



VALIDACIÓN PSICOMÉTRICA DE UN CUESTIONARIO PARA EVALUAR EMOCIONES ACADÉMICAS DURANTE EXÁMENES ESCOLARES

PSYCHOMETRIC VALIDATION OF A QUESTIONNAIRE FOR ASSESSING ACADEMIC EMOTIONS DURING SCHOOL EXAMINATIONS

Wilmer Ríos-Cuesta^{1*}

E-mail: wilmer.rios@alzate.edu.co

ORCID: <https://orcid.org/0000-0001-8129-2137>

¹Institución Educativa Gilberto Alzate Avendaño, Universidad de Antioquia, Medellín, Colombia.

*Autor para correspondencia

Cita sugerida (APA, séptima edición)

Ríos-Cuesta, W. (2026). Validación psicométrica de un cuestionario para evaluar emociones académicas durante exámenes escolares. *Revista Conrado*, 22(110), e4871.

RESUMEN:

El estudio tuvo como objetivo validar psicométricamente un cuestionario tipo Likert para identificar las emociones que experimentan estudiantes de secundaria de Medellín (Colombia) en los momentos antes, durante y después de un examen. El instrumento fue diseñado con base en la taxonomía tridimensional de Pekrun (2024) y sometido a un proceso de validación en seis fases: diseño de 37 ítems, juicio de expertos ($n=11$), cálculo de la V de Aiken, pilotaje con 51 estudiantes de los grados 9°, 10° y 11°, análisis factorial confirmatorio (AFC) y cálculo de coeficientes de fiabilidad. La V de Aiken global fue de 0,84 (IC 95%: 0,71–0,92), lo que indica alta validez de contenido. La consistencia interna resultó excelente (ω de McDonald = 0,941; α de Cronbach = 0,945). El AFC evidenció una estructura preliminar de cuatro factores coherente con el modelo teórico, con cargas factoriales significativas en todos los ítems, aunque los índices de ajuste global sugieren la necesidad de una validación confirmatoria con una muestra más amplia. El instrumento permite evaluar emociones como ansiedad, esperanza, orgullo y desesperanza, y puede contribuir a mejorar las prácticas evaluativas y el bienestar emocional en contextos educativos latinoamericanos. Se concluyó que este instrumento puede contribuir a mejorar la evaluación del estudiante, promover prácticas pedagógicas sensibles al bienestar emocional y fortalecer la investigación educativa en contextos latinoamericanos.

Palabras clave:

Evaluación del estudiante, Instrumento de medida, Psicometría, Emociones, Validación, Desarrollo emocional.

ABSTRACT:

This study aimed to psychometrically validate a Likert-type questionnaire to identify the emotions experienced by secondary school students in Medellín (Colombia) before, during, and after an examination. The instrument was designed based on Pekrun's (2024) three-dimensional taxonomy of emotions and validated through a six-phase process: item design (37 items), expert judgment ($n=11$), Aiken's V coefficient, piloting with 51 students in grades 9, 10, and 11, confirmatory factor analysis (CFA), and reliability estimation. The global Aiken's V was 0.84 (95% CI: 0.71–0.92), indicating high content validity. Internal consistency was excellent (McDonald's $\omega = 0.941$; Cronbach's $\alpha = 0.945$). The CFA revealed a preliminary four-factor structure consistent with the theoretical model, with statistically significant factor loadings for all items; however, global fit indices suggest that confirmatory validation with a larger sample is needed. The instrument enables assessment of emotions such as anxiety, hope, pride, and hopelessness, and can contribute to improving evaluative practices and emotional well-being in Latin American educational contexts.

Keywords:

Student Assessment, Measurement Instruments, Psychometrics, Emotions, Validation, Emotional Development.



INTRODUCCIÓN

Una diferencia entre evaluación y examen es el hecho de que la evaluación es un proceso sistemático y continuo que permite recoger información sobre la evolución de los aprendizajes de los estudiantes para la toma de decisiones y los exámenes son un instrumento para recoger información puntual sobre el estado de conocimiento de los estudiantes. A pesar de ello, los estudiantes conciben la evaluación como la prueba escrita que se realiza en clase para determinar qué tanto saben sobre un tema, lo que evidencia una falta de distinción entre estos dos conceptos.

Algunas emociones como la ansiedad al momento de enfrentar un examen son responsables del bajo rendimiento académico y de la pérdida de oportunidades a futuro (Zeidner, 1998). Sin embargo, se reconoce que la evaluación escolar no solo cumple una función académica, sino también emocional y motivacional. Investigaciones recientes destacan que las emociones académicas influyen directamente en la atención, la memoria y el rendimiento (Pekrun, 2024; Putwain et al., 2022).

Entre estas emociones, la ansiedad en los exámenes ha sido la más estudiada, pero hoy se reconoce un espectro más amplio tales como: gozo, alivio, esperanza, orgullo, vergüenza, ira y desesperanza (King & Chen, 2019; Pekrun et al., 2017).

En los últimos años, se han desarrollado y validado instrumentos específicos para la medición de emociones académicas. El Achievement Emotions Questionnaire (AEQ) (Pekrun et al., 2011) es el instrumento de mayor alcance en este campo: diseñado para estudiantes universitarios, evalúa emociones como el disfrute, la esperanza, el orgullo, el alivio, la ira, la ansiedad, la vergüenza y la desesperanza en tres situaciones académicas distintas (asistir a clases, estudiar y presentar exámenes). A partir del AEQ se ha derivado el Test Emotions Questionnaire (TEQ), instrumento de mayor especificidad centrado exclusivamente en el contexto de los exámenes, que constituye la base del cuestionario propuesto en este estudio. La aplicación de estas herramientas en estudios empíricos ha arrojado resultados consistentes: Villavicencio y Bernardo (2016), en un contexto asiático de educación matemática, confirmaron que las emociones positivas predicen el rendimiento académico, la autorregulación y la autoeficacia; Pekrun et al. (2017), en un estudio longitudinal de cinco años con 3.425 estudiantes alemanes de secundaria (grados 5° a 9°), demostraron que las emociones de logro y el rendimiento se influyen mutuamente a lo largo del tiempo; y Turnquest et al. (2024), en una muestra de 721 estudiantes universitarios en Estados Unidos,

hallaron que tanto la ansiedad como el disfrute predicen el promedio académico y la persistencia escolar. A nivel meta-analítico, Camacho-Morles et al. (2021) sintetizaron 68 estudios y encontraron que el disfrute se asocia positivamente con el rendimiento ($p = 0,27$), mientras que la ira y el aburrimiento lo hacen negativamente ($p = -0,35$ y $p = -0,25$, respectivamente), con efectos más pronunciados en estudiantes de secundaria que en universitarios.

