (2000). *Econometric Analysis*. Fourth ed., New York, NY: Macmillan.
- Guion, R. (1980). On Trinitarian doctrines of validity. *Professional Psychology*, 11(3), 385-398.
- Hair, J., Anderson, R., Tatham, R., y Black, W. (1999). *Análisis Multivariante*. Madrid: Prentice Hall.
- Haney, W. y Madaus, G. (1991). The evolution of ethical and technical standards for testing. En R. K. Hambleton y J. N. Zaal (Eds.). *Advances in educational and psychological testing: Theory and applications* (pp.395-425). Boston: Kluwer.

- Hartup, W. (1996). The company they keep: Friend-ships, and their developmental significance. *Child Dev* 67:1–13.
- Hatcher, L. (1994). *A Step-by-Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modelling*. Cary, NC: Sas Institute Inc.
- Hawker, D., y Boulton, M. (2000). Twenty years' research on peer victimization and psychosocial maladjustment: A meta-analytic review of cross-sectional studies. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 41(4), 441-455.
- Hawkins, L., Pepler, D., y Craig, W. (2001). Naturalistic observations of peer interventions in bullying. *Social Development*, 10, 512–527.
- Heinemann, P. (1972). *Mobbing-group elected among children and adults*. Stockholm: Nature and Culture.
- Holgado–Tello, F., Chacón–Moscoso, S., Barbero–García, I., y Vila–Abad, E. (2010). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality y Quantity*, 44(1), 153.
- Hu, L., y Bentler, P. (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.
- Hu, L., Bentler, P., y Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological bulletin*, 112(2), 351.
- INEE (2015). Los docentes en México. Informe 2015. México: INEE
- INEE (2017c). *Estudio Internacional de Educación Cívica y Ciudadana (ICCS, 2016). Informe nacional de resultados*. México: autor. Recuperado el 20 de marzo de 2018 de: <http://publicaciones.inee.edu.mx/buscadorPub/P2/A/332/P2A332.pdf>
- INEE (2017g). *Estudio Internacional de Educación Cívica y Ciudadana (ICCS, 2016). Informe nacional de resultados*. Presentación usada para una mesa pública realizada el 13 de noviembre de 2017.
- INEE (2017h). *Informe de resultados Planea 2015. El aprendizaje de los alumnos de sexto de primaria y tercero de secundaria en México. Lenguaje y Comunicación. Matemáticas*. México: autor. Recuperado de: <http://publicaciones.inee.edu.mx/buscadorPub/P1/D/246/P1D246.pdf>
- INEE (2018). *Educación obligatoria en México. Informe 2018*. Ciudad de México: INEE. Recuperado de: [https://www.inee.edu.mx/portalweb/informe2018/04\\_informe/capitulo\\_0601.html](https://www.inee.edu.mx/portalweb/informe2018/04_informe/capitulo_0601.html)
- International Test Commission. (2018). *ITC Guidelines for the Large-Scale Assessment of Linguistically and Culturally Diverse Populations*. [www.InTestCom.org]



- Jimerson, S. R., y Furlong, M. J. (2006). The handbook of school violence and school safety. *From research to practice*.
- Jöreskog, K., y Sörbom, D. (1989). LISREL 7: A Guide to the Program and Applications, SPSS. *Inc., Chicago, IL*.
- Juvonen, J., Graham, S., y Schuster, B. (2003). Bullying among young adolescents: The strong, weak, and troubled. *Pediatrics*, 112, 1231–1237.
- Kahn, J. (2006). Factor analysis in Counseling Psychology research, training and practice. *The Counseling Psychologist*, 34, 1-36.
- Kaiser, H. y Rice, J. (1974). Little jiffy, mark IV. *Educational and Psychological Measurement*, 34(1):111–117.
- Kaiser, H. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23(3), 187-200.
- Kaiser, H. (1970). A second generation little jiffy. *Psychometrika*, 35(4):401–415.
- Kane, M. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1-73.
- Kane, M. (2016). Validity as the evaluation of the claims based on test scores. *Assessment in Education: Principles, Policy y Practice*, 23(2), 309-311.
- Kline, P. (2000). *Handbook of Psychological Testing*. London: Routledge.
- Lagerspetz, K., Björkqvist, K., Berts, M., y King, E. (1982). Group aggression among school children in three schools. *Scandinavian Journal of Psychology*, 23(1), 45-52.
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., y Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Anales de psicología*, 30(3), 1151-1169.
- Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological reports*, 3(3), 635-694.
- Long, J. (1983). *Confirmatory factor analysis: A preface to LISREL* (Vol. 33). Sage Publications.
- Martinez Arias, M. R. (1999). *El análisis multivariante en la investigación científica*. Madrid: La Muralla.
- Martínez Arias, R. (1995). *Psicometría*. Madrid: Síntesis Psicológica
- Martínez Arias, R. (1999). *El análisis multivariante en la investigación científica*. La Muralla.

- Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, 35(11), 1012-1027.
- Messick, S. (1987). Validity. *ETS Research Report Series*, 1987(2), i-208.
- Messick, S. (1989b). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., 13-103). New York: Macmillan.
- Messick, S. (1993). Foundations of validity: Meaning and consequences in psychological assessment. *ETS Research Report Series*, 1993(2), i-18.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American psychologist*, 50(9), 741.
- Mislevy, R. (2009). Validity from the Perspective of Model-Based Reasoning. CRESST Report 752. *National Center for Research on Evaluation, Standards, and Student Testing (CRESST)*.
- Muthén, B., y Hofacker, C. (1988). Testing the assumptions underlying tetrachoric correlations. *Psychometrika*, 53(4), 563-577.
- Nansel, T., Craig, W., Overpeck, M., Saluja, G., y Ruan, W. (2004). Cross-national consistency in the relationship between bullying behaviors and psychosocial adjustment. *Archives of pediatrics y adolescent medicine*, 158(8), 730-736.
- Nielsen, M., Rugulies, R., Christensen, K., Smith-Hansen, L. y Kristensen, T. (2006). Psychosocial work environment predictors of short and long spells of sickness absence during a two-year follow-up. *Journal of Occupational and Environmental Medicine*, 48, 591-595.
- Nunnally, J. y Bernstein, I. (1995). *Teoría Psicométrica*. México: Mc Graw Hill.
- Nunnally, J., y Bernstein, I. (1994). *Psychometric Theory* (3a ed.). New York: McGraw-Hill
- Nunnally, J., y Bernstein, I. (1995) *Teoría Psicométrica*. México: Mc Graw Hill.
- Oliva, L., y Calderón, M. (2014). Los actores de la violencia escolar y el rol que desempeñan. *Entreciencias: Diálogos en la Sociedad del Conocimiento*, 2(5), 277-286.
- Olweus, D. (1973a). *Chickens and bully: Research on school bullying*. Stockholm: Almqvist y Wiksell.
- Olweus, D. (1973b). Personality and Aggression. In: Cole, J.K. and Jensen, D.D., Eds., *Nebraska Symposium on Motivation*, University of Nebraska Press, Lincoln.
- Olweus, D. (1978). *Aggression in the schools: Bullies and whipping boys*. Hemisphere.

