

LÍNEA DE INVESTIGACIÓN  
E INTEGRACIÓN CURRICULAR  
EN PSICOLOGÍA EDUCATIVA

# 3

## CALIDAD TÉCNICA DEL INSTRUMENTO DE EVALUACIÓN DE LOS PROYECTOS DE INVESTIGACIÓN Y TESIS DE LA MAESTRÍA

Carlos A. Pardo Adames\*

Vanessa Sánchez Mendoza

Universidad Católica de Colombia, Bogotá, Colombia.

### Introducción

El trabajo de grado de maestría en Psicología de la Universidad Católica de Colombia es un requisito central en el proceso de formación del estudiante, dado que se trata de un programa con énfasis investigativo, cuya pretensión, además de fortalecer habilidades en lo disciplinar, atiende a las orientaciones del Decreto 1295 de 2010, por cuanto busca

“[...] procurar el desarrollo de competencias científicas y una formación avanzada en investigación o creación que genere nuevos conocimientos, procesos tecnológicos u obras o interpretaciones artísticas de interés cultural según el caso [...] el trabajo de grado debe evidenciar las competencias científicas, disciplinares o creativas propias del investigador, del creador o del intérprete artístico” (Ministerio de Educación Nacional [MEN], 2010, p. 12).

En este sentido, el Centro de Estudios e Investigaciones en Psicología (CEIPS) de la Facultad de Psicología de la Universidad Católica de Colombia ha realizado jornadas de comité de investigaciones desde el año 2010 orientadas a la generación de un instrumento estandarizado que permita obtener una evaluación objetiva de los evaluadores de los proyectos. A lo largo de estos años se han generado tres versiones del formato de evaluación: la primera incluía criterios de calidad y solicitaba una evaluación cualitativa de la percepción del par evaluador; la segunda, elaborada en 2013, contenía los mismos aspectos que la anterior, pero en ella había la posibilidad

\*capardo@ucatolica.edu.co

de emitir un concepto en el que el jurado decidía si el proyecto quedaba pendiente de aprobación hasta que se verificara la completitud de las sugerencias o modificaciones; y, finalmente, la versión actual, vigente desde el año 2015, busca implementar un sistema de calificación de los criterios asociados a los proyectos de grado, utilizando una escala numérica con referencia en el formato de informe técnico de avance de proyectos implementado por Colciencias<sup>1</sup>.

Los estudiantes de la maestría en Psicología cursan la asignatura de trabajo de grado en cuatro semestres, en los cuales se adelantan actividades de acompañamiento, evaluación y arbitraje de sus proyectos. La elección de los temas de investigación está direccionada por las líneas de investigación e integración curricular, las cuales tienen como objetivo estudiar los procesos psicosociales y del comportamiento en los contextos educativo, organizacional, jurídico, comunitario y político (Centro de Estudios e Investigaciones en Psicología, 2014). La organización de las actividades del seminario de tesis es la siguiente: en el primer semestre se desarrolla el coloquio de investigadores, en donde los directores de tesis, los líderes de líneas y los grupos de investigación escuchan y evalúan la propuesta del tesista para emitir su aval; en el segundo semestre, el documento se envía a pares externos y se remite el formato de evaluación para su calificación; durante el tercer semestre, el estudiante realiza la recolección de datos e información; y en el cuarto semestre, el estudiante presenta un informe que se envía a pares externos para evaluación, y, si es aprobado, realiza la sustentación del mismo ante el jurado.

La finalidad de la revisión por pares mencionada apunta a los siguientes objetivos: (a) facilitar la detección y reparación de problemas a partir de la lectura de terceros, (b) buscar la generación y fortalecimiento de competencias asociadas a la presentación de documentos ante una audiencia, (c) contribuir a la identificación de criterios para la revisión de los documentos y (d) propiciar el diálogo entre pares para el fortalecimiento de la comunidad de investigadores (Carlino, 2008; Hotsmanshof & Conrad, 2003).

Sin embargo, la eficacia en el proceso de revisión de los proyectos de maestría ha sido una de las dificultades que con mayor frecuencia impacta en la culminación exitosa de este nivel de estudios, puesto que ocasiona el abandono tardío –entendido como el retiro del programa sin el cumplimiento del requisito de tesis para la graduación– (Frésan, 2013), y afecta directamente los indicadores de verificación de la calidad del programa por cuanto no se evidencia la capacidad institucional para egresar titulados a la sociedad (Gómez & Sánchez, 2013).