De igual modo, los estudios recientes confirman que las emociones influyen de manera diferenciada en el desempeño escolar según su valencia y nivel de activación. Mientras la ansiedad (negativa-activadora) puede saturar la memoria de trabajo, la desesperanza (negativa-desactivadora) suele desvincular al estudiante de la tarea; y emociones como el orgullo, la esperanza o el alivio favorecen la motivación, la persistencia y el aprendizaje autorregulado (Camacho-Morles et al., 2021; King & Chen, 2019). Sin embargo, estos instrumentos y estudios se circunscriben mayoritariamente a contextos de educación superior o a países de Europa, Norteamérica o Asia, lo que limita su aplicabilidad directa en el contexto de la educación secundaria latinoamericana (Pekrun et al., 2011; Putwain et al., 2022; Turnquest et al., 2024; Villavicencio & Bernardo, 2016).

En Colombia, y particularmente en la ciudad de Medellín, existe escasa evidencia empírica sobre cómo experimentan las emociones los estudiantes de secundaria frente a la evaluación escolar. Este vacío es relevante porque la adolescencia es un período crítico de desarrollo emocional y porque las políticas educativas nacionales promueven la evaluación como medio para la mejora del aprendizaje, un ejemplo de ello son las pruebas saber de 3°, 5°, 7°, 9° y 11° y la prueba evaluar para avanzar que se realizan anualmente.

Explorar estas emociones aporta tanto a la comprensión del bienestar estudiantil como a la optimización de prácticas evaluativas en el aula. Por ello, este estudio presenta el proceso de construcción y validación psicométrica de un cuestionario para medir las emociones que experimentan estudiantes de secundaria en el contexto de los exámenes, mediante un diseño cuantitativo con escala tipo Likert adaptado del TEQ. El objetivo es aportar evidencia empírica local y ofrecer un instrumento de investigación que aliente la reflexión sobre prácticas evaluativas orientadas al aprendizaje y al bienestar estudiantil.

MATERIALES Y MÉTODOS

Este estudio se inscribe en el paradigma positivista, con un diseño cuantitativo de tipo instrumental orientado a la construcción y validación de un instrumento de medida

(Montero & León, 2007). Su propósito es validar un cuestionario para medir las emociones que experimentan estudiantes de secundaria en el contexto de los exámenes. Para la construcción del instrumento se tuvo en cuenta el trabajo de Pekrun (2024) sobre las emociones (Tabla 1).

Tabla 1: Taxonomía tridimensional de las emociones

Enfoque del Objeto	Positivo		Negativo	
	Activador	Desactivador	Activador	Desactivador
Logro				
Actividad	Disfrute Emoción / Entusiasmo	Relajación	Ira / Enojo Frustración	Aburrimiento
Resultado / prospectivo	Esperanza Alegría anticipatoria	Seguridad / Certeza	Ansiedad	Desesperanza
Resultado / retrospectivo	Orgullo Alegría retrospectiva Gratiitud	Contentamiento Alivio	Vergüenza Ira / Enojo	Tristeza Decepción
Epistémico				
Incongruencia de la información	Sorpresa Curiosidad Deleite	Contentamiento	Confusión Frustración	Aburrimiento
Social				
Relacionado con uno mismo	Orgullo	Satisfacción (con uno mismo)	Vergüenza Culpa	Insatisfacción (con uno mismo)
Relacionado con otros	Amor Gratiitud Admiración Compasión	Simpatía / Empatía	Odio Ira / Enojo Desprecio Envidia	Antipatía
Existencial				
Salud, vida, enfermedad, muerte	Felicidad (salud)	Alivio (recuperación)	Ansiedad (enfermedad, muerte)	Desesperanza (enfermedad, muerte)

Fuente: Pekrun (2024, p. 10)

De esta forma, emociones como la alegría, la ira o la ansiedad se consideran de alta activación, ya que implican un estado de alerta y movilización corporal; mientras que emociones como la tristeza o la calma se clasifican como de baja activación, reflejando estados más pasivos o relajados. Esta dimensión de activación es fundamental para comprender cómo las emociones influyen en el comportamiento y en los procesos cognitivos, ya que determina la intensidad con que las personas responden ante los estímulos o situaciones que enfrentan.

Fases para la construcción y validación del instrumento

Para la construcción y validación del instrumento se siguieron las siguientes fases: en primer lugar, se diseñaron 37 afirmaciones para los tres momentos (antes, durante y después del examen), se identificó la emoción principal y el tipo de activación según Pekrun (2024)En la Tabla 2 se presentan las afirmaciones.



