






Validación en español de un cuestionario para la evaluación de la calidad de la mentoría en investigación en ciencias de la salud

Spanish validation of a questionnaire for the quality evaluation of research mentoring in health sciences

Sergio Adrián Terrasa^{a,b,c} , Fernando Vazquez Peña^c , María Lourdes Posadas Martínez^a , Natalia Rosli^d , Eduardo Durante^{b,e,c} 

Resumen

Introducción. Es relevante contar con un cuestionario breve y auto-administrado para evaluar la calidad de los procesos de mentoría en investigación en la comunidad científica hispanohablante.

Objetivos. Validar una versión abreviada, auto-administrada y en español del cuestionario multidimensional *Evaluación de las competencias de mentoría*.

Métodos. 1) Traducción y adaptación transcultural; 2) Reducción de la versión traducida; 3) Validación de la versión abreviada en una muestra local de 159 participantes respecto de su validez de constructo (análisis factorial confirmatorio) y de criterio (correlación de su puntaje total con el de una pregunta agregada y con la versión extensa del cuestionario).

Resultados. Obtuvimos un cuestionario abreviado, en español y auto-administrado de 17 ítems divididos en cinco dimensiones: 1) Mantenimiento de una comunicación efectiva, 2) Alineación de expectativas, 3) Evaluación del conocimiento, 4) Promoción de la independencia, 5) Promoción del desarrollo profesional. El modelo hipotetizado ajustó de manera aceptable a los datos de nuestra muestra de validación (RMSEA=0,045; CFI=0,992; TLI=0,99 y WRMR=0,564). Todos los parámetros estimados para los ítems individuales tuvieron significancia estadística, oscilando las cargas factoriales estandarizadas entre 0,571 y 0,936. La fiabilidad compuesta de los diferentes dominios estuvo comprendida entre 0,764 y 0,865, y la varianza media extractada, entre 0,522 y 0,644, en concordancia con lo recomendado. La validez convergente fue adecuada ya que las cargas factoriales estandarizadas de 94 % de los ítems del modelo final fueron mayores a 0,6 y con significancia estadística (valor $t > 1,96$). La correlación entre los cinco factores fue alta (0,740 a 0,944). En la mayoría de las comparaciones, el cuadrado de dichas correlaciones fue superior a la varianza media extractada, lo que refleja una validez divergente o discriminante baja.

El puntaje de la versión abreviada tuvo una correlación positiva y alta (Coeficiente de Pearson 0,847; $p < 0,001$) con la pregunta agregada como criterio externo concurrente de validación y muy alta correlación (coeficiente de Pearson: 0,988; $p < 0,001$) con el puntaje de la versión original traducida de 26 ítems.

Conclusiones. La versión abreviada en español del cuestionario *Evaluación de las competencias de mentoría* tiene adecuadas propiedades psicométricas y puede ser aplicada en forma auto-administrada, lo que la vuelve la primera herramienta en español para evaluar procesos de mentoría en investigación.

Abstract

Introduction. It is relevant to have a short and self-administered questionnaire to assess research mentoring process quality in the Spanish-speaking scientific community.

Objectives. To validate a short and self-administered version in Spanish of the multidimensional questionnaire *Mentoring Competency Assessment*.

Methods. 1) Translation and cross-cultural adaptation; 2) Reduction of the translated version; 3) Validation of the shortened version in a local sample of 159 participants regarding its construct validity (confirmatory factor analysis) and its criterion validity (correlation of its total score with that of an additional question and with the long version of the questionnaire).

Results. We obtained an abbreviated (17 items) and self-administered questionnaire in Spanish, divided into five constructs: 1) Maintaining effective communication, 2) Aligning expectations, 3) Addressing knowledge, 4) Promoting independence, 5) Promoting professional development. The hypothesized model fit acceptably to the data from our validation sample (RMSEA=0.045; CFI=0.992; TLI=0.99 and WRMR=0.564).