- Olweus, D. (1983). Low school achievement and aggressive behavior in adolescent boys. *Human development: An interactional perspective*, 353-365.
- Olweus, D. (1991). Bully/victim problems among schoolchildren: Basic facts and effects of a school based intervention program. *The development and treatment of childhood aggression*, 17, 411-448.
- Olweus, D. (1993). *Conductas de acoso y amenaza entre escolares*. Madrid: Morata (1998 fecha de la edición en castellano).
- Olweus, D. (1996). *The revised Olweus bully/victim questionnaire*. University of Bergen, Research Center for Health Promotion.
- Olweus, D. (2001). Peer harassment. A critical analysis and some important issues. En J. Juvonen y S. Graham (Eds.). *Peer harassment in school*. (pp. 3- 20). New York: The Guilford Press.
- Olweus, D. (2002). General information about the Revised *Olweus Bully/Victim Questionnaire, PC program and teacher handbook* (pp. 1-12). Bergen, Norway: Mineo, Reaserch Centre For Health Promotion (HEMIL), University of Bergen.
- Organización Panamericana de la Salud, OPS, Oficina Sanitaria Panamericana, OSP y Oficina Regional de la Organización Mundial de la Salud, OMS (2003). *Informe mundial de violencia y de la salud 2003*. Recuperado de: [http://www.paho.org/spanish/dd/pub/violencia\\_2003.htm](http://www.paho.org/spanish/dd/pub/violencia_2003.htm)
- Ortega, R. (2000). A Global, Ecological and Cultural Model for Dealing with Problems of Violence in European Compulsory Schools. *Paper presentado en 6th Meeting of TMR Programme Nature and Prevention of Bullying and Social Exclusion*. Cruz Quebrada-Dafundo. Lisboa
- Ortega, R. (2010). *Agresividad injustificada, bullying y violencia escolar*. Madrid: Alianza Editorial.
- Ortega, R., Sánchez, V., y Menesini, E. (2002). Violencia entre iguales y desconexión moral: un análisis transcultural. *Psicothema*, 14, 37-49.
- Ostvik, K., y Rudmin, F. (2001). Bullying and hazing among Norwegian army soldiers: Two studies of prevalence, context, and cognition. *Military Psychology*, 13(1), 17-39.
- Pagano, R. (1998). *Estadística para las ciencias del Comportamiento*. México: Thompson.
- Pearson, K. (1901). On lines and planes of closest fit to systems of point in space. *Philosophical Magazine*, 6, 559-572.
- Pepler, D., Jiang, D., Craig, W., y Connolly, J. (2008). Developmental trajectories of bullying and associated factors. *Child development*, 79(2), 325-338.

- Pérez, E. y Medrano, L. (2010). Análisis Factorial Exploratorio: Bases Conceptuales y Metodológicas. *Revista Argentina de Ciencias del Comportamiento*, 2(1), 58-66.
- Pérez-Gil, J., Chacón, S., y Moreno, R. (2000). Validez de constructo: el uso de análisis factorial exploratorio-confirmatorio para obtener evidencias de validez. *Psicothema*, 12(Supl2), 442-446.
- Pikas, A. (1975). *Entonces dejamos de intimidar*. Estocolmo: Prisma.
- Pulido, R. (2008). Estudio de los roles en la violencia entre iguales: víctima, agresor y observador. *International Journal of Developmental and Educational Psychology*, 3(1), 279-288.
- Reise, S., Waller, N., y Comrey, A. (2000). Factor analysis and scale revision. *Psychological assessment*, 12(3), 287.
- Revelle, W. (2008). *psych: Procedures for Personality and Psychological Research*. R package version 1.0-58. Recuperado de: <https://cran.r-project.org/web/packages/psych/psych.pdf>
- Richaud, M. (2005). Desarrollos del análisis factorial para el estudio de ítem dicotómicos y ordinales. *Revista Interdisciplinaria*, 22(2), 237-251.
- Rigby, K., y Slee, P. (1991). Bullying among Australian school children: Reported behavior and attitudes toward victims. *The journal of social psychology*, 131(5), 615-627.
- Rodicio-García, M., y Iglesias-Cortizas, M. (2011). *El acoso escolar. Diagnóstico y prevención*. 36-99. Madrid: Biblioteca Nueva.
- Rodkin, P., Farmer, T., Pearl, R., y Van Acker, R. (2006). They're cool: Social status and peer group supports for aggressive boys and girls. *Social Development*, 15(2), 175-204.
- Rodríguez-Macías, J., Pérez-Morán, J., y Sarabia-Ocampo, Z. (2017). *Violencia y convivencia escolar en escuelas primarias, secundarias y planteles de educación media superior de Baja California*. Reporte Técnico UEE 18-01/2017. Ensenada: Unidad de Evaluación Educativa.
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5-12 (BETA). *Journal of statistical software*, 48(2), 1-36.
- Salmivalli, C. (1999). Participant role approach to school bullying: implications for interventions. *Journal of Adolescence*, 22, 453-459.
- Salmivalli, C. (2010). Bullying and the peer group: A review. *Aggression and Violent Behavior*, 15, 112-120.
- Salmivalli, C. (2010). Bullying and the peer group: A review. *Aggression and Violent Behavior*, 15, 112-120.