<sup>1</sup> Colciencias es el Departamento Administrativo de Ciencia, Tecnología e Innovación que depende de la Presidencia de la República de Colombia y lidera el Sistema Nacional de Ciencia, Tecnología e Innovación del país. La guía para la presentación de informes técnicos se encuentra disponible en el enlace [http://www.colciencias.gov.co/sites/default/files/ckeditor\\_files/M301PR03G01-guia-presentacion%20-informes-tecnicos-ctei.pdf](http://www.colciencias.gov.co/sites/default/files/ckeditor_files/M301PR03G01-guia-presentacion%20-informes-tecnicos-ctei.pdf)

Al respecto, algunos autores han realizado trabajos que permiten reconocer los aspectos asociados con la deserción de los estudiantes en este nivel; sin embargo, el número de publicaciones sobre el tema es menor en comparación con aquellos que se encuentran asociados a la educación de pregrado. Por ejemplo, en el ámbito latinoamericano, Hernández, Pérez y González (2014) realizan un recorrido por algunos de los referentes sobre el tema, presentan el trabajo de Haro y Valverde (2002), desarrollado en Argentina, en el que plantean que un factor central en la deserción en los posgrados está asociado a la cantidad de trabajo y tiempo que requiere invertir el estudiante adicional a sus compromisos laborales y familiares; muestran el trabajo de Dubs (2005), realizado en Venezuela, en el que se hace explícita la influencia de la integración social y académica en el fortalecimiento de las competencias del estudiante, así como su influencia en la decisión de permanecer o no en el programa de estudio; y, en Costa Rica, aluden al trabajo de Barrientos y Umaña (2009), quienes abordan el problema de la deserción en tres sentidos: cualificación docente, oportunidad laboral y relación con la ocupación del estudiante. Finalmente, estos tres autores describen a los factores sociológicos y socioeconómicos como asociados a la deserción de estudiantes de posgrados, sin hacer evidente una tendencia asociada a aspectos de género, nivel académico de los padres ni rendimiento escolar de los exalumnos antes o durante el estudio del posgrado.

Por otra parte, en un estudio realizado por Barahona (2016) se muestra que los estudiantes que desertaron del programa de maestría en Docencia Universitaria aludieron a los siguientes factores, en orden de importancia, como causales de deserción: (a) problemas de salud o emocionales, (b) problemas económicos, (c) aspectos académicos y (d) aspectos institucionales. Específicamente, en los aspectos académicos el de mayor peso es el de la tesis de grado, de donde cerca del 35 % de los estudiantes deserta sin realizar su investigación o trabajo de grado.

Particularmente en Colombia, un estudio del Ministerio de Educación Nacional (MEN, 2015b) indica que en el periodo de 2001 a 2014 en Bogotá se graduaron de maestría 1710 personas, de los cuales 166 estudiaron algún programa de Ciencias Sociales y Humanas, y, de estos, cuatro cursaron programas de Psicología. Ese mismo estudio muestra que en Colombia, para el periodo de 2014, se matricularon en maestría 52 608 estudiantes, de los cuales, luego de cuatro semestres de observación, el 51.8 % de los estudiantes desertó –el 57.93 % perteneciente a programas de Ciencias Sociales y Humanas–. Específicamente, los mayores niveles de deserción se encuentran en las especializaciones, donde 53 de cada 100 estudiantes desertan, seguidos por las maestrías, con 47 estudiantes, y los doctorados, con 27.

Según el MEN (2015a), “se reconoce un salto importante en el paso de tercer a cuarto semestre, donde la deserción aumenta en 19 puntos, lo que puede estar asociado al

tema de trabajo de grado” (p. 2); y, en detalle, en el escenario del programa de maestría en Psicología de la Universidad Católica, durante el periodo de 2015 el 46% de proyectos quedaron “pendientes de aprobación” en la primera evaluación de pares y el 12% quedaron “no aprobados”; mientras que en el 2016, 32% quedaron “pendientes de aprobación” y 8%, “no aprobados”, aspecto que genera inquietud por cuanto no se ha establecido el efecto de las categorías en la evaluación de las tesis ni la pertinencia del instrumento para llevar a cabo la misma.

A este respecto, Sierra (2011) realizó un estudio cualitativo que tuvo como objetivo conocer las percepciones de estudiantes de maestría, magísteres y directores de tesis sobre los retos y dificultades asociados a la elaboración de la tesis para optar por la titulación, y los resultados mostraron que delimitar el tema, elaborar el estado de la cuestión y el marco metodológico, así como la elección del diseño son los obstáculos, identificados por los directores, que con mayor frecuencia se encuentran en la etapa inicial de tesis, mientras que redactar el informe final es el obstáculo que con mayor frecuencia identifican los estudiantes, sobre todo por recibir revisiones negativas por parte de los pares o no cumplir sus expectativas.

Lo que en palabras de Moreno, Cortés y Zamuya (2012), enmarca el proceso de evaluación por pares en una perspectiva cualitativa que parte de las experiencias y expectativas de los evaluadores acerca de la calidad del documento y de los indicadores de la misma, donde el concepto de calidad puede tener tres acepciones: primero, la *calidad cognitiva*, relacionada con la importancia del contenido específico de las ideas científicas; segundo, la *calidad metodológica*, que se cumple cuando el autor utiliza esquemas conceptuales eficaces ya comprobados; y, tercero, la *calidad estética*, que alude a la armonía, limpieza, exactitud y originalidad implícitas en las propuestas científicas (Sancho, 1990 citado en Moreno et al., 2012).

Teniendo como referente este contexto, surge la necesidad de establecer el peso que tiene en la calificación final cada categoría, de forma tal que se evidencien las principales dificultades o aspectos a los que deberán prestarle atención los investigadores, con el fin disminuir el desempeño de quienes presentan los proyectos de grado y mejorar el número de rechazos de las propuestas. Así, conocer el tamaño del efecto de cada aspecto en la puntuación final de las calificaciones permitirá identificar a qué aspecto le prestan más atención los evaluadores, para que los investigadores realicen un análisis más profundo de los mismos.