All the parameters estimated for the individual items were statistically significant, with their standardized factor loads ranging from 0.571 to 0.936. The composite reliability of the different domains ranged between 0.764 and 0.865, and the mean extracted variance, between 0.522 and 0.644, in accordance with the recommended range. Convergent validity was adequate since the standardized factor loads of 94% of the items in the final model were greater than 0.6 and with statistical significance (t -value $> 1,96$). The correlation between the five factors was high (0.740 to 0.944) and in most of the comparisons, the square of these correlations was higher than the average variance extracted, reflecting low divergent or discriminant validity. The score of the shortened version had a positive and high correlation (Pearson coefficient 0.847; $p < 0.001$) with the additional question as a concurrent external validation criterion and a very high correlation (Pearson coefficient: 0.988; $p < 0.001$) with the score of the original 26-item translated version.

Conclusions. The shortened Spanish version of the *Mentoring Competency Assessment* questionnaire has adequate psychometric properties and can be applied in a self-administered way, which makes it the first tool in Spanish to assess research mentoring processes.

Palabras clave: Mentoría, Investigación, Encuestas y Cuestionarios, Estudios de Validación. Keywords: Mentoring, Research, Surveys and Questionnaires, Validation Study.



Introducción

La *mentoría* o tutoría es una reconocida estrategia de formación de investigadores en el mundo científico en general¹ y en el de las ciencias de la salud en particular^{2,3}.

En tres artículos previos⁴⁻⁶ realizamos una descripción y una comparación de las diferentes definiciones y conceptualizaciones de la mentoría en investigación a partir de varios paradigmas, la evidencia con la que contamos respecto de su utilidad, y las características de una buena y una mala mentoría.

Dada la importancia documentada de la calidad de la interacción entre las personas que realizan mentoría en investigación y quienes las reciben, en las últimas décadas ha crecido el interés en la evaluación de estos procesos mediante instrumentos estructurados. Fue en este contexto que se publicó en 2016 una revisión con algunos criterios sistemáticos de búsqueda sobre las herramientas disponibles para evaluar las relaciones de mentoría⁷. Luego de revisar seis bases de datos bibliográficas (CINHAL, Medline, PsycINFO, Academic Search Premier, ERIC, Business premier) sus autores identificaron 22 escalas, tres de las cuales corresponden a ciencias de la salud y una sola, a mentoría en investigación en salud.

Entre ellas se destaca el instrumento publicado en 2013 por Fleming et al.⁸, de la Universidad de California San Francisco. Sus autores validaron en 286 díadas de *mentor-mentoreado* el instrumento para la *Evaluación de las competencias de la mentoría* en investigación clínica y traslacional (en inglés, *Mentoring Competency Assessment* o *MCA*). Luego de una extensa revisión bibliográfica de los instrumentos de evaluación existentes en ese momento histórico, entrevistas cognitivas a profesionales que estaban ejerciendo o recibiendo mentoría en investigación clínica y/o traslacional, y en forma alineada con los objetivos de un Programa de Educación de Mentores, el cuestionario fue construido por un panel de expertos de 16 ciudades de los EE.UU.⁸

El instrumento MCA es un cuestionario multidimensional con 26 ítems de tipo Likert con siete opciones ordinales que puntúan de 1 a 7, y que representan seis competencias básicas de la mentoría en investigación: 1) el mantenimiento de una comunicación efectiva; 2) la alineación de las expectativas; 3) la evaluación de la comprensión; 4) el abordaje de la diversidad; 5) la promoción de la independencia; 6) la promoción del desarrollo profesional. Su consistencia interna fue muy buena, y los dominios hipotetizados tuvieron un buen ajuste con los datos observados. Además, en las 286 díadas analizadas se observó una moderada correlación entre las evaluaciones realizadas por quién provee la mentoría y quién la recibe. Durante este proceso de validación este instrumento fue administrado en forma presencial por un entrevistador.

