

Capítulo 1

Consideraciones técnicas en la elección o diseño de instrumentos de medición

Reyes Piñuelas Erika Paola

<https://doi.org/10.61728/AE20256340>



Resumen

El propósito es describir las pautas técnicas y éticas para seleccionar o desarrollar instrumentos de medición en psicología y educación. Estos instrumentos son esenciales para evaluar fenómenos complejos como la ansiedad o las competencias docentes. Se destaca la importancia de la validez y confiabilidad, junto con consideraciones éticas respaldadas por estándares de APA, AERA y NCME. Se enfatiza la necesidad de definir variables y propósitos antes de elegir un método de recolección de datos, presentando un proceso de cuatro pasos: determinación inicial, delimitación y búsqueda, análisis comparativo y selección final. Se subraya la responsabilidad de los investigadores de adherirse a las mejores prácticas para asegurar validez y equidad en los resultados. En conclusión, el capítulo proporciona un marco exhaustivo que guía a los investigadores a seguir un enfoque ético y técnicamente sólido en el desarrollo y uso de instrumentos de medición.

Introducción

Los instrumentos de medición en el ámbito educativo y psicológico sirven para evaluar fenómenos como el aprendizaje de los estudiantes, competencias docentes, autorregulación del aprendizaje y ansiedad, por mencionar algunos. Estos tienen como característica común que no son directamente observables y mensurables; se requieren herramientas para acceder a su estudio.

En el proceso de investigación bajo un enfoque cuantitativo o mixto, una de las decisiones que tienen impacto en la generalización de los resultados y la validez de sus interpretaciones es la relacionada con el instrumento de medición que será utilizado. La selección o diseño de estos representa un paso dentro de las etapas del proceso investigativo que conducirá al cumplimiento de los objetivos específicos que fueron propuestos.

Sin embargo, en el proceso de búsqueda de un instrumento de medición sobre el constructo de interés, es probable que se encuentre con uno de un país o idioma diferente al de los participantes a quienes va dirigido el estudio, lo que implica, en principio, decidir si la ruta de la adaptación

de instrumento es la más viable, o bien, construir uno. No obstante, ¿qué criterios tomar en cuenta para decantarse por una de estas alternativas?

Tener claridad sobre las implicaciones de cada una de las opciones permite que los investigadores tomen decisiones éticas y responsables en el proceso de investigación. Por ello, el presente trabajo tiene como objetivo describir las pautas técnicas y éticas que se toman en cuenta en la decisión de diseñar o seleccionar un instrumento de medición. Es ofrecer una guía práctica, técnica y ética que permita a estudiantes e investigadores noveles tomar decisiones fundamentadas sobre la selección o diseño de instrumentos de medición en psicología y educación. A través de un enfoque experiencial, se busca proporcionar herramientas y orientaciones que faciliten la comprensión y aplicación de criterios clave para desarrollar instrumentos adecuados y fiables, considerando tanto los objetivos de la investigación como las implicaciones éticas del proceso.

Consideraciones éticas en el desarrollo de pruebas

A partir de preocupaciones y necesidades detectadas en el campo de la evaluación, asociaciones como la American Educational Research Association (AERA), American Psychological Association (APA) y National Council on Measurement in Education (NCME) desde 1985 —y en ediciones de 1999 y 2014—, se han dado a la tarea de desarrollar estándares que orienten hacia mejores prácticas en cuanto a los fundamentos, diseño, desarrollo y uso de las pruebas.

Preocupaciones compartidas por la International Test Commission (ITC, por sus siglas en inglés), que en el año 2014 publicó un documento que aborda los principios para el uso en investigación de tests psicológicos; en la misma tónica, se encuentra el modelo para evaluar la calidad de los test, propuesto por investigadores españoles (Hernández et al., 2016), en el que buscan atender las críticas sobre los problemas asociados con el uso de las pruebas en la evaluación. En México se desarrollaron los criterios técnicos para el desarrollo, uso y mantenimiento de instrumentos de evaluación, e impulsó el Instituto Nacional para la Evaluación de la Educación (DOF, 2017).

A pesar del avance en el establecimiento de criterios de las buenas prácticas en la elaboración y uso de los instrumentos de medida, con frecuencia los investigadores que diseñan o realizan adaptaciones de estos conocen de manera escasa literatura especializada y los métodos de medición (Díaz et al., 2017); reportan de manera parcial o insuficiente lo relativo a la confiabilidad y validez, o bien, no prestan importancia a los aspectos teóricos-conceptuales que subyacen en su estructura, con lo cual se pone en duda la calidad de las pruebas utilizadas y el aseguramiento de la equidad.

Los retos en cuanto a la calidad, las buenas prácticas en el diseño y uso de las pruebas en la evaluación persisten a casi cuatro décadas de la primera publicación de la APA, AERA y NCME sobre los estándares, ya algunas dificultades han sido enunciadas. De manera general, Hernández et al. (2020) declaran que, aun cuando son conocidas las directrices sobre este tema, en la práctica no se han implementado con la frecuencia que se desearía.