Tabla 2: Afirmaciones, emociones y tipo de activación

Momento	Afirmación	Emoción principal	Tipo
Antes del examen	Siento pena al ver que otros terminan y yo no	Vergüenza	Negativa – Activadora
	Confío en que tendré un buen desempeño	Esperanza	Positiva – Activadora
	Durante la prueba pienso que puedo equivocarme	Ansiedad	Negativa – Activadora
	Me preocupo por no poder recordar lo estudiado	Ansiedad	Negativa – Activadora
	Durante la prueba pienso que todo mi esfuerzo ha sido inútil	Desesperanza	Negativa – Desactivadora
	Me parece que las preguntas son injustas	Enojo	Negativa – Activadora
	Me siento muy nervioso cuando estoy haciendo el examen	Ansiedad	Negativa – Activadora
	Empiezo a pensar que, por mucho que me esfuerce, nunca ganaré el examen	Desesperanza	Negativa – Desactivadora
	Me enojo si las preguntas son difíciles o enredadas para mí	Enojo	Negativa – Activadora
	Me siento más tranquilo(a) a medida que voy respondiendo las preguntas del examen	Alivio	Positiva – Desactivadora
Después del examen	Me preocupo por la nota	Ansiedad	Negativa – Activadora
	Siento satisfacción porque sé que ganaré el examen	Orgullo	Positiva – Activadora
	Confío en que la nota reflejará mi esfuerzo	Esperanza	Positiva – Activadora
	Siento fracaso aunque terminé	Desesperanza	Negativa – Desactivadora
	Al terminar el examen siento que una gran carga se quita de encima	Alivio	Positiva – Desactivadora
	Cuando termina la prueba, dejo de sentir la tensión que tenía antes	Alivio	Positiva – Desactivadora
	Después del examen pienso que no me esforcé lo suficiente	Culpa	Negativa – Activadora
	Me siento culpable por no haber estudiado más	Culpa	Negativa – Activadora
	Sigo inquieto(a) por cómo me fue en el examen y solo me calmaré al conocer los resultados	Ansiedad	Negativa – Activadora

Fuente: Elaboración propia

En segundo lugar, se envió el instrumento a valoración mediante juicio de expertos (n=11), quienes usaron una escala de 1 a 5 para valorar la pertinencia y la redacción de las afirmaciones de acuerdo con el objetivo del estudio. La formación académica y la experiencia docente de los jueces se presenta en la Tabla 3.

Tabla 3: Formación académica y experiencia docente de los jueces

Máximo grado académico alcanzado	Título del posgrado alcanzado	Años de experiencia docente
Maestría	Magíster en Ciencias de la Educación	20
Maestría	Máster en Neuropsicología Clínica	13
Maestría	Magíster en Educación Matemática	19
Maestría	Magíster en Enseñanza de las Matemáticas	15
Maestría	Magíster en Gestión de la Tecnología Educativa	23
Maestría	Magíster en Artes	2
Doctorado	Doctor en Ciencias de la Educación	25
Maestría	Maestría en Ciencias de la Educación	15
Doctorado	Doctor en Ciencias de la Educación	15
Maestría	Maestría en Educación	15
Maestría	Maestría en Educación	20

Fuente: Elaboración propia



En tercer lugar, se calculó la V de Aiken para determinar la validez de las afirmaciones y del instrumento. En cuarto lugar, se hizo un pilotaje con una muestra de 51 estudiantes de secundaria que cursan los grados 9º, 10º y 11º. En quinto lugar, se procedió con un análisis factorial confirmatorio (AFC) para determinar el modelo teórico. En sexto lugar, se evaluó la fiabilidad interna mediante el Alfa de Cronbach y el omega de McDonald, junto con el gráfico del modelo y de desajuste.

RESULTADOS

Para la validación de contenido mediante juicio de expertos se contó con la participación de once especialistas. La elección de este número se fundamenta en las recomendaciones metodológicas de la literatura, que sugieren un mínimo de cinco jueces para garantizar la validez de los resultados y un rango óptimo entre siete y diez para lograr un equilibrio entre diversidad de criterios y viabilidad operativa (Escobar-Pérez & Cuervo-Martínez, 2008; Polit & Beck, 2006). Contar con once expertos permitió asegurar la representatividad de diferentes perspectivas disciplinares y, al mismo tiempo, disponer de una base suficiente para aplicar procedimientos estadísticos. El instrumento de valoración utilizado fue de naturaleza cuantitativa (escala de 1 a 5), ya que el objetivo era obtener datos comparables y susceptibles de ser procesados mediante la V de Aiken. Esta decisión metodológica es coherente con el enfoque positivista del estudio y con el propósito de cuantificar el nivel de acuerdo entre jueces, sin que ello implique una limitación en la rigurosidad del proceso: la precisión estadística del coeficiente V de Aiken ofrece una medida objetiva y replicable de la validez de contenido (Aiken, 1985; Escobar-Pérez & Cuervo-Martínez, 2008).

Validación del instrumento

La V de Aiken mide el nivel de coincidencia de los jueces o expertos en que cada ítem de un instrumento es adecuado con respecto a un criterio (pertinencia y redacción).

$$V = \frac{\bar{x} - l}{k} \quad (1)$$

En la ecuación 1 la $\bar{x}$ corresponde a la media aritmética de las puntuaciones de los jueces para la afirmación, l es el valor mínimo posible de la escala (1 en este caso) y k es el rango de la escala (valor máximo – valor mínimo, $k=4$ en este caso).

Para determinar los límites de los intervalos de confianza de la V de Aiken usamos:

$$L_{inf} = \frac{2nkV + z^2 - z\sqrt{4nkV(1-V) + z^2}}{2(nk + z^2)} \quad (2)$$

$$L_{sup} = \frac{2nkV + z^2 + z\sqrt{4nkV(1-V) + z^2}}{2(nk + z^2)} \quad (3)$$

En las ecuaciones 2 y 3 la n es el número de jueces, la k es el rango de la escala, la V de Aiken, $z=1.96$.

En este estudio, se consideraron valores para la V de Aiken entre $0.70 \leq V \leq 0.79$ como aceptables, $0.80 \leq V \leq 0.89$ son considerados buenos y $V \geq 0.90$ se consideran excelentes. Aquellos ítem que tenían una V de Aiken alta en pertinencia y baja en redacción fueron reformulados (se marcan con asterisco) y los que tenían una pertinencia baja y redacción alta fueron eliminados ($V < 0.70$). En la Tabla 4 se presenta el ítem en su primera fila la pertinencia y en la segunda la redacción.