Esta herramienta fue destacada por la Academia Nacional de Ciencias, Ingeniería y Medicina de los EE.UU.⁹ debido a la prolijidad de la metodología empleada durante su desarrollo. Fue usado para la evaluación de la eficacia de una intervención educativa tendiente a desarrollar las competencias en la mentoría en los EE.UU. y mostró sensibilidad para detectar cambios en la percepción de los mentoreados que habían sido asignados a la rama

intervención de un ensayo clínico aleatorizado con grupo comparador¹⁰, y también en un estudio antes-después sin grupo control¹¹. Fue recientemente validado en el idioma turco¹² y en la actualidad es utilizado como herramienta de evaluación de diversos programas tendientes al desarrollo de habilidades en mentoría^{13,14}.

Justificación de esta investigación

Consideramos de gran utilidad para las instituciones que forman investigadores y para las agencias que monitorean su formación, contar con alguna herramienta breve, autoadministrada, que pueda ser completada sin la necesidad de un entrevistador, que permita evaluar la percepción de efectividad del proceso de mentoría por parte de quien la recibe. La escala MCA fue desarrollada en inglés⁸. Luego de una revisión bibliográfica en las bases de datos MEDLINE y LILACS (ver material suplementario 1: *Estrategia de búsqueda bibliográfica*) no pudimos documentar que hubiera sido traducida y/o adaptada transculturalmente al español.

Destacamos que en nuestro medio, además de las restricciones presupuestarias por todos conocidas, existe poca cultura de medición de la calidad de los procesos y de los resultados en mentoría de investigación. Por eso consideramos que contar con una versión breve de este cuestionario adaptado para su uso local podría facilitar la implementación en forma masiva de este tipo de mediciones en Argentina y en la región de habla hispana a nivel de maestrías y doctorados, en especial teniendo en cuenta que suele ser necesario aplicarlas en forma repetida, por ejemplo, antes y después de brindar alguna intervención educativa a quienes ejercen roles de mentoría, para mejorar su desempeño.

En este artículo comunicamos el proceso y los resultados de la validación de una versión corta, autoadministrada y en español del cuestionario multidimensional *Evaluación de las competencias de la mentoría*⁸ para la evaluación de la calidad de la mentoría en investigación en ciencias de la salud.

MÉTODOS

Diseño

Traducción, adaptación transcultural, reducción (obtención de una versión abreviada) y validación psicométrica de un cuestionario estructurado a través de un diseño mixto en el que fueron empleadas técnicas cualitativas y cuantitativas.

Población

Fueron invitados a participar aquellos profesionales de la salud que estaban, o habían estado en el pasado, involucrados en una relación de mentoría en investigación definida como aquella en la que se cumplían los siguientes tres criterios de elegibilidad: 1) la existencia de un vínculo académico entre dos personas, una de ellas con mayor experiencia y/o desarrollo profesional o académico en alguna disciplina; 2) la persona de mayor experiencia y/o desarrollo profesional (mentora) enseñaba o aconsejaba a la otra (mentoreada), o bien tenía el rol institucional de supervisarla, acompañarla o respaldarla regularmente en algún proyecto de

^a Departamento de Investigación, Hospital Italiano de Buenos Aires. maria.posadas@hospitalitaliano.org.ar

^b Departamento de Medicina, Instituto Universitario Hospital Italiano de Buenos Aires.

^c Servicio de Medicina Familiar y Comunitaria, Hospital Italiano de Buenos Aires. sergio.terrasa@hospitalitaliano.org.ar, fernandoramon.vazquez@hospitalitaliano.org.ar, eduardo.durante@hospitalitaliano.org.ar

^d Departamento de Educación, Instituto Universitario Hospital Italiano de Buenos Aires. natalia.rosli@hospitalitaliano.org.ar

^e Departamento de Investigación, Centro de Investigación en Educación para Profesionales de la Salud, Instituto Universitario Hospital Italiano de Buenos Aires.

investigación; 3) este vínculo debía haber durado un mínimo de seis meses, con un mínimo de cuatro encuentros (presenciales o virtuales) entre las partes.