Los problemas en el diseño, adaptación y desarrollo de los instrumentos de medida también han sido declarados por diversos autores; algunos ejemplos de esto son las revisiones sistemáticas reportadas por: Ossa-Cornejo et al. (2017), con un estudio sobre pensamiento crítico, afirman que los instrumentos desarrollados han mostrado omisiones importantes respecto a la validez y han puesto en tela de juicio su viabilidad; asimismo, en la evaluación docente, Navarro y Ramírez (2018) señalan como principal dificultad la calidad técnica de los instrumentos; o en el caso de los exámenes de ingreso a las IES en México, pruebas consideradas de alto impacto, una constante es la ausencia de publicaciones que reporten el sustento metodológico en el proceso de elaboración, así como del análisis de sus resultados (Sánchez et al., 2020).

En otras palabras, como lo argumentan Hernández et al. (2022), la publicación de estándares y recomendaciones sobre el uso adecuado de los instrumentos ha sido infructuosa para lograr un cambio notorio en las prácticas evaluativas. Sin embargo, reconocer los desafíos en el ámbito del diseño, desarrollo y uso de los instrumentos de medición es imprescindible para superarlos.

Por otro lado, un aspecto que complejiza este tema son los alcances y limitaciones que presentan estos instrumentos de medida. Una de las

limitaciones inherentes a la naturaleza misma de los constructos es la precisión en su medición, dado que los constructos son una construcción teórica que permite comprender y medir aspectos psicológicos complejos e inobservables mediante indicadores observables en el comportamiento o productos de este (Soler, 2013).

Martínez et al. (2014) mencionan algunos aspectos asociados al diseño de instrumentos de medición, (1) no existe una única aproximación a la medición de un constructo que sea aceptada de manera universal, diversas definiciones del marco general (el primer paso en el diseño de pruebas) pueden llevar a diversas conclusiones; (2) los instrumentos se basan en muestras limitadas pero representativas de conductas del dominio que se pretende medir, la determinación de dicha muestra y variedad es uno de los mayores problemas; (3) toda medida obtenida a partir de los instrumentos contiene un componente de error, que puede deberse a una muestra limitada de contenido, errores de muestreo o ser aleatorios, entre otros; (4) las limitaciones derivadas de las escalas de medida; y (5) los constructos deben establecer su relación con otros constructos y las conductas observables.

Ahora bien, en el proceso de investigación, los instrumentos de medición sirven para registrar información sobre las variables que se estudian, las que están enunciadas en el planteamiento del problema, en la pregunta de investigación y los objetivos; elementos que se espera presenten congruencia entre sí. Yuni y Urbano (2020) sostienen que los instrumentos de obtención de datos son dispositivos que permiten medir los fenómenos empíricos y son aparatos especialmente elaborados para recoger información de la realidad; además, como lo afirma Gregori (2012), tienen como característica contar con un procedimiento estandarizado de una muestra de la conducta, y esta es descrita a partir de categorías o puntuaciones.

El diseño de instrumentos de medición requiere de diversos pasos. Muñiz (2018) plantea de manera detallada al menos 10 necesarios — para mayor información, consultar su obra —, en los cuales el énfasis se concentra en contar con evidencias de validez y confiabilidad. Estas evidencias son imprescindibles dado que respaldan las inferencias que se realizan a partir de los instrumentos, proporcionando un marco sólido para su interpretación.

La AERA, APA y NCME (2018) sostienen que la confiabilidad se trata de “la coherencia de puntajes entre replicaciones de un procedimiento de evaluación” (p. 35), mientras que la validez hace referencia al “grado en que la evidencia y la teoría respaldan las interpretaciones de los puntajes de una prueba para usos propuestos de las pruebas” (p. 11), con cinco fuentes de evidencias basadas en el contenido de la prueba, en los procesos de respuesta, en la estructura interna, en las relaciones con otras variables y, de validación y consecuencias de las pruebas. Para el estudio de cada una de estas dos propiedades imprescindibles, se utilizan diversos procedimientos para acumular evidencias que justifican la pertinencia de su utilización.

Consideraciones técnicas en la elección de un instrumento de medición

Para el desarrollo de un proyecto de investigación, existen trabajos de diversos autores que enuncian de manera explícita la decisión de desarrollar o identificar un instrumento de recolección, entre los que se encuentran Creswell (2012) y Pacheco (2015). El primero sugiere decidir sobre el tipo de técnica de obtención de información, identificar aspectos centrales del propósito de la investigación, así como las características de la población, muestra y procedimiento de ingreso a campo, antes de la toma de decisión sobre el desarrollo o localización de un instrumento (ver Tabla 1).