Con esta investigación se pretende responder al interrogante ¿cuál es el tamaño del efecto de las categorías en la evaluación de las tesis de maestría?, y se busca alcanzar el objetivo de establecer el tamaño del efecto de las categorías en la evaluación de tesis de maestría en Psicología, de modo que sea posible identificar

la calidad técnica del formato de evaluación y determinar la relación de los componentes del instrumento con las puntuaciones globales obtenidas.

## Materiales y método

### Datos

El presente trabajo se realizó con un total de 57 evaluaciones realizadas por los pares. Debido a que la modalidad empleada por el CEIPS implica dos evaluaciones para cada proyecto, en algunos casos dos pares evaluaron el mismo documento o proyecto. La evaluación se realizó sobre tres categorías: antecedentes teóricos y empíricos (7 ítems), justificación y planteamiento del problema de investigación (3 ítems) y método (7 ítems).

### Instrumentos

El formato de evaluación describe los criterios para la evaluación de los proyectos, tiene como finalidad realizar la verificación de las competencias investigativas del estudiante en el diseño y presentación del trabajo de grado, y recibe una puntuación diferencial. Los criterios de calificación son:

- a. *Título* (5 puntos): se establecen como indicadores de calidad que el título sea descriptivo; que los términos que lo componen faciliten su identificación junto a trabajos similares en bibliografías e índices que permitan búsquedas por palabras clave; y que sea conciso, claro y preciso para el lector. Al evaluador se le plantean interrogantes que orientan su revisión, como, por ejemplo: ¿es coherente con el problema, el objetivo general y la(s) hipótesis/supuestos de trabajo?
- b. *Resumen* (5 puntos): deberá contener máximo 500 palabras y la información necesaria para darle al lector una idea precisa de la pertinencia y calidad del estudio llevado a cabo, así como una síntesis del objetivo general, la descripción de la metodología (diseño, muestra, instrumentos, –técnicas de recolección de datos– y análisis de datos) y las palabras clave.
- c. *Antecedentes teóricos y empíricos* (25 puntos): en este apartado se evalúa si el documento contiene una síntesis del contexto teórico general en el cual se ubica el tema del proyecto y el estado actual del conocimiento del problema (nacional y mundial), así como las brechas que existen y el vacío que se quiere llenar con el estudio. En este sentido, se formula a los evaluadores algunas preguntas que orientan su revisión, como, por ejemplo: ¿la revisión responde a preguntas sobre por qué y cómo la investigación desarrollada, con fundamento en investigaciones

previas, contribuye con probabilidades de éxito, a la solución o comprensión del problema planteado o al avance del sector interesado? Acá, el evaluador debe privilegiar la utilización de las fuentes primarias, como artículos de revistas indexadas impresas o provenientes de bases de datos especializadas y textos actuales que sirvan como referente empírico. De igual forma, deberán estar citados los autores que se encuentran en el apartado de referencias y se espera que el contenido del marco teórico se haya desarrollado completamente y se hayan asumido las recomendaciones que hubieran surgido de la revisión previa.

- d. *Justificación y planteamiento del problema de investigación* (30 puntos): en este criterio se busca evaluar la precisión y la claridad con que se ha formulado el problema de investigación y en qué medida se evidencia la relación entre el problema de investigación y los conocimientos existentes sobre el tema, además de su importancia y las razones que le justifican. En este sentido, se remite a los evaluadores a responder interrogantes orientadores como, por ejemplo: ¿se presenta una descripción precisa y completa de la naturaleza y magnitud del problema?, ¿se justifica la necesidad de la investigación en función de su pertinencia social y disciplinar?, ¿brinda argumentos convincentes de que los conocimientos disponibles son insuficientes para dar cuenta del problema?, o ¿da un fundamento convincente de la necesidad de someter a prueba si lo que se conoce o se da como un hecho verdadero, puede debatirse, a partir de nuevos hallazgos o nuevas situaciones? Acá, se debe tener en cuenta que la descripción del problema refleja que el investigador se ha documentado y ha realizado una exhaustiva revisión bibliográfica sobre el tema. Igualmente, para la justificación se proponen algunos interrogantes al evaluador: ¿cuál es el aporte del estudio a la generación de conocimiento sobre el tema?, ¿cuál es el vacío o vacíos en el conocimiento, identificados a partir de la revisión teórica y empírica, que ha abordado esta investigación?, y ¿cuál es la importancia científica y tecnológica del estudio en la producción de conocimiento? Al final de esta reflexión se debe formular claramente la pregunta concreta que se quiere responder en el contexto del problema a cuya solución o entendimiento se ha contribuido con la ejecución del estudio.
- e. *Objetivos* (5 puntos): en esta categoría se describe la necesidad de formular objetivos claros y viables, en donde no se debe confundir objetivos con actividades o procedimientos metodológicos; se plantee un solo objetivo general, coherente con el problema; y se presenten tantos objetivos específicos como sean necesarios para lograr el objetivo general. La evaluación de este aparte se orienta a través de interrogantes como: ¿muestran una relación clara y consistente con la descripción del problema y, específicamente, con la pregunta o hipótesis que se quieren resolver?, ¿muestra una relación clara con los resultados y discusión?, y ¿cuál es el conocimiento generado con la realización del estudio?