Tabla 4: V de Aiken para pertinencia y redacción

Momento	Ítem	Media	Desv	V de Aiken	L_inf	L_Sup
Antes del examen	Espero con entusiasmo el examen porque me permite demostrar lo que aprendí	4,00	0,85	0,75	0,61	0,85
		4,45	0,66	0,86	0,73	0,94
	Me preocupa olvidar lo que estudié	4,36	0,88	0,84	0,71	0,92
		4,55	0,66	0,89	0,76	0,95
	Me da pena pensar que puedo ser el único en perder	3,82	1,27	0,70	0,56	0,82
		4,45	0,78	0,86	0,73	0,94
	Me siento nervioso y con el estómago revuelto por el miedo a olvidar lo que estudié	4,00	1,21	0,75	0,61	0,85
		4,36	0,77	0,84	0,71	0,92
	Me pongo feliz al demostrar mis capacidades	4,64	0,64	0,91	0,79	0,96
		3,73	1,48	0,68*	0,53	0,80
	Antes del examen siento una energía positiva	4,27	1,05	0,82	0,68	0,90
		4,36	0,88	0,84	0,71	0,92
	Antes del examen me siento nervioso	4,27	0,96	0,82	0,68	0,90
		4,64	0,64	0,91	0,79	0,96
	Siento que no importa cuánto estudie, me irá mal	3,55	1,44	0,64**	0,49	0,76
		4,36	1,07	0,84	0,71	0,92
	Antes del examen siento que puedo superar con éxito esta prueba	4,55	0,66	0,89	0,76	0,95
		4,55	0,66	0,89	0,76	0,95
	Siento que no soy capaz de aprobar	3,91	1,38	0,73	0,58	0,84
		4,36	1,07	0,84	0,71	0,92
	Me siento seguro de que mi esfuerzo dará buenos resultados	4,55	0,66	0,89	0,76	0,95
		4,64	0,64	0,91	0,79	0,96
	Me preocupa si habré estudiado lo suficiente	4,09	1,16	0,77	0,63	0,87
		4,64	0,64	0,91	0,79	0,96
	Me preocupa que el examen sea demasiado difícil	4,09	1,31	0,77	0,63	0,87
		4,64	0,64	0,91	0,79	0,96
	Preferiría no presentarme al examen porque siento que no lo ganaré	3,64	1,37	0,66**	0,51	0,78
		4,00	1,28	0,75	0,61	0,85
	Me siento orgulloso de mis habilidades y de mostrarlas ante mis compañeros y profesores	4,18	0,83	0,80	0,65	0,89
		4,45	0,78	0,86	0,73	0,94

Momento	Ítem	Media	Desv	V de Aiken	L_inf	L_Sup
Durante el examen	El examen es un desafío que disfruto	4,45	0,78	0,86	0,73	0,94
		4,45	0,89	0,86	0,73	0,94
	La ansiedad hace que no me acuerde de la respuesta	4,09	1,31	0,77	0,63	0,87
		4,55	0,78	0,89	0,76	0,95
	Me irrita que hayan preguntas que no vimos en clase	4,00	1,21	0,75	0,61	0,85
		4,27	0,96	0,82	0,68	0,90
	Siento pena al ver que otros terminan y yo no	3,82	1,11	0,70	0,56	0,82
		4,27	0,86	0,82	0,68	0,90
	Confío en que tendré un buen desempeño	4,64	0,64	0,91	0,79	0,96
		4,55	0,78	0,89	0,76	0,95
	Durante la prueba pienso que puedo equivocarme	4,64	0,64	0,91	0,79	0,96
		4,64	0,77	0,91	0,79	0,96
	Me preocupo por no poder recordar lo estudiado	4,09	1,24	0,77	0,63	0,87
		4,45	0,89	0,86	0,73	0,94
	Durante la prueba pienso que todo mi esfuerzo ha sido inútil	3,64	1,23	0,66**	0,51	0,78
		4,18	0,83	0,80	0,65	0,89
	Me parece que las preguntas son injustas	3,45	1,30	0,61**	0,47	0,74
		4,18	0,83	0,80	0,65	0,89
	Me siento muy nervioso cuando estoy haciendo el examen	4,18	1,27	0,80	0,65	0,89
		4,64	0,77	0,91	0,79	0,96
	Empiezo a pensar que, por mucho que me esfuerce, nunca ganaré el examen	3,55	1,50	0,64**	0,49	0,76
		4,36	0,77	0,84	0,71	0,92
	Me enoja si las preguntas son difíciles o enredadas para mí	4,09	1,24	0,77	0,63	0,87
		4,55	0,78	0,89	0,76	0,95
	Me siento más tranquilo(a) a medida que voy respondiendo las preguntas del examen	4,64	0,64	0,91	0,79	0,96
		4,64	0,64	0,91	0,79	0,96
	Me preocupo por la nota	4,64	0,77	0,91	0,79	0,96
		4,64	0,64	0,91	0,79	0,96
	Siento satisfacción porque sé que ganaré el examen	4,45	0,78	0,86	0,73	0,94
		4,36	0,77	0,84	0,71	0,92
	Confío en que la nota reflejará mi esfuerzo	4,27	1,21	0,82	0,68	0,90
		4,36	1,15	0,84	0,71	0,92
	Siento fracaso aunque terminé	3,73	1,48	0,68**	0,53	0,80
		4,18	1,11	0,80	0,65	0,89
Después del examen	Al terminar el examen siento que una gran carga se quita de encima	4,45	0,78	0,86	0,73	0,94
		4,55	0,66	0,89	0,76	0,95
	Cuando termina la prueba, dejo de sentir la tensión que tenía antes	4,36	1,23	0,84	0,71	0,92
		4,64	0,64	0,91	0,79	0,96
	Después del examen pienso que no me esforcé lo suficiente	4,00	1,41	0,75	0,61	0,85
		4,45	0,89	0,86	0,73	0,94
	Me siento culpable por no haber estudiado más	3,82	1,40	0,70	0,56	0,82
		4,18	1,11	0,80	0,65	0,89
	Sigo inquieto(a) por cómo me fue en el examen y solo me calmaré al conocer los resultados	4,64	0,64	0,91	0,79	0,96
		4,55	0,78	0,89	0,76	0,95

Fuente: Elaboración propia (*reformulada, **eliminada)

Al determinar la media de las puntuaciones tras eliminar las afirmaciones que obtuvieron una $V < 0,70$ se obtienen los siguientes resultados (Tabla 5):