Tabla 1. *La elección del instrumento de medición en el proceso investigativo*

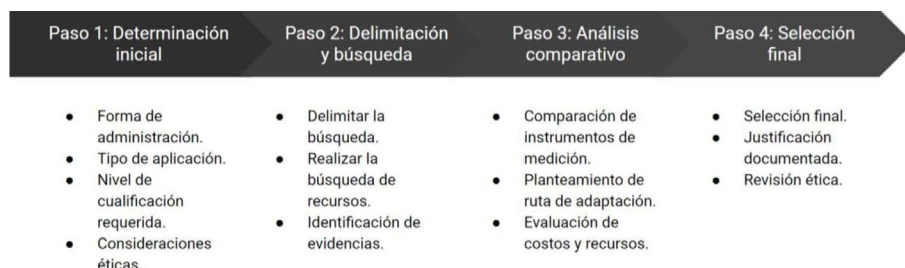
Autor	Pasos en una investigación
Creswell (2012)	1. Decidir si una encuesta es el mejor diseño a utilizar. 2. Identificar las preguntas o hipótesis de investigación. 3. Identificar la población, el marco muestral y la muestra. 4. Determinar el diseño de la encuesta y los procedimientos de recopilación de datos. 5. Desarrollar o localizar un instrumento. 6. Administrar el instrumento. 7. Analizar los datos para abordar las preguntas o hipótesis de investigación. 8. Escribir el informe.
Pacheco (2015)	1. Listar las variables que se pretende medir u observar. 2. Revisar su definición conceptual y comprender su significado. 3. Revisar cómo han sido definidas operacionalmente las variables, esto es, cómo se ha medido cada variable. 4. Elegir el instrumento o los instrumentos (ya desarrollados) que hayan sido favorecidos por la comparación y adaptarlos al contexto de la investigación. 5. Indicar el nivel de medición de cada ítem y, por ende, el de las variables.

Nota. Elaboración propia a partir de Creswell (2012) y Pacheco (2015)

Mientras que Pacheco (2015) le otorga mayor peso a la comprensión clara de las variables, como enunciarlas, definirlas y explorar su medición en otros estudios, antes de elegir un instrumento. Ambos autores plantean aspectos centrales que se deben realizar previo a la decisión de diseñar o seleccionar un instrumento para utilizar en un proyecto de investigación.

Determinación del instrumento de medición en un proyecto de investigación

Al elegir un instrumento de medición para un proyecto de investigación, se deben considerar diversos elementos que aseguren una selección adecuada y fundamentada. La Figura 1 presenta un proceso estructurado en cuatro pasos que facilita tomar una decisión informada y coherente con los objetivos del estudio.

Figura 1*Proceso para la determinación del instrumento de medición*

Nota. Elaboración propia.

Paso 1. Determinación inicial

Forma de administración. Conocer las características y contexto de los participantes permite valorar la forma de aplicación que es más pertinente; si bien de manera general se ubican dos, en papel y a través del uso de tecnología, también se encuentra una combinación de estas formas que dependerá de las condiciones y recursos humanos, materiales y económicos con los que se cuente. En el caso de aplicación en papel, se cuenta con mayor control, al posibilitar la detección de ciertas fuentes de error, y según Cisneros-Caicedo et al. (2022), se obtienen altas tasas de respuestas, que fluctúan entre 80-85 %.

Por otro lado, Muñoz et al. (2020) argumentan que los avances informáticos, multimedia e internet han provocado un cambio en los procesos de evaluación; en la aplicación de los instrumentos, cada vez es más frecuente, por ejemplo, cuando se envía por correo electrónico o a partir del llenado de un formulario. Algunos autores han reportado una tasa de respuesta baja, alrededor del 50 %, señala Creswell (2012); sin embargo, dependerá del seguimiento, las notificaciones adecuadas, interés de los participantes, calidad del instrumento e incentivos. Además, debe considerarse factores como el acceso a tecnología y la familiaridad de los participantes con las herramientas digitales y la extensión del instrumento.

Tipo de aplicación. Elegir el modo de administración, ya sea grupal o individual, influye tanto en la logística del estudio como en la calidad de los datos recolectados y, por ende, las conclusiones que se puedan

extraer. Se debe considerar que algunos constructos, como las habilidades cognitivas generales o ciertas actitudes y opiniones, pueden ser medidos eficazmente a través de aplicaciones grupales. Sin embargo, otros constructos, como la inteligencia, pueden requerir una atención personalizada que solo puede lograrse mediante la administración individual. La elección entre aplicación grupal e individual debe estar alineada con el objetivo central de la investigación y el fenómeno que se desea estudiar.

Un segundo elemento es el contexto en el que se llevará a cabo la investigación. En un entorno escolar, la aplicación grupal puede ser más práctica y económica, permitiendo evaluar a grandes grupos de estudiantes al mismo tiempo; esto es relevante cuando el objetivo es obtener un panorama general de ciertas capacidades o conocimientos en una población extensa. Empero, en un entorno clínico, donde se busca un diagnóstico más detallado, la aplicación individual favorece una mayor observación y seguimiento de las respuestas del participante. .

Las características de la muestra también juegan un papel relevante en esta decisión. Algunas poblaciones, como la niñez o personas con discapacidad intelectual, pueden requerir la atención individual para garantizar una comprensión adecuada de las instrucciones y facilitar respuestas precisas. Mientras que, en poblaciones más homogéneas, la aplicación grupal puede ser más adecuada.