- Salmivalli, C., y Voeten, M. (2004). Connections between attitudes, group norms, and behaviour in bullying situations. *International Journal of Behavioral Development*, 28, 246–258.
- Salmivalli, C., Huttunen, A., y Lagerspetz, K. M. J. (1997). Peer networks and bullying in schools. *Scandinavian Journal of Psychology*, 38, 305– 312.
- Salmivalli, C., Lagerspetz, K., Bjorkqvist, K., Osterman, K., y Kaukialnen, A. (1996). Bullying as a group process: participant roles and their relations to social status within the group. *Aggressive Behavior*, 22,1-15.
- Salmivalli, C., Lappalainen, M., y Lagerspetz, M. (1998). Stability and change of behavior in connection with bullying in schools: A two year follow up. *Aggressive Behavior*, 24, 205-218.
- Schwartz, D. (2000). Subtypes of victims and aggressors in children's peer groups. *Journal of abnormal child psychology*, 28(2), 181-192.
- SEP y UNICEF México (2009). *Informe nacional sobre violencia de género en educación básica en México*. México:SEP. Recuperado en: [https://www.unicef.org/mexico/spanish/Estudio\\_violencia\\_genero\\_educacion\\_basica\\_Part1.pdf](https://www.unicef.org/mexico/spanish/Estudio_violencia_genero_educacion_basica_Part1.pdf)
- Sharp, S., y Cowie, H. (1994). Empowering pupils to take positive action against bullying. *School bullying: Insights and perspectives*, 108-131.
- Sierra Bravo, R. (1981). *Análisis Estadístico y Modelos Matemáticos*. Madrid: Paraninfo.
- Sierra. C. (2010). Violencia escolar. Perfiles psicológicos de agresores y víctimas. *Poliantea* 10. Volumen VI, pp. 53-71. Enero - Junio 2010 / Bogotá, Colombia.
- Smith, P., y Sharp, S. (1994). The problem of school bullying. In: Smith PK, Sharp S, eds. *School bullying*. London, Routledge.
- Spearman, C. (1904). General Intelligence, objectively determined and measured. *American Journal of Psychology*, 15, 201-270.
- Springer, A., Cuevas, M., Ortiz, Y., Keis, K., y Wilkinson, A. (1757). The association of bullying and school cohesion in Colombia secondary school students. *Global Health Promotion*, 9759, 1-12. Lappalainen Lappalainen
- SRSG-VAC (2016). *Ending the Torment: Tackling bullying from the schoolyard to cyberspace*. New York: United Nations. Recuperado de: <https://hivhealthclearinghouse.unesco.org/library/documents/ending-torment-tackling-bullying-schoolyard-cyberspace>
- Sutton, J., y Smith, P. (1999). Bullying as a group process: An adaptation of the participant role approach. *Aggressive Behavior*, 25, 97 –111.

- Tabachnick, B., y Fidell, L. (2001). *Using multivariate statistics*. New York: Harper y Row.
- Tabachnick, B., Fidell, L., y Ullman, J. (2007). *Using multivariate statistics* (5th edi). Boston, MA: Pearson.
- Thompson, B. (2004). Exploratory and confirmatory factor analysis. Washington, DC: American Psychological Association.
- Thornberg, R., y Jungert, T. (2014). School bullying and the mechanisms of moral disengagement. *Aggressive behavior*, 40(2), 99-108.
- Trautmann A. (2008). Maltrato entre pares o bullying. Una visión actual. *Revista Chilena de Pediatría*; 79 (1) 13-20.
- Twemlow, S., Fonagy, P., y Sacco, F. (2004). The role of the bystander in the social architecture of bullying and violence in schools and communities. *Annals of the New York Academy of Sciences*, 1036(1), 215-232.
- Ullman, J. (1996). Structural Equation Modelling In Tabachnick B. And Fidell L. (1996) *Using Multivariate Statistics*, 3.
- UNESCO y Ministerio de Educación de España (2012). Informe de progreso. Educación para la Convivencia y la Cultura de la Paz en América Latina y El Caribe. Recuperado de: <http://www.unesco.org/new/fileadmin/MULTIMEDIA/FIELD/Santiago/pdf/Educacion-Convivencia-Paz.pdf>
- UNESCO (2017). *School Violence and Bullying: Global Status Report*. Francia: UNESCO.
- UNESCO Institute for Statistic (2011). Global Education Digest 2011: Comparing Education Statistics Across the World. Recuperado de: [http://www.ungei.org/global\\_education\\_digest\\_2011\\_en.pdf](http://www.ungei.org/global_education_digest_2011_en.pdf)
- UNICEF (2005). Violence against children in the Caribbean region/ Regional assessment/ UN Secretary General's Study on Violence. Child Protection Section. UNICEF Regional Office for Latin America and the Caribbean.
- UNICEF (2011). *Violencia escolar en América Latina y el Caribe Superficie y fondo*. Recuperado de: [https://www.unicef.org/lac/violencia\\_escolar\\_OK.pdf](https://www.unicef.org/lac/violencia_escolar_OK.pdf).
- UNICEF (2016). *Annual Results Report 2016. Child Protection*. Recuperado de: [https://www.unicef.org/protection/files/2016\\_CP\\_ARR.pdf](https://www.unicef.org/protection/files/2016_CP_ARR.pdf)
- UNICEF. (2009). Informe nacional sobre violencia de género en la educación básica en México. *México, DF*.
- UNICEF. (2012). Progress for children. A report card on adolescents. (UNICEF Publication No. 10).