Tabla 5: Media para el instrumento

Media	D_Estandar	V de Aiken	L_inferior	L_Superior
4,36	0,91	0,84	0,71	0,92

Fuente: Elaboración propia

Los resultados presentados en la Tabla 4 evidencian que el instrumento alcanza una media global de 4,36 con una desviación estándar de 0,91 en la evaluación de sus ítems, lo cual indica un alto nivel de acuerdo entre los jueces respecto a la pertinencia y redacción de los enunciados. Asimismo, el coeficiente V de Aiken alcanza un valor de 0,84 con un intervalo de confianza entre 0,71 y 0,92, se encuentra dentro de los rangos considerados adecuados y satisfactorios para la validez de contenido según los criterios psicométricos establecidos por Aiken (1985). Estos valores sugieren que el instrumento presenta una consistencia conceptual sólida, y que las afirmaciones reflejan de manera apropiada los constructos teóricos que se pretenden medir, por lo que puede considerarse válido y adecuado para su aplicación en estudios sobre emociones.

Estudio piloto

Una vez revisado el instrumento y de haber reformulado las afirmaciones se aplicó el instrumento a 51 estudiantes de educación secundaria seleccionados mediante un muestreo intencional por accesibilidad (Figura 1). Estos estudiantes cursan los grados 9°, 10° y 11°. De ellos, el 33% han repetido algún grado académico y 67% no lo ha hecho.

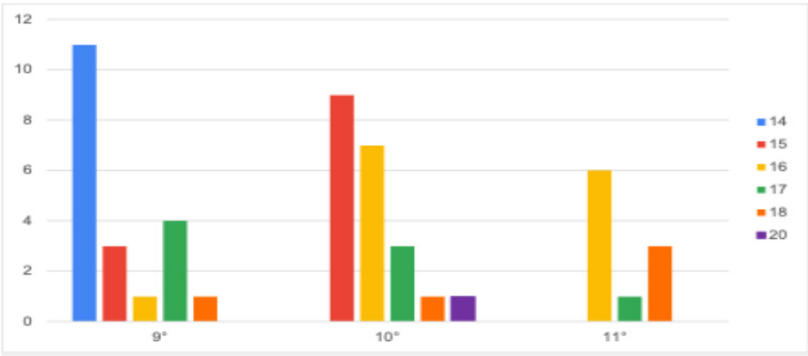


Figura 1: Distribución por grado y edad de los participantes de la prueba piloto

Fuente: Elaboración propia

En la Tabla 6 se presenta la media de las emociones que han experimentado los estudiantes participantes del pilotaje.

Tabla 6: Media y desviación estándar de las emociones de los participantes

Momento	Emoción	Media	Desviación
Antes del examen	Disfrute	2,95	1,07
	Ansiedad	2,60	1,26
	Vergüenza	2,84	1,41
	Orgullo	3,45	1,29
	Esperanza	3,25	1,08
	Desesperanza	3,70	1,08

Momento	Emoción	Media	Desviación
Durante el examen	Disfrute	3,57	1,28
	Ansiedad	2,75	1,26
	Enojo	3,05	1,22
	Vergüenza	1,83	1,40
	Esperanza	3,45	1,16
	Alivio	3,41	1,14
Después del examen	Ansiedad	2,17	1,25
	Orgullo	3,69	1,28
	Esperanza	3,55	1,26
	Desesperanza	3,33	1,31
	Alivio	3,53	1,28
	Culpa	3,14	1,23

En la Tabla 6 se observa que, antes del examen hay una ligera inclinación hacia la esperanza y el orgullo, sin embargo, la vergüenza y la ansiedad aparecen con valores medios lo que indica cierta tensión anticipatoria antes de la prueba, también se observa que la desesperanza aparece con la media más alta, lo cual indica dudas sobre el desempeño de los participantes. Durante el examen hay un incremento en el disfrute y la esperanza así como un descenso de la vergüenza, lo que sugiere un mejor ajuste emocional durante la ejecución de la tarea. Hay un ligero aumento de la ansiedad y emociones como el alivio y el enojo permanecen en niveles medios. Finalmente, después del examen, se observa una disminución leve de la ansiedad y un aumento del orgullo y el alivio, esto refleja una recuperación emocional y sensación de logro. La culpa se mantiene en un nivel medio lo cual puede reflejar la autocrítica o la sensación de poder haberlo hecho mejor.

Lo anterior sirve como evidencia de que la aplicación del cuestionario permite rastrear las emociones de los estudiantes en los tres momentos (antes, durante y después del examen). Además, se observan variaciones de las emociones en los tres momentos lo cual es un indicio de que se puede hacer una trazabilidad de ellas.

Análisis Factorial Confirmatorio (AFC)

En este análisis se usó el modelo de emociones de Pekrun (2024)control-value theory describes and explains achievement emotions. More recently, the theory has been expanded to also explain epistemic, social, and existential emotions. In this article, I outline the development of the theory, from preliminary work in the 1980s to early versions of the theory and the recent generalized control-value theory. I provide summaries of the theory's evidence-based propositions on antecedents, outcomes, and regulation of emotions, including the fundamentally important role of control and value appraisals across different types of human emotions that are relevant to education (and beyond teniendo en cuenta su clasificación en: 1) Positivas activadoras, 2) Positivas desactivadoras, 3) Negativas activadoras y 4) Negativas desactivadoras. Estas fueron utilizadas para establecer los factores. Además, se usó el software JASP (Versión 0.95.3) para dicho análisis.

Fiabilidad del instrumento

La escala utilizada para evaluar las emociones frecuentes presentó adecuados niveles de consistencia interna. El coeficiente ω de McDonald y el α de Cronbach evidencian una confiabilidad excelente según los criterios psicométricos establecidos (George & Mallery, 2019). Estos resultados indican que los ítems del instrumento son altamente homogéneos y coherentes entre sí, reflejando una medición estable y precisa del constructo emocional evaluado. En consecuencia, el instrumento puede considerarse válido y confiable para la aplicación en contextos académicos e investigativos relacionados con el estudio de las emociones (Tabla 7).