Nivel de cualificación requerida. Para garantizar el cumplimiento de las normas éticas durante la administración del instrumento, es necesario identificar el nivel de cualificación necesario para aplicar e interpretar el instrumento. Los aplicadores deben estar preparados para manejar situaciones sensibles, responder a las preguntas de los participantes y proporcionar apoyo si surge algún malestar o incomodidad durante la administración.

Una categorización con frecuencia utilizada en función de la cualificación requerida por los usuarios es Nivel A, para los instrumentos de rendimiento y conocimientos, accesible a profesionales de otras disciplinas; Nivel B, pruebas colectivas de aptitudes e inteligencia, accesible para profesionales afines a la psicología según las especificaciones de la prueba; y Nivel C, los de aplicación individual de inteligencia, personalidad y otros instrumentos complejos, de uso exclusivo para psicólogos acreditados (Hernández et al., 2016).

Consideraciones éticas. Proteger la confidencialidad y el anonimato de los participantes es de suma importancia; implica que, en los datos recolectados, las identidades de los participantes no puedan ser reveladas y contar con un almacenamiento seguro de los datos para evitar accesos no autorizados. El consentimiento informado es fundamental en la investigación y en estudios que involucren a menores de edad; es necesario obtener el consentimiento de los progenitores o tutores y el asentimiento informado por parte de la niñez y juventud (Oliu y Bosch, 2021). Este doble proceso asegura que tanto los responsables legales como los propios menores comprendan y acepten su participación en el estudio; sin embargo, podría requerir un periodo de tiempo mayor.

Asimismo, al seleccionar un instrumento de medición, es importante considerar la sensibilidad del contenido y el impacto potencial en los participantes. En tales casos, se debe evaluar cuidadosamente la pertinencia del instrumento y considerar el uso de aplicaciones en un ambiente donde los participantes se sientan cómodos respondiendo con honestidad.

Paso 2: Delimitación y búsqueda

Delimitar la búsqueda. Identificar las palabras clave en coherencia con el objeto de estudio, es un aspecto importante en la búsqueda de instrumentos de medición, elegir los términos específicos que sean adecuados puede hacer la diferencia para tener éxito en las búsquedas; el uso de un vocabulario controlado como un tesoro facilita la identificación de los términos acertados, entre los más utilizados se encuentran el tesoro de la UNESCO (s. f.) por sus siglas en inglés, el catálogo del Índice de Revistas de Educación Superior e Investigación Educativa (Universidad Nacional Autónoma de México, 2015), y el tesoro del Education Resources Information Center [ERIC] (2023) del Institute of Education Sciences.

Para efectos del objetivo que se persigue en este capítulo, relacionado con la identificación de instrumentos de medición, encontrar los términos específicos del estudio y añadir en combinación con los operadores de búsqueda las palabras: “instrumentos de medición”, “validez”, “validez de las pruebas”, por citar algunas, puede dar resultados más exitosos, pues aumenta la posibilidad de encontrar información relevante y de alta

calidad para la selección de un instrumento de medición en investigaciones cuantitativas.

En cuanto a los operadores de búsqueda, Freitas et al. (2023) destaca la importancia de los operadores booleanos en la recuperación de información, explicando que estos se basan en la lógica algebraica desarrollada por el matemático Boole, fundamentales en la búsqueda de datos precisos, ya que permiten la creación de algoritmos eficientes mediante el uso de símbolos que cumplen funciones específicas. Los principales operadores booleanos son “AND”, “OR” y “NOT”; cada uno de estos ayuda a unir, diversificar o excluir términos clave, lo que optimiza la búsqueda de información científica al hacerla enfocada en lo que se necesita.

Por su parte, Carranza (2018) sugiere un procedimiento formal para realizar búsquedas efectivas utilizando operadores booleanos, los cuales se consideran una parte integral de las herramientas de las TIC. El procedimiento propuesto consta de cuatro etapas: (1) Formular las palabras clave: Identificar términos precisos y relevantes para la búsqueda; (2) Formular el perfil de búsqueda: Emplear operadores booleanos como OR, AND, NEAR, etc., para estructurar la búsqueda; (3) Llevar a cabo la búsqueda: Ejecutar la búsqueda en las bases de datos seleccionadas; (4) Discriminar la información: Evaluar y seleccionar la información relevante de los resultados obtenidos.

Realizar la búsqueda de recursos. Buscar en bases de datos, repositorios, con herramientas de inteligencia artificial y catálogos de instrumentos ofrece un acceso valioso a instrumentos previamente desarrollados y validados, así como a la literatura relevante que puede guiar el diseño de nuevos instrumentos.

Bases de datos académicas. Las bases de datos académicas son esenciales para identificar instrumentos de medición que hayan sido utilizados y validados en estudios previos. Entre las bases de datos más relevantes se encuentran: Google Académico, PsycINFO, Scientific Electronic Library Online (SCIELO), Dialnet, Latindex, Red de Revistas Científicas de América Latina y el Caribe, España y Portugal (R

edaly), EBSCO HOST, por citar algunos, cada una con características diversas.