- f. *Método* (25 puntos): este aspecto ha contado con la asignación de mayor relevancia debido al carácter investigativo de la maestría en Psicología; y los aspectos evaluados incluyen:
- La descripción y justificación del *tipo de investigación* y el diseño que se utilizó, los cuales deben ser coherentes con los objetivos propuestos y con la muestra seleccionada.
  - Las *variables* o *categorías* (según la naturaleza del estudio) propuestas, tipología, definición, referentes para su elección, entre otros aspectos.
  - Las *hipótesis* o supuestos de trabajo, que deben ser coherentes con los objetivos, las variables propuestas y el marco teórico.
  - Los *participantes*, en donde se debe describir y argumentar la cantidad y las características de los participantes que efectivamente hicieron parte del estudio, el tipo de selección utilizada y las formas de asignación a las condiciones de investigación.
  - Los *instrumentos*, en donde se incluyen escalas, cuestionarios, listas de chequeo, protocolos, matrices de análisis y cualquier otra estrategia usada en el estudio para recolectar información. Por cada uno se solicita la descripción de los autores, el número de ítems, el tipo de ítems, la forma de calificación y los indicadores de confiabilidad y validez según el caso. Adicionalmente, deberá incluirse en el apéndice una copia del instrumento utilizado. En caso de que la investigación sea de corte cualitativo, se solicita describir a los investigadores –quiénes son y cuántos son–. Además, se solicita la descripción y argumentación de las técnicas de recolección de datos utilizadas (entrevista en profundidad –individual o grupal–, observación participante, conversación informal o revisión documental).
  - El *procedimiento*, el cual deberá tener coherencia con el diseño planteado y describir en forma detallada (preferiblemente, por fases) cuándo, cómo, quién y dónde se desarrolló la recolección de datos o los procesos de triangulación, entre otras estrategias. Las acciones llevadas a cabo en la investigación deberán describirse en su totalidad. Igualmente, se pide la revisión de la inclusión de aspectos éticos considerados en la legislación colombiana para el oficio como anonimato, confidencialidad y consentimiento informado, así como las formas en que se garantizó el bienestar de los participantes (seres humanos o animales) en la investigación. Además, debe incluirse como apéndice el tipo de consentimiento informado utilizado.
- g. *Aspectos formales* (5 puntos);, en donde se incluye la revisión del estilo de presentación (APA), referencias y apéndices.

Estas categorías evalúan la precisión y claridad con que se ha formulado la investigación y, además, en qué medida se evidencia el establecimiento de relaciones entre el problema de investigación y los conocimientos existentes sobre el tema, su importancia y las razones que le justifican al responder a cuestionamientos sobre el aporte de la propuesta a la generación de conocimiento, los vacíos que se pretenden suplir luego de la revisión empírica y teórica, la inclusión de aspectos éticos como el anonimato, la confidencialidad, el consentimiento informado, las formas de garantizar el bienestar de los participantes (seres humanos o animales) en la investigación y los resultados de la actividad investigativa desarrollada. Estos criterios se califican con la siguiente escala:

Tabla 3. Escala de evaluación de los criterios del formato

Calificación	Criterio	Descripción
5	Satisfactorio	El indicador ha sido abordado en profundidad y no requiere ajustes.
4	Adecuado	El indicador ha sido abordado, pero requiere algunos ajustes.
3	Aceptable	El indicador ha sido abordado con inexactitud y debe ser modificado.
2	Inadecuado	El indicador ha sido abordado parcialmente y requiere ajustes de fondo.
1	Insuficiente	El indicador no ha sido abordado o requiere reformularse en su totalidad.

Fuente: CEIPS (2014).

### Diseño y procedimiento

De acuerdo con Montero y León (2007), este trabajo consiste en un estudio *ex post facto*, retrospectivo y de un solo grupo simple, debido a que la información de las categorías de valoración y la puntuación total se recogió en un mismo momento, luego de ser realizado el proceso de valoración por parte de los pares evaluadores. Adicionalmente, se puede clasificar como una investigación instrumental (según los mismos autores), por cuanto se establecerá la calidad técnica del instrumento utilizado para la evaluación de pares.

### Procesamiento de datos

El análisis de datos para verificar la calidad técnica del instrumento se realizó dentro de la perspectiva de la Teoría de Respuesta al Ítem y, en particular, con el modelo de Rasch (Linacre, 2016), el cual genera valores en el nivel de intervalo a partir de observaciones estocásticas. En este caso, se trabajó específicamente con el modelo de Rasch de Andrich para escalas de valoración. De acuerdo con Linacre (2016), este modelo se expresa de la siguiente manera:

$$\log(P_{nij}/P_{ni(j-1)}) = B_n - D_i - F_j \quad (1)$$

Donde:

$P_{nij}$  = Probabilidad de asignar un valor particular de la escala (que en el caso del instrumento presente se encuentra en una escala de 5 puntos).