Tabla 7: Confiabilidad del instrumento

Coeficiente	Estimar	Error típico	IC del 95%	
			Lower	Upper
Coefficient ω	0.941	0.012	0.917	0.964
Coefficient α	0.945	0.018	0.909	0.981

Fuente: Elaboración propia con el software JASP

Gráfico del modelo

Los resultados del AFC arrojaron los siguientes índices de ajuste global: $\chi^2(554) = 1075,64$, $p < 0,001$; CFI = 0,68; TLI = 0,66; RMSEA = 0,135; GFI = 0,52; NFI = 0,52. Estos valores se encuentran por debajo de los umbrales convencionales aceptables (CFI y TLI $\geq 0,90$; RMSEA $\leq 0,08$), lo que indica un ajuste global insatisfactorio del modelo. No obstante, este resultado debe interpretarse en el contexto de las limitaciones inherentes a la prueba piloto: la muestra de 51 estudiantes resulta insuficiente para la estabilidad estadística de un AFC con 37 ítems y cuatro factores, dado que la literatura psicométrica recomienda entre 5 y 10 observaciones por ítem (Hair et al., 2019), lo que exigiría entre 185 y 370 participantes para este instrumento. A pesar de lo anterior, las cargas factoriales de todos los ítems resultaron estadísticamente significativas ($p < 0,001$), lo que indica que cada ítem guarda una relación coherente con su factor teórico asignado, y las covarianzas entre factores son consistentes con la estructura de la taxonomía de Pekrun (2024). En consecuencia, los resultados del AFC deben considerarse como evidencia preliminar de la estructura factorial del instrumento, sujeta a confirmación con una muestra de mayor tamaño (Figura 2).

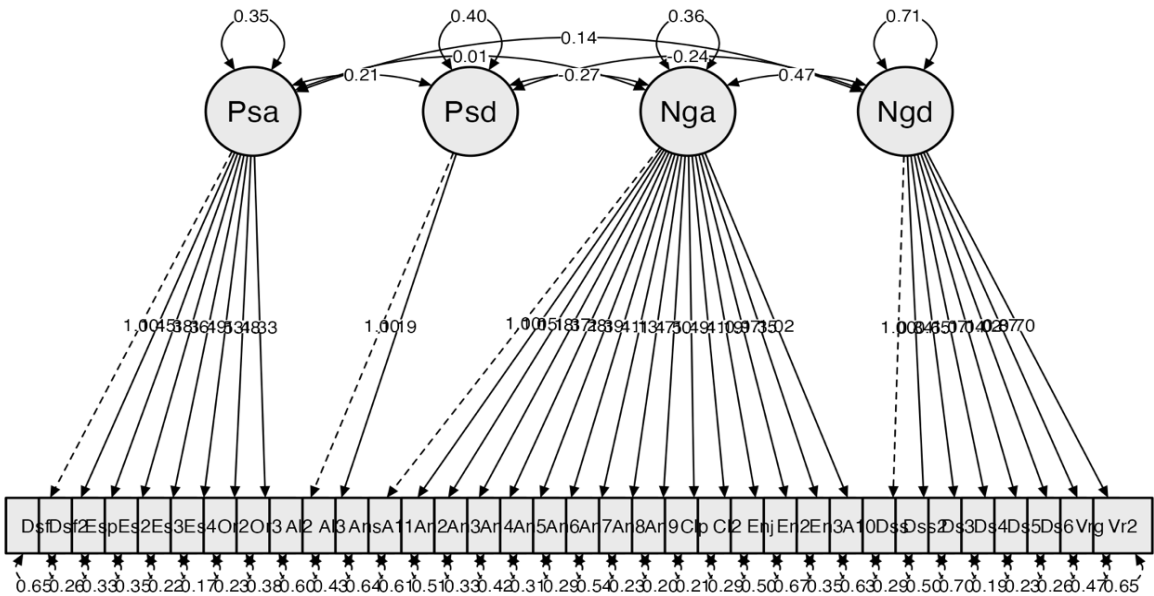


Figura 2: Gráfico del modelo

Fuente: Elaboración propia con el software JASP

Gráfico de desajuste

El gráfico muestra los residuos estandarizados entre las covarianzas observadas y las estimadas por el modelo factorial. En general, los residuos se mantienen en valores cercanos a cero (representados por tonos claros en la matriz), lo que sugiere un ajuste local adecuado entre los ítems y las dimensiones teóricas de las emociones (Figura 3).