Dos bases de datos tienen una mención especial, una es SCOPUS, cuenta con una función avanzada de análisis de los resultados de la búsqueda “Analyze results”, en la cual es posible delimitar no solo año o tipo de documento como es común en los filtros de la mayoría de las bases de datos, sino también por país, de qué institución se encuentra afiliado, así como disciplina o área a la que pertenecen; función que ayuda a identificar si existen instrumentos en el contexto donde se realiza el estudio. La otra mención la tiene la colección de ERIC (2023), que ofrece una sección dedicada a identificar instrumentos de medición. Al realizar búsquedas, los usuarios pueden acceder a la pestaña de “Assessments and surveys”, donde se listan los instrumentos utilizados en los recursos recuperados. Esta función proporciona a los investigadores los nombres de los instrumentos que se utilizan para el tema de interés y se logra identificar por número de apariciones cuáles son los más utilizados.

Repositorios institucionales. Son otra fuente valiosa para encontrar instrumentos de medición, contienen trabajos académicos, tesis y disertaciones que pueden proporcionar información sobre instrumentos utilizados en estudios locales o específicos de una región, entre los que destaca el de TESIUNAM (UNAM, 2020); el repositorio del Consejo Mexicano de Ciencias Sociales (s/f) se trata de recursos electrónicos y bibliotecas digitales de instituciones dedicadas a las ciencias sociales; y del Gobierno de España, el Ministerio de Universidades cuenta con un fichero de tesis doctorales TESEO (Ministerio de Universidades, s. f.).

Herramientas de inteligencia artificial. De manera reciente emergen herramientas derivadas de la inteligencia artificial; se ubican dos: una es el Research Rabbit (2023); como lo señala su página, se trata del Spotify para artículos, ayuda a descubrir literatura académica relevante. Al agregar artículos a colecciones personalizadas, la plataforma utiliza algoritmos para aprender sobre el interés del usuario y mejora las recomendaciones de recursos. Ofrece visualizaciones interactivas y genera listas de artículos relacionados basados en similitud temática y redes de autores. Esto permite no solo encontrar instrumentos relevantes, sino también explorar nuevas perspectivas y enfoques innovadores en el área

de investigación. Y otra es, el Semantic Scholar desarrollado por el Allen Institute for Artificial Intelligence utiliza IA e ingeniería innovadoras para comprender la literatura científica (Allen Institute, 2024).

Catálogos de pruebas. Ya sea de editoriales como son los de Manual Moderno y TEA Ediciones, o bien, de instituciones o autores que buscan compilar instrumentos de medición, por citar algunos ejemplos, se encuentran, el catálogo de escalas psicosociales y de la salud en español de Calleja (s/f), que nace de la experiencia de la autora al acompañar a sus estudiantes en la búsqueda de instrumentos válidos y confiables, y encontrar que con frecuencia se recurre a instrumentos inadecuados para su proyecto (López, 2019); el catálogo de pruebas de la Universidad Autónoma de San Luis Potosí (s. f.); la compilación de García et al. (2016) de escalas de evaluación psicológica, con el propósito de divulgar el trabajo derivado de investigaciones y con ello abonar a la democratización del conocimiento; y de otros países de Iberoamérica, el buscador de evaluación de tests editados del Consejo General de la Psicología en España (2023), el cual brinda una valoración cualitativa y cuantitativa de cada uno de los instrumentos incluido; y el catálogo de pruebas e instrumentos de Posada (2016) de la Universidad de San Buenaventura Medellín Colombia.

Identificación de evidencias de validez y confiabilidad. Existen elementos a considerar en la valoración de un instrumento de medición para utilizar dentro de una investigación; si bien, en general, se enuncia la confiabilidad y validez como las principales propiedades psicométricas, existen algunos recursos que pueden ayudar a registrar y analizar las características de los instrumentos que se encuentran en las búsquedas. Uno es el Cuestionario de Evaluación de Tests Revisado (CET-R) de Hernández et al. (2016), el cual consta de tres apartados a completar:

1. Descripción general del test a evaluar.
2. Características del test (la calidad de sus materiales, instrucciones, adaptación, desarrollo, sus ítems, etc.) y propiedades psicométricas (análisis de ítems, validez, fiabilidad e interpretación de las puntuaciones)
3. Valoración global del test.

El segundo recurso es el proporcionado por la Comisión Internacional de Test (ITC, 2014), que sugiere mínimamente proporcionar información sobre: (a) los fundamentos teóricos del test y su finalidad, (b) el sistema seguido para la selección inicial de los ítems, (c) cómo se analizaron posteriormente y los criterios de selección, (d) número de ítems de cada dimensión, (e) los métodos de escalamiento utilizados, (f) información sobre las evidencias de validez y la precisión de las mediciones, por ejemplo, fiabilidad u otros indicadores de la precisión de las mediciones, según el modelo de medida del test.