$B_n$  = Desempeño del estudiante en la escala de valoración (a partir de los valores asignados por los pares).

$D_i$  = Complejidad relativa de los ítems planteados en términos de la valoración asignada.

$F_j$  = Medida calibrada de los distintos valores de cada ítem.

Aiken (1996) plantea que las escalas de valoración son de amplio uso en investigación y psicología tanto básica como aplicada, y que son apropiadas para la evaluación de diversos aspectos, entre los que se encuentran las publicaciones –así como los proyectos de grado–. En este sentido, el modelo de Andrich permite obtener resultados sobre las valoraciones de los pares expertos, agrupándolas en las diferentes dimensiones del instrumento, e indicando las dimensiones que recibieron menor y mayor puntuación. Para la producción de datos se utilizó el software Winsteps (Linacre, 2016), que trabaja con base en los planteamientos y algoritmos del modelo de Rasch, y produce, igualmente, algunos indicadores de la Teoría Clásica de los Test.

Los indicadores para el análisis del instrumento utilizado son:

- a. *Confiabilidad*: desde la perspectiva del alfa de Cronbach, ya sea obtenida con los datos brutos o con las probabilidades asignadas por el modelo de Rasch; se refiere a la repetibilidad de los resultados, es decir, es un indicador de la consistencia interna (Pardo & Rocha, 2004). Igualmente, se produce el índice de separación del Modelo de Rasch, que permite analizar la sensibilidad del instrumento para clasificar a los evaluados de acuerdo con su desempeño (es decir, valorar la tesis de acuerdo con su calidad) (Linacre, 2016).
- b. *Dimensionalidad*: se analiza a través de los valores propios, o eigenvalues. Estos valores se producen con el Análisis de Componentes Principales de residuos (Linacre, 2016) y se utilizan con el fin de analizar la dimensionalidad del instrumento, es decir, si las categorías en las cuáles se asignaron los ítems son unidimensionales o no. Se espera que el valor propio de la varianza no explicada por el primer contraste sea inferior a 2 (es decir, 2 ítems).

Asimismo, los indicadores para analizar la calidad técnica de los ítems son:

- a. *Dificultad*: que en este caso se refiere a la valoración asignada por los pares. Si el valor de la dificultad es alto, es decir, si el ítem es difícil, quiere decir que el valor asignado es poco, y viceversa.
- b. *Ajuste al modelo*: se refiere al ajuste que tienen los datos al modelo, es decir, a la aproximación de las valoraciones a las probabilidades planteadas por el modelo.

- c. *Correlación punto biserial*: que a pesar de ser un indicador de la Teoría Clásica de los Test y de verse afectado por la población en la cual se estima, proporciona información clave en el sentido de saber si un ítem en particular mide lo mismo que el conjunto de ítems de su categoría. Aquí, se esperan valores altos y positivos.

Para el análisis del tamaño de los efectos y la significatividad de las relaciones entre las categorías evaluadas y la calificación final, así como para procesar los datos con ecuaciones estructurales y calcular los efectos directos e indirectos, se utilizó el software AMOS (Arbuckle, 2013).

## Resultados

Como ya se mencionó, los resultados apuntan tanto a la verificación de la calidad del instrumento de valoración empleado como a estimar el efecto que tienen las categorías de valoración en los resultados finales relacionados con la aprobación del proyecto de investigación.

### Calidad técnica del instrumento

En la Tabla 4 se presentan los resultados de los indicadores técnicos obtenidos con el programa Winsteps. Uno de los resultados más llamativos es el del valor propio del instrumento considerado como un todo. En este caso, el valor obtenido supera ligeramente el valor esperado (2.0 o menos), lo que indica la posibilidad de multidimensionalidad. En el contexto de esta investigación, eso quiere decir que, en efecto, hay varias dimensiones representadas por las categorías en las que se ha dividido el instrumento. Al analizar ese indicador para cada una de las categorías se encuentra que estas pueden considerarse unidimensionales. El menor promedio es el de la categoría método, lo que indica que es el aspecto o categoría que recibe la menor valoración por parte de los pares. En todos los casos, tanto para la prueba total como para cada una de las categorías, se obtienen valores altos de confiabilidad.

Tabla 4. Valores de los indicadores de la calidad técnica del instrumento

	N. ° ítems	Promedio	Desv. Est.	Alfa de C	KR # 20 Rasch	Separación	Eigenvalues
Antecedentes	7	1.79	1.95	0.88	0.81	2.05	1.82
Justificación	3	1.74	2.14	0.87	0.83	2.23	1.67
Método	7	1.49	2.01	0.92	0.84	2.31	1.78
Todo	17	1.45	1.62	0.95	0.92	3.30	2.10

En la Tabla 5 se encuentran los valores de los indicadores de los ítems. En general, se observa que todos los ítems cumplen con los criterios establecidos, especialmente el de ajuste al modelo en sus dos perspectivas: próximo y lejano –solo el ítem 7 de antecedentes supera el valor de 1.50 que utilizan varios autores como criterio (Linacre, 2016), no obstante, aún no supera el valor de 2.0, que indicaría ruido en la medición–. De esta manera, se concluye que el instrumento cumple con criterios de alta calidad técnica.