La elegibilidad fue independiente del grado de satisfacción que estuviera generando o hubiera generado el vínculo con el mentor o la mentora.

Descripción del procedimiento de traducción y adaptación transcultural del cuestionario

Realizamos la traducción y adaptación transcultural de acuerdo a las recomendaciones de los consensos de expertos^{15,16}:

1. Traducción a la lengua objetivo por dos traductores con el español como lengua nativa y conciliación posterior de ambas versiones;
2. Retro-traducción al inglés por dos personas con este idioma como lengua nativa, cegadas al cuestionario original, y posterior conciliación de ambas versiones;
3. Comparación de las equivalencias conceptuales, semánticas y de contenido entre el cuestionario original y la versión retro-traducida al inglés por un comité bilingüe conformado por los cuatro traductores y el grupo investigador (fueron invitados participar los autores de la herramienta original, pero no recibimos respuesta a esa invitación);
4. Lectura final en busca de errores gramaticales y ortográficos;
5. Cinco entrevistas cognitivas de acuerdo a lo sugerido por la bibliografía¹⁶ de potenciales personas usuarias del cuestionario para identificar las dificultades que percibían mientras lo completaban, realizadas hasta lograr la saturación del discurso;
6. Formulación de la versión final del instrumento.

Descripción del procedimiento de reducción de los ítems del instrumento original

Una vez concretada la traducción y adaptación transcultural, el cuestionario fue sometido a un proceso matemático¹⁷ y conceptual para eliminar los ítems menos representativos de cada dominio.

La etapa cuantitativa de la eliminación de ítems fue efectuada tomando en cuenta las cargas factoriales estandarizadas del análisis factorial confirmatorio comunicadas por en el estudio original⁸, que habían sido lo suficientemente diferentes entre sí como para permitir la selección desde el punto de vista cuantitativo de los ítems potencialmente eliminables. El hecho de que la validez convergente en los factores del cuestionario original no haya sido alta, nos permitió argumentar que sería razonable plantear la posibilidad de reducir la cantidad de sus ítems.

Aclaremos que los ítems que son indicadores de un mismo constructo o *dimensión específicos* deberían compartir una alta proporción de la varianza (validez convergente) y que existen varias maneras de estimar esa validez convergente a través de las mediciones de los ítems. Entre ellas, se destaca la magnitud de las cargas factoriales, ya que cuando existe una alta validez convergente, altas cargas factoriales indicarían que los ítems convergen en un punto común: el constructo latente. Por eso y como mínimo, todas las cargas factoriales deberían poseer significancia estadística. Sin embargo, y dado que podría constataarse este tipo de significancia en presencia de una carga factorial débil, una regla bastante aceptada es que, además de la significancia estadística, las cargas estandarizadas sean idealmente superiores a 0,6¹⁸.

Dado que el cuestionario original⁸ cuenta con 26 ítems divididos en seis dominios y que para optimizar la estabilidad de cada dominio es recomendable que cada uno de ellos esté representado al menos por tres ítems¹⁸, estimábamos que podríamos lograr

una versión abreviada del cuestionario de aproximadamente 18 preguntas (como veremos más adelante, finalmente fueron 17).

El cuestionario resultante del proceso de la traducción y adaptación transcultural del instrumento original fue sometido a una discusión de un grupo de expertos integrado por ocho médicos y una bioquímica con formación en mentoría en investigación, para definir cuáles ítems serían eliminados del cuestionario final sobre la base de su potencial redundancia conceptual con otros ítems que sí permanecerían en el cuestionario final abreviado y/o su menor relevancia, y también teniendo en cuenta las cargas factoriales estandarizadas del análisis factorial confirmatorio comunicadas originalmente por Fleming et al.⁸ Esta discusión tomó lugar durante una reunión de una hora y media de duración y fue coordinada por el investigador principal y por su director de tesis doctoral.