- UNICEF-IRC (2018). *Bullying: a global challenge requires a global measure*. Recuperado en: <https://hivhealthclearinghouse.unesco.org/library/documents/bullying-global-challenge-requires-global-measure>
- Uriel, E., y Aldas, J. (2005). *Análisis Multivariante Aplicado*. España: Thomson.
- Vaillancourt, T., Hymel, S., y McDougall, P. (2003). Bullying is power: Implications for school-based intervention strategies. *Journal of Applied School Psychology, 19*(2), 157–176.
- Valdés, A., Yañez, A., y Carlos, E. (2013). Diferencias entre subgrupos de estudiantes involucrados en el bullying: víctimas, agresores-víctimas y agresores. *Liberabit. Revista de Psicología, 19*(2), 215-222.
- Veenstra, R., Lindenberg, S., Oldehinkel, A., De Winter, A., Verhulst, F., y Ormel, J. (2005). Bullying and victimization in elementary schools: a comparison of bullies, victims, bully/victims, and uninvolved preadolescents. *Developmental psychology, 41*(4), 672.
- Whitney, I., y Smith, P. (1993). A survey of the nature and extent of bullying in junior/middle and secondary schools. *Educational research, 35*(1), 3-25.
- Winter, J., y Dodou, D. (2012). Factor recovery by principal axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size. *Journal of Applied Statistics, 39*(4), 695-710.
- Yela, M. (1966). Los tests y el análisis factorial. En B. Szekeli (Ed.), *Los Tests*, 153-178. Buenos Aires: Kapelusz.
- Yela, M. (1996). Los tests y el análisis factorial. *Psicothema, 8*(Sup), 73-88.

## Apéndice 1. Descriptivos de las preguntas de la Escala de los Roles de los Estudiantes en el Ciclo de la Violencia Escolar

### Sección I. Descriptivos correspondientes a las preguntas de la Subescala del Rol de la Víctima en el Ciclo de la Violencia Escolar

Pregunta S144.1. Sufro violencia por parte de mis compañeros(as) de clase.

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	8292	80.0	80.0
Raramente	1599	15.4	95.4
Mensualmente	157	1.5	96.9
Semanalmente	149	1.4	98.4
Diariamente	170	1.6	100.0
Total	10367	100.0	

Pregunta S144.2. Sufro violencia por parte de compañeros(as) de otros salones de clase.

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	9197	88.7	88.7
Raramente	907	8.7	97.5
Mensualmente	118	1.1	98.6
Semanalmente	59	0.6	99.2
Diariamente	86	0.8	100
Total	10367	100	

Pregunta S144.3. Sufro violencia por parte de alguno de los docentes de mi escuela.

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	9873	95.2	95.2
Raramente	319	3.1	98.3
Mensualmente	70	0.7	99
Semanalmente	44	0.4	99.4
Diariamente	61	0.6	100
Total	10367	100	

Pregunta S144.4. Sufro violencia por parte de algún miembro del personal administrativo de la escuela.