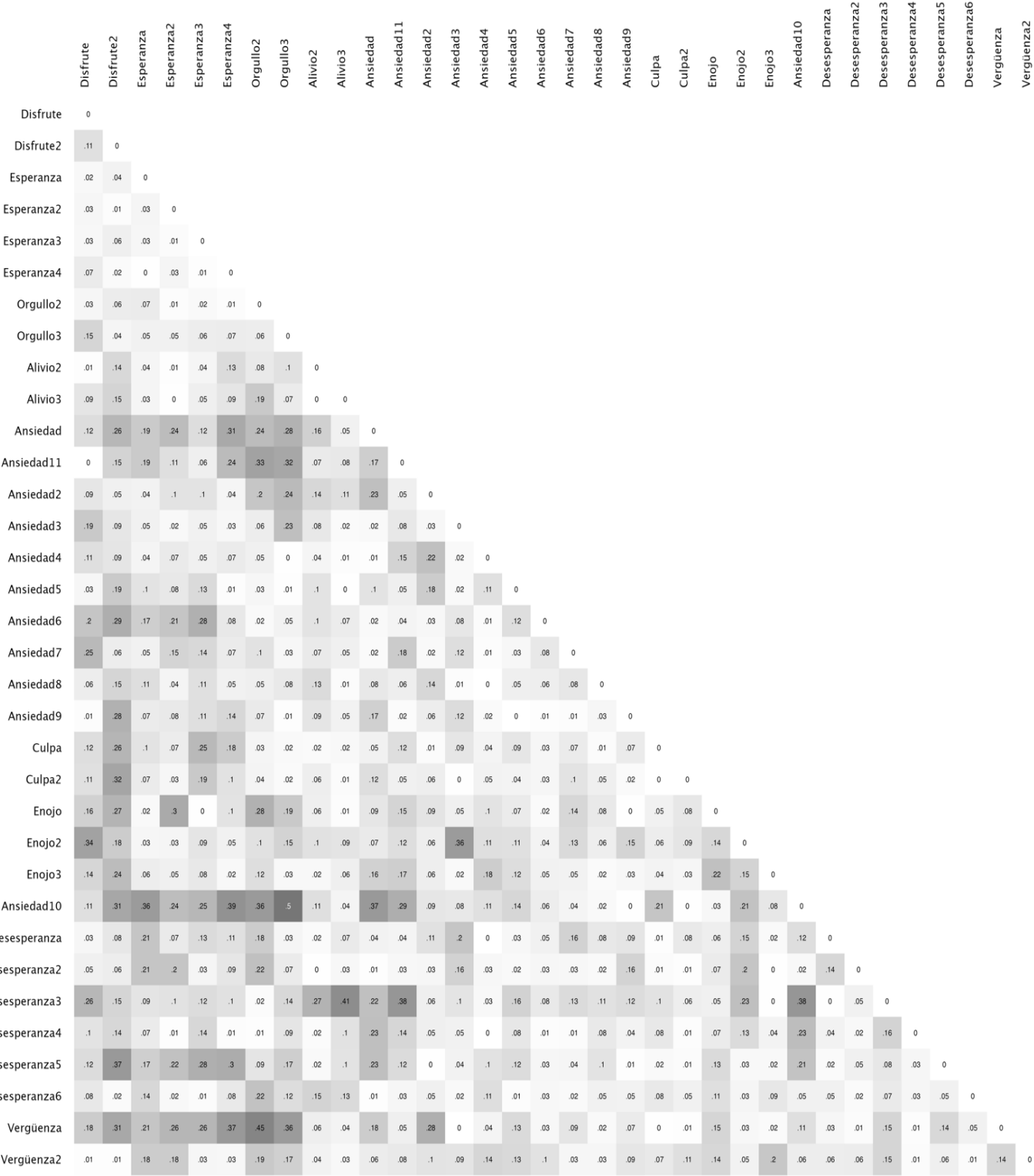


Figura 3: Modelo desajustado
Fuente: Elaboración propia software JASP

Se observan algunos desajustes puntuales de mayor intensidad (tonos más oscuros) entre ciertos ítems, esto podría indicar redundancia semántica o solapamiento de contenido entre ellos. Este patrón es común cuando varios reactivos evalúan manifestaciones similares de una misma emoción.

En el caso de las emociones negativas desactivadoras, también se aprecian residuos moderados, lo que sugiere una alta intercorrelación interna entre estos ítems, posiblemente por tratar aspectos cercanos de la misma experiencia emocional.

No se observan desajustes significativos entre ítems de factores distintos, lo que respalda la discriminación adecuada entre emociones positivas y negativas, y entre las dimensiones activadoras y desactivadoras.

DISCUSIÓN

Los análisis psicométricos demostraron una alta validez de contenido, avalada por un coeficiente V de Aiken global de 0,84, superando los umbrales estandarizados para la concordancia entre jueces (Aiken, 1985). El Análisis Factorial Confirmatorio confirmó que todos los ítems presentan cargas factoriales estadísticamente significativas sobre sus respectivos factores ($p < 0,001$), y que la organización de los ítems es coherente con la clasificación en función de la valencia (positiva/negativa) y la activación (activadora/desactivadora), replicando los modelos hallados en instrumentos de educación superior como el Achievement Emotions Questionnaire (AEQ) (Pekrun et al., 2011). Sin embargo, los índices de ajuste global del modelo (CFI = 0,68; RMSEA = 0,135) no alcanzaron los umbrales convencionales, lo cual se atribuye principalmente al tamaño reducido de la muestra piloto ($n = 51$), insuficiente para la complejidad del modelo de 37 ítems. Esta limitación es característica de los estudios piloto de validación y no invalida la estructura teórica del instrumento, pero subraya la necesidad de una validación confirmatoria con una muestra ampliada (Hair et al., 2019). La consistencia estructural disponible, sumada a coeficientes de confiabilidad (Alfa de Cronbach y Omega de McDonald) superiores a 0,85 (George & Mallery, 2019), posiciona al instrumento con niveles de precisión excelentes, garantizando la homogeneidad de las subescalas para la medición del constructo en población adolescente.

Un hallazgo distintivo de esta investigación radica en la sensibilidad del instrumento para capturar la variabilidad temporal y la dinámica de activación emocional. Los resultados del estudio piloto evidencian fluctuaciones significativas entre los momentos previos, durante y posteriores al examen, coincidiendo con investigaciones recientes que describen la tensión anticipatoria y la posterior recuperación emocional (Putwain et al., 2022). La capacidad del cuestionario para discriminar entre emociones de valencia negativa, pero distinta activación resulta fundamental, dado que estos estados tienen impactos divergentes en el procesamiento cognitivo y el desempeño, en consonancia con el modelo teórico de Pekrun (2024). Mientras la ansiedad (activación) puede saturar la memoria de trabajo, la desesperanza (desactivación) suele desvincular al estudiante de la tarea, lo que subraya la importancia teórica y práctica de no tratar las emociones negativas como un bloque monolítico.

Finalmente, la validación de este instrumento responde a una necesidad concreta en el contexto educativo latinoamericano: la escasez de herramientas psicométricas culturalmente ajustadas para la educación secundaria. A diferencia de las escalas diseñadas para población universitaria o de adultos, este cuestionario fue construido y pilotado directamente con estudiantes de secundaria en Colombia, lo que refuerza su pertinencia contextual. La identificación precisa de perfiles emocionales en tres momentos del examen permite diseñar intervenciones pedagógicas más específicas y eficaces: estrategias para reducir la ansiedad anticipatoria, promover la esperanza durante la tarea y capitalizar el orgullo y el alivio como refuerzo posterior. Esto posiciona al instrumento no solo como un recurso de investigación, sino como una herramienta práctica para docentes y orientadores escolares, en consonancia con las prioridades educativas contemporáneas orientadas al bienestar estudiantil (King & Chen, 2019).