Descripción del procedimiento de validación psicométrica de la versión abreviada

Una vez lograda la versión abreviada del cuestionario, confirmamos en una muestra local la carga factorial de los ítems que fueron conservados, siguiendo la estructura del trabajo original⁸. De esta manera utilizamos la muestra original como muestra de *calibración* y nuestra muestra, como muestra para realizar una *validación cruzada*.

El tamaño muestral para la realización del análisis factorial confirmatorio fue calculado sobre la base de fórmula propuesta por Jöreskog: $(k+1)(k+2)/2$; donde $k = n^{\circ}$ de ítems del instrumento en cuestión¹⁹. Para realizar un análisis factorial confirmatorio de un cuestionario abreviado de 18 ítems, habría sido necesario contar con 185 cuestionarios completos, y en el caso de obtener un cuestionario abreviado con 17 ítems, estimábamos necesario contar con 171 observaciones.

Procurando optimizar el poder de nuestra investigación y tratando de minimizar los sesgos que podrían haber resultado de haber realizado un análisis que hubiera incluido solamente las observaciones para las que contáramos con los datos completos –solamente justificable cuando es posible argumentar que la falta del dato en cuestión responde a un mecanismo puramente aleatorio²⁰ –, realizamos un proceso de imputación múltiple²¹ de los ítems que no habían sido respondidos (datos perdidos), asignando a cada dato faltante el valor promedio de las respuestas que el resto de los participantes había otorgado al ítem en cuestión.

Dado que nuestro tamaño muestral terminó siendo ligeramente más pequeño que el previsto inicialmente, decidimos realizar el análisis factorial confirmatorio con la metodología más apropiada para muestras pequeñas: la estimación robusta de mínimos cuadrados ponderados, disponible en el programa Mplus^{®22}. Esta metodología cuenta con evidencia proveniente de ensayos de simulación²³ que respalda su desempeño para estimar los parámetros con una precisión aceptable, tanto en modelos simples como complejos, en presencia de tamaños muestrales entre 100 y 1000 observaciones con distribución normal y no normal.

Con el objetivo de representar de la manera lo más versátil posible los vínculos matemáticos²⁴ entre dos o más preguntas con cierta superposición conceptual y/o de su estructura lingüística, previmos ajustar el análisis factorial confirmatorio a través de la correlación de sus errores de medida, siempre que dichas preguntas estuvieran incluidas en el mismo factor y se constatará entre ellas una asociación residual más allá de la correlación compartida por pertenecer al mismo dominio.

Fue evaluada la fiabilidad compuesta de cada dominio y la validez convergente y divergente del instrumento. Para evaluar la validez convergente verificamos que las cargas factoriales estandarizadas en los modelos finales fueran todas mayores a 0,6 y/o 0,7, y con significancia estadística ($t > 1,96$)¹⁸. Para determinar la

validez discriminante, evaluamos si la varianza que cada dominio compartió con sus indicadores propios superaba a la compartida con los otros dominios incluidos en el modelo (situación deseable)²⁵.

Por otro lado, realizamos una validación de criterio externa concurrente del cuestionario mediante el agregado de la pregunta: *¿Recomendaría a esta persona como tutor o mentor?* Esta pregunta tuvo la misma escala de Likert que el resto y, basándonos en la bibliografía especializada²⁶, esperábamos que su puntuación correlacionara en forma moderada con el promedio de todos los ítems del cuestionario (coeficiente de Pearson entre 0,3 y 0,7).

Por último, realizamos una exploración gráfica de la concordancia entre los promedios de ambas versiones (la de 26 ítems y la versión reducida) a través de un gráfico de Bland y Altman²⁷, calculando además la media de las diferencias entre sus puntajes globales y los límites del 95 % de su acuerdo.