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	10030	96.7	96.7
Raramente	211	2	98.8
Mensualmente	49	0.5	99.3
Semanalmente	33	0.3	99.6
Diariamente	44	0.4	100
Total	10367	100	

Pregunta S144.5. Sufro violencia por parte de compañeros(as) fuera de mi escuela o en la comunidad en donde vivo.

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	9674	93.3	93.3
Raramente	514	5	98.3
Mensualmente	65	0.6	98.9
Semanalmente	47	0.5	99.4
Diariamente	67	0.6	100
Total	10367	100	



## Sección II. Descriptivos correspondientes a las preguntas de la subescala del Rol del estudiante agresor solitario en el ciclo de la violencia escolar

Pregunta S144.6. Practico violencia sobre mis compañeros(as) de clase.

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	9465	91.3	91.3
Raramente	679	6.5	97.8
Mensualmente	89	0.9	98.7
Semanalmente	60	0.6	99.3
Diariamente	74	0.7	100
Total	10367	100	

Pregunta S144.5. Practico violencia sobre compañeros(as) de otros salones de clase.

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	9711	93.7	93.7
Raramente	487	4.7	98.4
Mensualmente	58	0.6	98.9
Semanalmente	48	0.5	99.4
Diariamente	63	0.6	100
Total	10367	100	

Pregunta S144.8. Practico violencia sobre compañeros(as) fuera de mi escuela o en la comunidad en donde vivo.

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	9610	92.7	92.7
Raramente	527	5.1	97.8
Mensualmente	91	0.9	98.7
Semanalmente	57	0.5	99.2
Diariamente	82	0.8	100
Total	10367	100	

## Sección III. Descriptivos correspondientes a las preguntas de la subescala del Rol del estudiante agresor social en el ciclo de la violencia escolar

Pregunta S143.1. Me reúno con un grupo de amigos(as) para practicar violencia sobre compañeros(as) de mi escuela.

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	8832	85.2	85.2
Raramente	974	9.4	94.6
Mensualmente	170	1.6	96.2
Semanalmente	143	1.4	97.6
Diariamente	248	2.4	100
Total	10367	100	

Pregunta S143.3. Me reúno fuera de mi escuela con un grupo de amigos(as) para realizar actos vandálicos.

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	9278	89.5	89.5
Raramente	609	5.9	95.4
Mensualmente	167	1.6	97
Semanalmente	121	1.2	98.1
Diariamente	192	1.9	100
Total	10367	100	

Pregunta S143.5. Realizo actos de vandalismo para recibir la admiración de mis amigos(as).

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	9511	91.7	91.7
Raramente	544	5.2	97
Mensualmente	113	1.1	98.1
Semanalmente	89	0.9	98.9
Diariamente	110	1.1	100
Total	10367	100	

Pregunta S143.7. Practico violencia sobre otros compañeros(as) para sentirme superior a ellos.

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	9350	90.2	90.2
Raramente	694	6.7	96.9
Mensualmente	111	1.1	98
Semanalmente	103	1	98.9
Diariamente	109	1.1	100
Total	10367	100	

## Sección IV. Descriptivos correspondientes a las preguntas de la subescala del Rol del estudiante observador en el ciclo de la violencia escolar

Pregunta S144.9. Veo violencia entre mis compañeros(as) de clase.

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	5189	50.1	50.1
Raramente	3263	31.5	81.5
Mensualmente	575	5.5	87.1
Semanalmente	513	4.9	92
Diariamente	827	8	100
Total	10367	100	

Pregunta S144.10. Veo violencia entre compañeros(as) de otros salones de clase.

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	5403	52.1	52.1
Raramente	3105	30	82.1
Mensualmente	623	6	88.1
Semanalmente	502	4.8	92.9
Diariamente	734	7.1	100
Total	10367	100	

Pregunta S144.11. Veo violencia entre compañeros(as) de otros salones de clase.