Tabla 5. Valores de los indicadores de los ítems

ENTRY	Dificultad	Ajuste próximo	Ajuste lejano	Correlación r pbis
Antecedentes 1	0.27	1.05	1.02	0.70
Antecedentes 2	-0.65	0.99	1.08	0.69
Antecedentes 3	-0.02	1.01	1.12	0.69
Antecedentes 4	-0.03	0.83	0.74	0.80
Antecedentes 5	0.11	0.80	0.78	0.80
Antecedentes 6	0.38	0.81	0.82	0.78
Antecedentes 7	-0.06	1.45	1.54	0.66
Justificación 1	-0.16	0.87	1.01	0.88
Justificación 2	-0.03	1.16	1.11	0.86
Justificación 3	0.19	0.84	0.80	0.91
Método 1	0.28	1.01	0.92	0.74
Método 2	0.04	0.84	0.80	0.83
Método 3	-0.40	1.48	1.48	0.71
Método 4	-0.30	0.89	0.96	0.80
Método 5	-0.27	0.76	0.72	0.83
Método 6	0.44	1.00	1.02	0.76
Método 7	0.21	1.05	1.03	0.72

### Relación de las categorías con la valoración total

Inicialmente, en este caso se trabajó con un modelo en el que se establecieron las relaciones directas de las categorías en la calificación total (véase Figura 5).

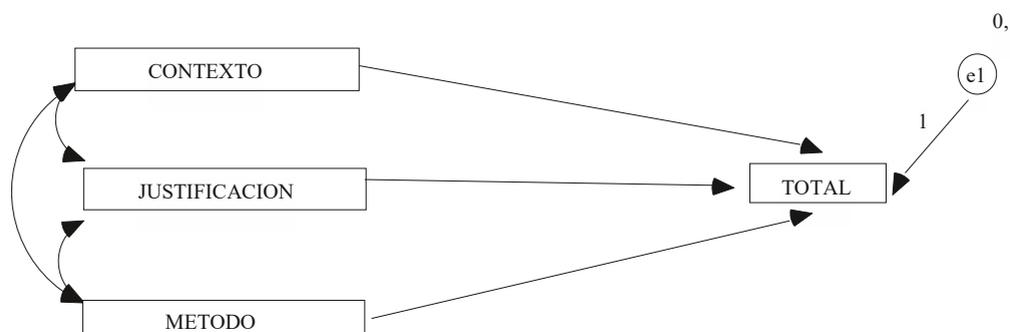


Figura 5. Relación de las categorías con la puntuación total.

Ahora bien, el programa de computador AMOS establece los pesos de regresión que valoran la relación entre las variables. Los pesos de regresión obtenidos y su valor de significatividad se pueden observar en la Tabla 6.

Tabla 6. Pesos de regresión obtenidos

	Relación	Peso de regresión	P
Total	<--- Contexto	0.39	***
Total	<--- Justificación	0.29	***
Total	<--- Método	0.38	***

En todos los casos se observa que la relación es significativa al 1 % (los asteriscos indican que el valor de p es el más bajo posible). Los valores de los pesos de regresión son estandarizados, es decir, corresponden a desviaciones estándar.

Para este modelo de ecuación planteado se obtuvieron los tamaños de efecto estandarizados: método = 0.381; justificación = 0.287; y contexto = 0.390. Resultado que indica que tanto el método como los antecedentes presentan tamaños relativamente altos e importantes, y que la categoría relacionada con la justificación no lo es tanto.

Adicionalmente, se consideró que la valoración realizada por un par es articulada entre las partes del documento, lo que implica que va estableciendo la relevancia y

pertinencia de los planteamientos del documento a medida que avanza en la lectura y que su valoración requiere de un ir y venir entre las diferentes partes del texto, además de una anticipación de lo que puede encontrar en los apartados que aún no ha leído. Es decir, que su actividad como evaluador es dinámica, analítica y crítica. En este sentido, se propuso el modelo de ecuación representado en Figura 6; sus valores del tamaño de efecto se muestran en la Tabla 7.

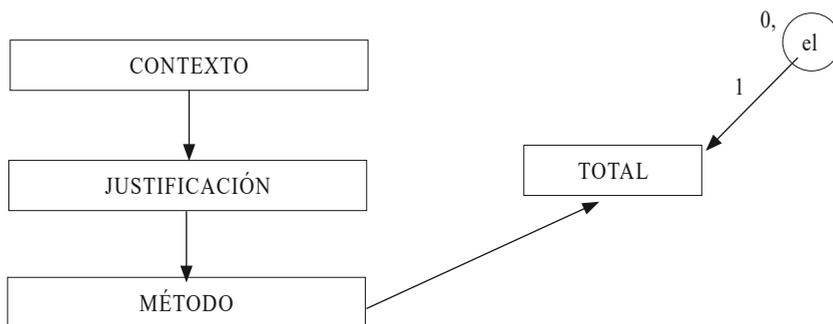


Figura 6. Relación de las categorías con la valoración total del proyecto.