Los análisis estadísticos fueron realizados con los programas Stata 13[®] y MPlus[®].

Aspectos éticos

Dado que permite garantizar que quien participe de la investigación, lo haga en concordancia con sus valores, intereses y preferencias²⁸, les propusimos a las personas que participaron en esta investigación que brindaran su consentimiento informado luego de informarles: el propósito de la investigación, el tiempo estimado para completar el cuestionario (seis minutos) y que sus respuestas serían tratadas en forma confidencial. Dado que el llenado del cuestionario supuso participación activa, se asumió que quien completara el cuestionario luego de haber leído el consentimiento informado, estaría consintiendo participar.

El protocolo de esta investigación fue concordante con las Guías Éticas de Investigación en Educación²⁹ vigentes y fue aprobado por el Comité de Evaluación de Protocolos de Investigación del Hospital Italiano de Buenos Aires con el número 3150.

RESULTADOS

Traducción y adaptación transcultural del cuestionario

El material suplementario 2 muestra el texto original del encazado del cuestionario y de las preguntas en inglés, y el texto final resultante del proceso de la traducción, retro-traducción y conciliación de las diversas versiones obtenidas, luego de incorporar las sugerencias recolectadas a través de las entrevistas cognitivas con personas potencialmente usuarias del instrumento.

Describimos a continuación la justificación de los principales cambios realizados durante el proceso de adaptación transcultural.

Uso de la primera persona en las preguntas

Como parte de la adaptación transcultural, quienes realizamos las dos primeras traducciones al español –una psicóloga experta en ciencias de la educación y un médico con experiencia en mentoría en investigación–, estuvimos de acuerdo en construir la mayoría de las preguntas en forma de afirmaciones en primera persona del singular, a diferencia del cuestionario original en inglés, en el que dichas preguntas estaban formuladas como preguntas en segunda o tercera persona del singular. Esta decisión estuvo respaldada por las siguientes dos razones:

1. Procurar reflejar con mayor potencia lingüística la experiencia personal de quien recibe la mentoría en investigación a través de una oración escrita en primera persona.
2. Frente a la única opción disponible en inglés (you), evitar tener que optar por algún pronombre personal de la segunda persona del singular de los disponibles en español (tú, vos o usted). Si bien actualmente estas diferencias

son cada vez menores^{30–32}, estos pronombres personales tienen un uso muy diferente en distintas regiones (el pronombre personal vos es de uso eminentemente rioplatense, mientras que en el resto de las regiones de habla hispana predomina el empleo del pronombre personal tú). Por otro lado, el pronombre personal de cortesía (usted) implica mayor respeto o distancia social, menor confianza entre quienes hablan, o bien refleja que una persona más joven se está dirigiendo en forma respetuosa a una persona de mayor edad o prestigio social.

Modificaciones realizadas a partir de las entrevistas cognitivas

Ítem 1. Incluimos la aclaración de lo que significa brindar una *escucha activa*.

Ítem 2. Si bien en el ámbito de las ciencias de la educación, la traducción exacta al español de la palabra inglesa *feedback* es *retroalimentación*, decidimos traducirla como *devolución* ya que esta expresión es más usada en el ámbito de las ciencias de la salud. Decidimos además conservar dicho término en inglés, ya que es muy utilizado en nuestro medio evocando el verdadero concepto de retroalimentación.

Ítem 3. Reemplazamos la palabra *confianza* por la expresión *confianza mutua* ya que a las personas entrevistadas no les quedaba claro la direccionalidad respecto de la percepción de confianza en la relación que estábamos evaluando, vale decir si se trataba de confianza de la persona que recibe la mentoría respecto de quien la brinda, viceversa o ambas, y nuestra intencionalidad era evaluar ambas.