CONCLUSIONES

Al aplicar el cuestionario se puede identificar el perfil emocional de los estudiantes por grado y rango etario, lo que permitiría determinar si las emociones evolucionan a medida que avanzan en la escolaridad y si la edad predice determinadas respuestas emocionales ante los exámenes. Asimismo, un diseño longitudinal con este instrumento podría aportar evidencia sobre tendencias de cambio emocional a lo largo del ciclo escolar y sobre el efecto acumulado de las prácticas evaluativas en el bienestar estudiantil.

El instrumento mostró una V de Aiken global de 0,84, lo que indica un alto nivel de acuerdo entre los expertos en cuanto a la pertinencia y redacción de los ítems. Esto respalda su solidez conceptual y su capacidad para representar adecuadamente las emociones académicas en contextos evaluativos. Este proceso incluyó el juicio de expertos, un análisis factorial confirmatorio y el cálculo del coeficiente de fiabilidad (Alfa de Cronbach y Omega de McDonald), lo que garantiza la robustez del instrumento desde una perspectiva metodológica. No obstante, los índices de ajuste del AFC (CFI = 0,68; RMSEA = 0,135) evidencian la necesidad de una validación confirmatoria con una muestra de mayor tamaño ($n \geq 185$), como paso siguiente para consolidar la evidencia de validez estructural del instrumento (Hair et al., 2019).

Por otro lado, este instrumento permite medir emociones activadoras y desactivadoras, tanto positivas como negativas, en tres momentos del examen (antes, durante y después), incluyendo ansiedad, esperanza, orgullo, vergüenza, alivio y desesperanza. Además, puede ser utilizado por profesores, orientadores y equipos de investigación para comprender el impacto emocional de las prácticas evaluativas, promover el bienestar estudiantil y diseñar estrategias pedagógicas más sensibles y eficaces.

Finalmente, este trabajo enriquece el campo de la psicometría educativa al ofrecer un instrumento validado que integra dimensiones emocionales en la evaluación del estudiante, para el desarrollo de investigaciones sobre el desarrollo emocional en contextos escolares. Además, se puede revisar cómo las secuencias de aprendizaje y las tareas de los libros de textos pueden desencadenar emociones que incidan en el rendimiento escolar.

REFERENCIAS BIBLIOGRÁFICAS

- Aiken, L. R. (1985). Three Coefficients for Analyzing the Reliability and Validity of Ratings. *Educational and Psychological Measurement*, 45(1), 131-142. <https://doi.org/10.1177/0013164485451012>
- Camacho-Morles, J., Slemp, G. R., Pekrun, R., Loderer, K., Hou, H., & Oades, L. G. (2021). Activity Achievement Emotions and Academic Performance: A Meta-analysis. *Educational Psychology Review*, 33(3), 1051-1095. <https://doi.org/10.1007/s10648-020-09585-3>
- Escobar-Pérez, J., & Cuervo-Martínez, Á. (2008). Validez de contenido y juicio de expertos: Una aproximación a su utilización. *Avances en Medición*, 6, 27-26.
- George, D., & Mallery, P. (2019). *IBM SPSS Statistics 26 Step by Step: A Simple Guide and Reference* (6.ª ed.). Routledge. <https://doi.org/10.4324/9780429056765>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis* (8.ª ed.). Cengage Learning.
- King, R. B., & Chen, J. (2019). Emotions in Education: Asian Insights on the Role of Emotions in Learning and Teaching. *The Asia-Pacific Education Researcher*, 28(4), 279-281. <https://doi.org/10.1007/s40299-019-00469-x>
- Pekrun, R. (2024). Control-Value Theory: From Achievement Emotion to a General Theory of Human Emotions. *Educational Psychology Review*, 36(3), 83. <https://doi.org/10.1007/s10648-024-09909-7>
- Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology*, 36(1), 36-48. <https://doi.org/10.1016/j.cedpsych.2010.10.002>
- Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement Emotions and Academic Performance: Longitudinal Models of Reciprocal Effects. *Child Development*, 88(5), 1653-1670. <https://doi.org/10.1111/cdev.12704>
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? critique and recommendations. *Research in Nursing & Health*, 29(5), 489-497. <https://doi.org/10.1002/nur.20147>
- Putwain, D. W., Wood, P., & Pekrun, R. (2022). Achievement emotions and academic achievement: Reciprocal relations and the moderating influence of academic buoyancy. *Journal of Educational Psychology*, 114(1), 108-126. <https://doi.org/10.1037/edu0000637>
- Turnquest, K. N., Fan, W., Snodgrass Rangel, V., Dyer, N., & Master, A. (2024). Achievement emotions predict transfer student academic success. *Social Psychology of Education*, 27(4), 1481-1508. <https://doi.org/10.1007/s11218-023-09858-z>
- Villavicencio, F. T., & Bernardo, A. B. I. (2016). Beyond Math Anxiety: Positive Emotions Predict Mathematics Achievement, Self-Regulation, and Self-Efficacy. *The Asia-Pacific Education Researcher*, 25(3), 415-422. <https://doi.org/10.1007/s40299-015-0251-4>
- Zeidner, M. (1998). *Test Anxiety: The State of the Art*. Springer. <https://doi.org/10.1007/b109548>
- Montero, I., & León, O. G. (2007). A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology*, 7(3), 847-862.

CONFLICTO DE INTERESES:

El autor declara que no mantiene relaciones financieras ni personales con entidades públicas o privadas que pudieran influir de manera inapropiada en los resultados o en la interpretación del presente trabajo. Asimismo, manifiesta que no existen conflictos de interés de carácter académico, institucional, ideológico, político o religioso relacionados con el contenido del manuscrito.

Autor	Roles
Autor 1	Encargado de: Conceptualización, Investigación, Metodología, Software, Análisis formal, Validación, Escritura – borrador original, Redacción – revisión y edición.

Contribución de los autores

La Revista Conrado publica sus artículos bajo una licencia Creative Commons: Atribución/Reconocimiento-NoComercial-SinDerivados 4.0 Internacional

<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>