Paso 3: Análisis comparativo

Comparación de los instrumentos de medición encontrados. En el ejercicio de comparar los instrumentos a partir de las búsquedas realizadas, es importante destacar que no se trata de elegir entre opciones binarias; la decisión debe estar respaldada por un análisis exhaustivo de varios factores clave que aseguren la idoneidad del instrumento en el contexto específico de la investigación. Es esencial considerar la fundamentación teórica que sustenta el instrumento, que esta sea congruente con los preceptos teóricos del estudio; asimismo, contar con evidencias sobre consistencia interna, confiabilidad, error de medición, validez de contenido, validez de constructo, validez de criterio e interpretación, aspectos que el CET-R contempla (Hernández et al., 2016). Además, el contexto donde fue realizado, para valorar los pasos posibles ante procedimientos de adaptación, así como el año de elaboración o revisión que se reporta.

Planteamiento de ruta de adaptación. Entre los aspectos a considerar para establecer una ruta para adaptar un instrumento de medición se encuentran los trabajos de Hernández et al. (2020), Muñiz et al. (2013), Muñiz (2018) y AERA, APA y NCME (2018), los cuales convergen en cinco aspectos a considerar; Uno, realizar un análisis exhaustivo del instrumento original, en contenido, y su propósito, evaluar las características psicométricas, identificar elementos culturales que pudieran afectar; Dos, traducir el instrumento es una etapa clave en la adaptación de un test, se sigue un procedimiento para garantizar la equivalencia lingüística y

conceptual, que incluye usualmente la traducción directa, traducción inversa y la revisión de expertos; Tres, realizar una adaptación cultural que va más allá de la traducción literal e implica modificar elementos para que sean culturalmente relevantes y equivalentes; Cuatro, realizar una prueba piloto con análisis de validez, confiabilidad y de equivalencia que permita comprobar la equivalencia factorial y la invarianza métrica entre el test original y el adaptado; y cinco, la revisión para posibles ajustes adicionales y el reporte son necesarios para asegurar que el test adaptado cumple con las normas y estándares internacionales para la adaptación de instrumentos.

Evaluación de costos y recursos. Sin el análisis pertinente, se puede poner en riesgo la conclusión de un estudio en el tiempo que fue proyectado. Por ello, se requiere analizar los recursos humanos, materiales y financieros con los que se cuenta para la realización de posibles procedimientos de adaptación o costos por compra del instrumento, así como los generados por la aplicación del mismo en el contexto de estudio y asegurar una muestra adecuada.

Paso 4: Selección final

Selección final. Elegir el instrumento que mejor cumpla con los criterios establecidos y ofrezca las mejores garantías psicométricas, o bien, optar por el diseño de uno que garantice su ajuste al objetivo y preceptos teóricos del proyecto, así como al contexto de los participantes.

Justificación documentada. Elaborar un informe que incluya las razones de la elección y las evidencias que respaldan la decisión, el cual abona al apartado del método en un reporte del estudio y permite dejar constancia sobre los argumentos que llevaron a la elección de determinado instrumento sobre otro.

Revisión ética. Obtener la aprobación de uso de los autores del instrumento de medición (International Test Commission, 2017), así como el análisis de las normas éticas y legales al utilizar, adaptar o comprar uno, según sea el caso.

En concreto, dar cuenta de los criterios a tomar en consideración en la determinación de un instrumento sobre otro, o bien el diseñar uno propio,

es una tarea compleja y depende de diversos factores. Por un lado, se encuentran autores como Creswell (2012), que se decantan por invertir esfuerzos en la búsqueda y análisis de potenciales instrumentos para realizar procedimientos de adaptación; lo anterior, por el conocimiento que se requiere sobre la construcción de los test y los numerosos pasos a seguir, que podrían desencadenar en una calidad métrica cuestionable.

Por otro lado, se encuentran Morales et al. (2003), que se inclinan más hacia el diseño de instrumentos, al argumentar que los constructos o rasgos no son universales y pueden ser distintos de una cultura a otra, en este mismo tema, existen determinados instrumentos y autores que han alcanzado cierto prestigio, pero eso no es garantía de la calidad del instrumento cuando se traduce para ser utilizado en otro idioma o cultura, y cuando son instrumentos que cuentan con cierta antigüedad debe considerarse que los constructos varían en el tiempo y tener connotaciones diferentes a las de origen. En la Tabla 2 se sintetizan algunos aspectos a favor y en contra sobre la decisión antes planteada.

Tabla 2*Decisión de adaptar o diseñar un instrumento de medición*

	Seleccionar o adaptar un instrumento	Diseñar un instrumento
A favor	* Existen numerosos instrumentos disponibles de diversas editoriales que pueden ser usados para propósitos evaluativos. * Logra ahorrar mayor tiempo y recursos * Más sencillo localizar un instrumento y ajustarlo. * Posibilita la comparación de los resultados con otros estudios . * Cuentan con evidencias de confiabilidad y validez.	* La teoría o modelo en el que se sustenta el instrumento es coherente con los preceptos de la investigación. * Contextualizado a las características de la población a la que va dirigido.
En contra	* Los instrumentos comerciales pueden ser caros. * Están dirigidos a una población muy específica. * No adecuarse al propósito evaluativo requerido.	* Consta de varios pasos rigurosos que requieren de dominio del proceso de construcción. * Se requiere de la colaboración de otros profesionales expertos. * El tiempo necesario para llevar a cabo algunas etapas depende de la disponibilidad de terceros.