Ítem 19. Donde se podría haber traducido *'... negocia conmigo un camino para promover mi independencia profesional...'*; decidimos traducir *'... negocia y consensúa...'* ya que las personas que integraron las entrevistas cognitivas atribuían al verbo negociar una connotación predominantemente económica.

Ítem 20. Donde se podría haber traducido *'...identifica sus prejuicios...'*; decidimos traducir *'...identifica sus propios prejuicios...'* para darle mayor claridad a la afirmación, ya que a algunas de las personas entrevistadas les quedaban dudas respecto de a quién estaba formulado dicho interrogante (de la persona que ejerce la mentoría o de la entrevistada).

Ítem 22. Si bien el Diccionario Cambridge³³ estipula que la traducción al español de la palabra *network* es *red*, procurando ser más específicos, en nuestra versión traducida y adaptada transculturalmente al español decidimos expresar dicho concepto como *'... red de contactos profesionales...'*

Ítem 24. En vez de traducir literalmente el enunciado originalmente expresado en inglés *'... me ayuda a balancear mi trabajo y mi vida personal...'*, decidimos cambiarlo a *'... me ayuda a organizarme para balancear mi vida profesional con mi vida personal...'*; debido a que cuando realizamos la traducción literal, a las personas entrevistadas les daba la impresión de que la ayuda en el balance podía tener una connotación negativa semejante a cierta intromisión en la vida privada de quien recibe la mentoría en investigación. La modificación de la pregunta desactivó dicha percepción.

Reducción del instrumento original preservando los ítems más representativos de cada dominio

Fueron invitados a participar 453 profesionales a través de un formulario en línea que les fue enviado por correo electrónico. Luego de varias solicitudes la tasa de respuesta fue de 35%; obtuvimos un total de 159 cuestionarios completos (93% del número previsto para un cuestionario de 17 ítems, número definitivo de ítems del instrumento). Hubo una mediana de 5% (Intervalo intercuartilo 3,1 a 17) de datos faltantes.

La cantidad y la proporción de respuestas completas de cada uno de los ítems, así como la media del puntaje obtenido y su des-



vío estándar, se muestran en el material suplementario 3. Durante la sesión de consenso fueron discutidos en forma sucesiva el conjunto de ítems que integran cada uno de los dominios y consensuada la eliminación de nueve de ellos: los ítems 1, 5, 6, 11, 15, 17, 20, 24 y 25.

El dominio *Abordaje de la diversidad*, que integraba el cuestionario original extenso⁸, contaba con dos ítems (20 y 21), de los que el comité de expertos decidió eliminar el número 20. Para evitar que quedara un dominio representado con un sólo ítem (21), el grupo de expertos acordó incluirlo en el dominio *Mantenimiento de una comunicación efectiva*, y el dominio *Abordaje de la diver-*

sidad fue eliminado.

De esta forma obtuvimos un cuestionario final integrado por 17 ítems que representaban un total de cinco dominios: 1) Mantenimiento de una comunicación efectiva (ítems 2, 3, 4 y 21); 2) alineación de expectativas (ítems 7, 8, 9 y 10); 3) evaluación del conocimiento (ítems 12, 13 y 14); 4) promoción de la independencia (ítems 16, 18 y 19); 5) promoción del desarrollo profesional (ítems 22, 23 y 26). La Tabla 1 y el material suplementario 4 muestran la versión abreviada del cuestionario con la nueva numeración de sus preguntas en forma correlativa.

Tabla 1. Versión final abreviada (17 ítems) del cuestionario de MCA: *Evaluación de las Competencias de Mentoría* obtenida luego del consenso de expertos.