Tabla 7. Tamaño de efecto estandarizados directos e indirectos

	Contexto	Justificación	Método
Justificación	0.790	0.000	0.000
Método	0.666	0.843	0.000
Total	0.623	0.788	0.935

En este caso, se observa que al aceptar que la valoración de los pares obedece a una lectura dinámica –es decir, aquella en la que el par va leyendo y al mismo tiempo relaciona lo nuevo que lee con lo ya leído y, en ocasiones, vuelve a leer partes previas del documento–, la valoración se convierte en un ejercicio activo de lectura –de leer y volver a leer–, y los efectos se potencian ostensiblemente. Es en este tipo de ejercicio, al potenciarse por la lectura previa del contexto y la justificación, el efecto del método es de casi una desviación estándar del resultado total obtenido. Este valor de 0.935 indica que al aumentar en una desviación estándar la valoración del método, se aumenta en 0.935 desviaciones estándar la valoración total del proyecto.

## Discusión

Los resultados muestran conclusiones positivas, ya que permiten confirmar la multidimensionalidad del formato en las categorías expresadas en su diseño, además de su calidad. El tema de la multidimensionalidad se hace relevante al considerar que, efectivamente, el formato de evaluación apunta a valoraciones de diferentes competencias relacionadas con el proceso de investigación y expresadas a través del documento del proyecto, lo que va más allá de una valoración simple de la competencia lingüística de escribir correctamente.

Adicionalmente, a partir de los resultados se puede establecer que el método es la categoría que recibe la menor valoración por parte de los pares, lo que resulta importante debido a la tendencia investigativa de la maestría y el énfasis en la formación metodológica de los estudiantes, aunque contradictorio con los hallazgos de Alston et al. (2005), Frésan (2013) y Barahona (2016), pero que permitiría orientar acciones de capacitación para los pares con el fin de fortalecer competencias asociadas al *peer review*, al tiempo que favorecería la revisión de la gradación de las puntuaciones asignadas a esta categoría, la cualificación de los procesos que orientan el ejercicio de los pares al momento de realizar las valoraciones, la posibilidad de apreciar con mayor detalle si la valoración a la dimensión de metodología obedece a un ejercicio independiente o si es consecuencia de la lectura anterior, y tener en cuenta si el proyecto está en sus etapas iniciales o si ya corresponde a la entrega de resultados.

Por otra parte, el análisis de la relación entre las categorías evidencia significancia al 1 % entre estas, y los tamaños del efecto muestran que el método y los antecedentes obtienen valores altos –a diferencia de la justificación, que obtiene valores bajos–, sin embargo, al plantear un modelo que refleja un ejercicio activo de lectura del evaluador a través del abordaje articulado y dinámico de las partes del documento, se evidencia una potenciación del efecto de las categorías en la evaluación global. Acá, para superar esta aparente contradicción es necesario comprender que la valoración que dan los pares evaluadores indica aspectos diferentes a lo que indica el efecto que esa valoración tiene; es decir, que la valoración de los pares es la calificación que ellos dan a las distintas partes del documento de acuerdo con las instrucciones del formato, mientras que el efecto es una medida de la importancia de la variable en la puntuación general del proyecto de investigación.

Por otra parte, estos resultados también llevan a tener más cuidado por parte de los directores de los proyectos a la hora de considerar el método de la investigación y los detalles que pueda tener al momento de su planteamiento, y con esto orientar al estudiante con mayor tiempo para que se pueda asumir una mejor escritura, atendiendo a indicadores de profundización y calidad. Estos resultados llevan a la maestría a analizar las opciones de formación que ofrece en términos de las asignaturas relacionadas con la metodología de investigación y a plantear, posiblemente, opciones de

acompañamiento a los proyectos de los estudiantes para mejorar las posibilidades de aprobación por parte de los pares evaluadores.

## Conclusiones

Teniendo en cuenta el tamaño del efecto encontrado, es posible establecer las siguientes conclusiones acerca del formato de evaluación de los proyectos de maestría: (a) sirve como guía para reorientar el proceso de formación de los estudiantes hacia los indicadores mejor evaluados por los pares; (b) permite diseñar una guía de presentación de los proyectos para los estudiantes; y (c) favorece su propia actualización hacia la estandarización de los criterios. Las aplicaciones de estos resultados se enmarcan en las actividades de evaluación por pares adelantadas en diferentes unidades y centros de investigación en cualquier disciplina, y se convierten en un referente para la identificación de factores que tienen mayor relevancia en la evaluación de la producción científica en el proceso de pares.