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	5403	52.1	52.1
Raramente	3105	30	82.1
Mensualmente	623	6	88.1
Semanalmente	502	4.8	92.9
Diariamente	734	7.1	100
Total	10367	100	

## Sección V. Descriptivos correspondientes a las preguntas de la subescala del Rol del estudiante pacificador en el ciclo de la violencia escolar

Pregunta S143.2. Me reúno con un grupo de compañero(as) para realizar actividades extraescolares.

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	725	7	7
Raramente	1105	10.7	17.7
Mensualmente	1196	11.5	29.2
Semanalmente	3480	33.6	62.8
Diariamente	3861	37.2	100
Total	10367	100	

Pregunta S143.4. Me reúno fuera de mi escuela con un grupo de amigos(as) para realizar actividades deportivas o culturales.

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	1165	11.2	11.2
Raramente	1177	11.4	22.6
Mensualmente	1002	9.7	32.3
Semanalmente	3142	30.3	62.6
Diariamente	3881	37.4	100
Total	10367	100	

Pregunta S143.6. Realizo actividades de voluntariado con el fin de ayudar a personas en condiciones de vulnerabilidad

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	569	5.5	5.5
Raramente	722	7	12.5
Mensualmente	1187	11.4	23.9
Semanalmente	3658	35.3	59.2
Diariamente	4231	40.8	100
Total	10367	100	

Pregunta S144.8. Ayudo a otros compañeros(as) que se encuentran en problemas para mejorar su situación.

Válidos	Frecuencia	Porcentaje	Porcentaje acumulado
Nunca	1836	17.7	17.7
Raramente	1431	13.8	31.5
Mensualmente	1516	14.6	46.1
Semanalmente	3597	34.7	80.8
Diariamente	1987	19.2	100
Total	10367	100	

## Apéndice 2. Estimaciones de los coeficientes de correlación de Spearman entre los ítems de la ERECVE

Ítems	43.1	43.2	43.3	43.4	43.5	43.6	43.7	43.8	44.1	44.2	44.3	44.4	44.5	44.6	44.7	44.8	44.9	44.10	44.11
43.1	1.000																		
43.2	.158**	1.000																	
43.3	.350**	.116**	1.000																
43.4	.113**	.338**	.099**	1.000															
43.5	.291**	.086**	.416**	.088**	1.000														
43.6	.107**	.264**	.098**	.348**	.086**	1.000													
43.7	.312**	.100**	.284**	.101**	.379**	.115**	1.000												
43.8	.066**	.244**	.046**	.262**	.030**	.411**	.042**	1.000											
44.1	.058**	.081**	-.006	.068**	.017	.104**	.086**	.101**	1.000										
44.2	.069**	.089**	.029**	.086**	.030**	.101**	.095**	.089**	.435**	1.000									
44.3	.053**	.029**	.045**	.023*	.059**	.048**	.058**	.033**	.200**	.254**	1.000								
44.4	.068**	.033**	.059**	.039**	.071**	.036**	.111**	.024*	.113**	.179**	.346**	1.000							
44.5	.077**	.053**	.066**	.072**	.054**	.062**	.101**	.049**	.221**	.299**	.219**	.170**	1.000						
44.6	.164**	.072**	.100**	.057**	.107**	.063**	.240**	.044**	.181**	.169**	.103**	.131**	.188**	1.000					
44.7	.172**	.065**	.123**	.081**	.102**	.068**	.204**	.050**	.117**	.193**	.086**	.078**	.206**	.491**	1.000				
44.8	.152**	.063**	.130**	.078**	.120**	.059**	.156**	.049**	.082**	.144**	.098**	.128**	.256**	.351**	.411**	1.000			
44.9	.032**	.151**	.034**	.120**	-.013	.196**	.038**	.202**	.292**	.166**	.065**	.047**	.141**	.183**	.135**	.125**	1.000		
44.10	.035**	.156**	-.023*	.138**	-.009	.198**	.031**	.210**	.215**	.170**	.062**	.048**	.134**	.157**	.135**	.139**	.718**	1.000	
44.11	.056**	.158**	.005	.158**	.012	.207**	.033**	.200**	.152**	.121**	.063**	.049**	.153**	.112**	.118**	.151**	.542**	.639**	1.000

\*\* La correlación es significativa al nivel 0.01 (bilateral)

\* La correlación es significativa al nivel 0.05 (bilateral)