Nota. Elaboración propia con información de Covacevich (2014); Creswell (2012), Morales (2003); Yuni y Urbano (2020).

A manera de conclusión

La selección adecuada de instrumentos de medición en psicología y educación no solo depende de consideraciones técnicas, sino también de una serie de principios éticos que son fundamentales para el éxito de cualquier investigación.

Un aporte del texto es la propuesta de un proceso de selección de instrumentos de medición, que abarca la identificación de las variables y el contexto de aplicación, hasta la consideración de las cualificaciones necesarias y las implicaciones éticas. Esto ayuda a garantizar la validez

y confiabilidad de los instrumentos seleccionados, y asegura que se respeten los derechos y bienestar de los participantes.

Por otro lado, con el avance en herramientas de inteligencia artificial, presentan un panorama prometedor para la búsqueda de recursos y la identificación eficiente de instrumentos de medición. Además, la utilización de repositorios institucionales y catálogos en línea permite un acceso más amplio a instrumentos ya validados, lo que favorece, en palabras de García et al. (2016), la democratización del conocimiento y la estandarización de las prácticas de medición a nivel global.

Los alcances de esta propuesta son proporcionar un marco comprensivo para investigadores, psicólogos y educadores, no solo en la elección de instrumentos de medición, sino también de la posible decisión de una adaptación cultural. Asimismo, fomenta el cumplimiento de normativas éticas y legales, resaltando la importancia de la confidencialidad y el consentimiento informado, elementos esenciales en cualquier estudio de investigación. Finalmente, se contribuye promoviendo prácticas de investigación más rigurosas y responsables que son esenciales para la generación de conocimientos valiosos y aplicables en las disciplinas de la psicología y la educación.

Referencias

- Allen Institute (2024). *About Semantic Scholar*. <https://www.semanticscholar.org/about>
- American Educational Research Association, American Psychological Association y National Council on Measurement in Education (2018). *Estándares para Pruebas Educativas y Psicológicas*. APA.
- Calleja, N. (s. f.). *Catálogo de Escalas Psicosociales y de la Salud en Español*. Universidad Nacional Autónoma de México. <https://planeacion.psicol.unam.mx/escalas/inventarios/buscar>
- Carranza, J. (2018). Aplicación de las tecnologías de la información y la comunicación para la búsqueda de información científica en el posgrado de especialidades médicas. *Anestesia en México*, 30(1), 18-25. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2448-87712018000100018&lng=es&tlng=es.

- Cisneros-Caicedo, A., Guevara-García, A., Urdánigo-Cedeño, J., y Garcés-Bravo, J. (2022). Técnicas e instrumentos para la recolección de datos que apoyan a la investigación científica en tiempo de pandemia. *Dominio de las ciencias*, 8(1), 1165–1185. <https://doi.org/10.23857/dc.v8i1.2546>
- Consejo General de la Psicología de España (2023). Evaluación Tests Editados en España. <https://www.jornadas.cop.es/evaluacionTest/>
- Consejo Mexicano de Ciencias Sociales (s. f.). *Repositorio*. COMECESO. <https://www.comecso.com/repositorio-digital>
- Covacevich, C. (2014). *Cómo seleccionar un instrumento para evaluar aprendizajes estudiantiles* [Nota técnica]. Banco Interamericano de Desarrollo.
- Creswell, J. (2012). *Educational Research Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Pearson.
- Criterios técnicos para el desarrollo, uso y mantenimiento de instrumentos de evaluación, Instituto Nacional para la Evaluación de la Educación, Diario Oficial de la Federación [D.O.F.], 28 de abril de 2017, (México).
- Díaz, C., Caso, J., y Contreras, L. (2017). Estándares y directrices que orientan el desarrollo y adaptación de instrumentos de medición. En J. Rodríguez y J. Caso (Coords.). *Prácticas de investigación aplicada a contextos educativos*, (p. 14-30). Universidad de Guadalajara.
- Education Resources Information Center (2023, agosto). *ERIC Thesaurus*. Institute of Education Sciences. <https://eric.ed.gov/>
- Freitas, B. F. de, Castro, C. S., Alves, E. L., Mota, E. M. de B., Brito, I. E. de, Miranda, M. A., Muniz, N. D., Lopes, P. C., Souza, T. L. P. de, Oliveira, T. P. de, & Aquino, R. L. de. (2023). O uso dos operadores como estratégia de busca em revisões de literatura científica. *Brazilian Journal of Implantology and Health Sciences*, 5(3), 652–664. <https://doi.org/10.36557/2674-8169.2023v5n3p652-664>
- García, B., Lozano, M., y Díaz, I. (2016). *19 Escalas de Evaluación Psicológica*. 19-Escalas-De-Evaluacion-Psicologica.pdf (ujat.mx)
- Gregori, R. (2012). *Pruebas psicológicas*. Pearson.
- Hernández, A., Elosua, P., Fernández-Hermida, J., y Muñiz, J. (2022). Test Commission: twenty five years working on test quality. *Papeles*