## Referencias

- Aiken, L. (1996). *Rating scales y checklist. Evaluating behavior, personality and attitude*. New York: John Wiley & Sons, Inc.
- Alston, M., Allan, J., Bell, K., Brown, A., Dowling, J., ... Williams, R. (2005). "SERPS Up": support, engagement and retention of postgraduate students – a model of postgraduate support". *Australian Journal of adult Learning*, 45(2), 172-190. Recuperado de <http://files.eric.ed.gov/fulltext/EJ797629.pdf>
- Arbuckle, J. (2013). IBM® SPSS® Amos™ 22. User's Guide. Recuperado de [http://www.sussex.ac.uk/its/pdfs/SPSS\\_Amos\\_User\\_Guide\\_22.pdf](http://www.sussex.ac.uk/its/pdfs/SPSS_Amos_User_Guide_22.pdf)
- Barahona, N. (2016). Deserción estudiantil en la Maestría en Docencia Universitaria para estudiantes de primer ingreso. *Análisis de la realidad nacional*, 102, 89-133. Recuperado de <http://revistasdigi.usac.edu.gt/ipn/index.php/IPN/article/view/59/89>
- Carlino, P. (2008). Exploración de géneros, diario de tesis y revisión entre pares: análisis de un ciclo de investigación – acción en talleres de tesis de postgrado. En E. Narvaja de Arnoux (Dir.), *Escritura y producción de conocimiento en las carreras de posgrado*. Buenos Aires, Argentina: Santiago Arcos. Recuperado de <http://media.utp.edu.co/referencias-bibliograficas/uploads/referencias/capitulo/carlino2008exploraciondegenerospdf-zNDdR-libro.pdf>
- Centro de Estudios e Investigaciones en Psicología (2014). *Líneas de investigación e integración curricular*. Bogotá: Universidad Católica de Colombia.
- Fréсан, M. (2013). Factores que propician el abandono y obstaculizan la culminación de los estudios de posgrado. Conferencia: Tercera conferencia latinoamericana sobre el abandono en la educación superior. 13, 14 y 15 de noviembre. Gestión Universitaria Integral del Abandono. México. Recuperado de [http://www.alfaguia.org/www-alfa/images/ponencias/clabesIII/LT\\_1/ponencia\\_completa\\_80.pdf](http://www.alfaguia.org/www-alfa/images/ponencias/clabesIII/LT_1/ponencia_completa_80.pdf)

- Gómez, H., & Sánchez, V. (2013). Indicadores cualitativos para la medición de la calidad en la educación. *Educación y Educadores*, 16(1), 9-24. doi: 10.5294/edu.2013.16.1.1
- Hernández, O, Pérez, R., & González, G. (2014). La deserción en los posgrados, un problema no menor. *Diálogos sobre educación*, 8, 1-15. México: Universidad de Guadalajara. Recuperado de <http://revistascientificas.udg.mx/index.php/DSE/article/view/3766/3550>
- Hortsmanshof, L., & Conrad, L. (2003). *Postgraduate peer support programme: enhancing community*. Brisbane, Australia: Griffith University. Recuperado de [http://www98.griffith.edu.au/dspace/bitstream/handle/10072/1838/22703\\_1.pdf;sequence=1](http://www98.griffith.edu.au/dspace/bitstream/handle/10072/1838/22703_1.pdf;sequence=1)
- Linacre, J. M. (2016). *Winsteps® Rasch measurement computer program User's Guide*. Beaverton, Oregon: Winsteps.com. Recuperado de <http://www.winsteps.com/>
- Ministerio de Educación Nacional [MEN]. (2010). Decreto 1295 de 2010. Bogotá: Colombia. Recuperado de [http://www.mineducacion.gov.co/1621/articles-229430\\_archivo\\_pdf\\_decreto1295.pdf](http://www.mineducacion.gov.co/1621/articles-229430_archivo_pdf_decreto1295.pdf)
- Ministerio de educación nacional [MEN]. (2015a). Boletín de la educación superior. Por primera vez el país mide la deserción de los programas de posgrado. Boletín N° 1, Febrero de 2015. Recuperado de [http://www.mineducacion.gov.co/1621/articulos-349245\\_recurso.pdf](http://www.mineducacion.gov.co/1621/articulos-349245_recurso.pdf)
- Ministerio de Educación Nacional [MEN]. (2015b). Caracterización de graduados por departamento (2001-2014). Bogotá, Colombia. Recuperado de <http://bi.mineducacion.gov.co:8380/eportal/web/men-observatorio-laboral/ubicacion-geografica>
- Montero, I., & León, O. (2007). A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology*, 7(3), 847-862. Recuperado de [http://www.aepc.es/ijchp/GNEIP07\\_es.pdf](http://www.aepc.es/ijchp/GNEIP07_es.pdf)
- Moreno, F, Cortés, J., & Zumaya, R. (2012). Usos, limitaciones y prospectiva de la evaluación por pares. *Revista Interamericana de Bibliotecología*, 35(2), 201-210. Medellín: Colombia. Recuperado de [http://www.scielo.org.co/scielo.php?script=sci\\_arttext&pid=S0120-09762012000200007](http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0120-09762012000200007)
- Pardo, C., & Rocha, M. (2004). *Glosario de términos técnicos en psicometría y medición y evaluación educativa*. Bogotá, Colombia: ICFES.
- Sierra, L. (2011). La elaboración de una tesis de maestría: exigencias y dificultades percibidas por sus protagonistas. *Revista Entornos, Universidad Surcolombiana. Vicerrectoría de Investigación y Proyección social*, 24, 171-183. Recuperado de <https://dialnet.unirioja.es/servlet/articulo?codigo=3798839>