- del Psicólogo*, 43(1), 55-62. Epub . <https://dx.doi.org/10.23923/pap.psicol.2978>
- Hernández, A., Hidalgo, M., Hamblenton, R., y Gómez-Benito, J. (2020). International Test Commission guidelines for test adaptation: A criterion checklist. *Psicothema*, 32(3), 390-398. doi: 10.7334/psicothema2019.306
- Hernández, A., Ponsoda, V., Muñiz, J., Prieto, G., & Elosua, P. (2016). Cuestionario de Evaluación de tests Revisado (CET-R). *Papeles del Psicólogo*, 37(3), 192-197. <https://www.cop.es/uploads/pdf/CET-R.pdf>
- Hernández, A., Ponsoda, V., Muñiz, J., Prieto, G., y Elosua, P. (2016). Revisión del modelo para evaluar la calidad de los tests utilizados en España. *Papeles del Psicólogo*, 37(3), 192-197.
- International Test Commission y Consejo General de la Psicología de España (2014). *El uso de los tests y otros Instrumentos de evaluación en investigación*. ITC.
- International Test Commission. (2017). *The ITC Guidelines for Translating and Adapting-Tests*. www.InTestCom.org
- López, P. (2019). Base de datos en línea. Presentan catálogo de escalas psicosociales. *Gaceta UNAM*. Mayo, sección Comunidad.
- Martínez, M., Hernández, M. y Hernández, M. (2014). *Psicometría*. Alianza.
- Ministerio de Universidades (s. f.). *Tesis doctorales*: TESEO. Ministerio de Universidades, Gobierno de España. <https://www.educacion.gob.es/teseo/irGestionarConsulta.do;jsessionid=40169A62212EF9B509F-BA6AD55944697>
- Morales, P., Urosa, B., y Blanco, A. (2003). *Construcción de escalas de actitudes tipo Likert. Una guía práctica*. La Muralla.
- Muñiz, J. (2018). *Introducción a la Psicometría, Teoría clásica y TRI*. Pirámide.
- Muñiz, J. Elosua, P. y Hambleton, R. (2013). Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema*, 25(2), 151-157. doi: 10.7334/psicothema2013.24
- Muñiz, J., Hernández, A., y Fernández-Hermida, J. (2020). Utilización de los test en España: el punto de vista de los psicólogos. *Papeles del Psicólogo*, 41(1), 1-15. <https://dx.doi.org/10.23923/pap.psicol2020.2921>

- Navarro, C., y Ramírez, M. (2018). Mapeo sistemático de la literatura sobre evaluación docente (2013-2017). *Educação e Pesquisa, São Paulo*, 44,(e185677), 1-23. doi: <http://dx.doi.org/10.1590/S1678-4634201844185677>
- Oliu Bosch, S., y Bosch Núñez, A. (2021). El asentimiento informado en la investigación clínica. *Revista UNIMED*, 3(2). <https://revunimed.sld.cu/index.php/revestud/article/view/72>
- Ossa-Cornejo, C., Palma-Luengo, M., Lagos-San M., Quintana-Abello, I., y Díaz-Larenas, C. (2017). Análisis de instrumentos de medición del pensamiento crítico. *Ciencias Psicológicas*, 11(1), 19-28. <https://doi.org/10.22235/cp.v11i2.1343>
- Pacheco, P. (2015). Construcción y validación de los instrumentos para la medición de la influencia de los campos emocionales en los aprendizajes significativos. *Revista Internacional de Educación y Aprendizaje*, 3(1), 1-24. <http://sobrelaeducacion.com/>
- Posada, D. (2016). *Catálogo de pruebas e instrumentos*. Universidad de San Buenaventura Medellín.
- Research Rabbit (2023). *Research Rabbit*. <https://researchrabbitapp.com/>
- Sánchez, M., García, M., Martínez, A., y Buzo, E. (2020). El Examen de Ingreso a la Universidad Nacional Autónoma de México: Evidencias de validez de una prueba de alto impacto y gran escala. *Revista Iberoamericana de Evaluación Educativa*, 13(2), 107-128. <https://doi.org/10.15366/riee2020.13.2.006>
- Soler, S. (2013). Los constructos en las investigaciones pedagógicas: cuantificación y tratamiento estadístico. *Atenas Revista Científica Pedagógica*, 3(23), 84-101. <https://atenas.umcc.cu/index.php/atenas/article/view/585>
- UNESCO (s. f.). *Tesaurus de la UNESCO*. UNESCO.
- Universidad Autónoma de San Luis Potosí (s/f). *Área de pruebas psicológicas*. UASLP. https://cihbyp.uaslp.mx/cihbyp-content/pruebas_psicologicas.html
- Universidad Nacional Autónoma de México (2015). *Índice de revistas de educación superior e investigación educativa* (6.^a ed.). UNAM.
- Universidad Nacional Autónoma de México (2020). TESIUNAM. UNAM. https://tesiunam.dgb.unam.mx/F?func=find-b-0&local_base=TES01

Yuni, J., y Urbano, C. (2020). *Metodología y técnicas para investigar Recursos para la elaboración de proyectos, análisis de datos y redacción científica*. Editorial Brujas.