Dominio	Número de ítem	Texto de las preguntas
Mantenimiento de una comunicación efectiva	1	Da devoluciones (feedback) constructivas
	2	Generó conmigo una relación de confianza mutua
	3	Es capaz de identificar y adaptarse a los estilos comunicacionales de las distintas personas con las que trabaja
	4	Es capaz de trabajar efectivamente con personas de características diferentes a las de él/ella (edad, raza, sexo, clase social, lugar de nacimiento, cultura, religión, composición familiar, etc.)
Alineación de las expectativas	5	Trabaja conmigo para establecer expectativas claras respecto de nuestra relación de tutoría/mentoría
	6	Alinea sus expectativas con las mías
	7	Considera el posible impacto de nuestras diferencias personales y profesionales sobre las expectativas de cada uno de nosotros
	8	Trabaja conmigo para definir mis propósitos de investigación
Evaluación del conocimiento	9	Detecta con bastante exactitud mi nivel de conocimiento científico
	10	Detecta con bastante exactitud mi habilidad para llevar a cabo investigaciones
	11	Emplea estrategias para mejorar mi comprensión de la investigación
Promoción de la independencia	12	Contribuye a que desarrolle confianza en mi mismo
	13	Reconoce mis contribuciones a la investigación
	14	Negocia y consensúa conmigo un camino para promover mi independencia profesional
Promoción del desarrollo profesional	15	Me ayuda efectivamente a desarrollar mi red de contactos profesionales
	16	Me ayuda a definir mis metas profesionales
	17	Me ayuda a obtener recursos económicos (subsidios, becas, etc.)

Evaluación de las propiedades psicométricas de la versión abreviada

Validez de constructo

Los resultados del análisis factorial confirmatorio de la versión abreviada, realizados mediante la estimación robusta de mínimos cuadrados ponderados, se describen en la Figura 1 y en la Tabla 2.

Dado que los ítems 9 y 10 de la versión breve tenían muy alta correlación residual, definida esta como la que ocurre más allá de la correlación que tienen dos variables por el sólo hecho de integrar el mismo dominio, el análisis realizado por el programa MPlus[®] nos propuso incluir la correlación de los errores de medida de ambas variables. Dado que ambas preguntas integran el mismo factor, tienen la misma estructura gramatical –comienzan con las cinco mismas palabras: 'Detecta con bastante exactitud mi...'- y se encuentran ubicadas en forma contigua en el cuestionario, aceptamos la sugerencia, procurando representar lo mejor posible su relación matemática y conceptual²⁴.

Tal como recomienda la bibliografía especializada, las cargas factoriales de la mayoría de los ítems fueron mayores a 0,6 y con significancia estadística. Por otro lado, la fiabilidad compuesta de todos los dominios fue mayor a 0,7 y la varianza media extractada,

mayor a 0,5, ambos valores en concordancia con lo recomendado por la bibliografía³⁴.

La validez convergente fue adecuada ya que las cargas factoriales estandarizadas de 16 de los 17 ítems (94 %) del modelo final fueron mayores a 0,6 y con significancia estadística ($t > 1,96$)¹⁸.

Respecto de la validez discriminante o divergente, sería deseable que la varianza media extractada (VME) de cada factor fuera mayor que el cuadrado de las correlaciones entre el propio factor y cada uno de todos los demás dominios incluidos en el modelo²⁶. Como lo muestra la Tabla 3, el cuadrado de las correlaciones entre los factores fue en general muy alta, salvo entre el factor *Promoción del desarrollo profesional* y los factores *Mantenimiento de una comunicación efectiva*, *Alineación de las expectativas* y *Evaluación del conocimiento*. Dado que esta situación implica una validez discriminante o divergente pobre, testeamos el ajuste de tres tipos de modelos de análisis factorial confirmatorio (ver Tabla 4) y documentamos que, a pesar de que los dominios no se individualizan totalmente, el ajuste fue mejor considerando cinco factores que considerando uno solo factor o dos factores (el que incluía las tres preguntas del dominio *Promoción del desarrollo profesional* vs. un segundo factor con el resto de las preguntas).

Figura 1. Resultados del análisis factorial de la versión abreviada

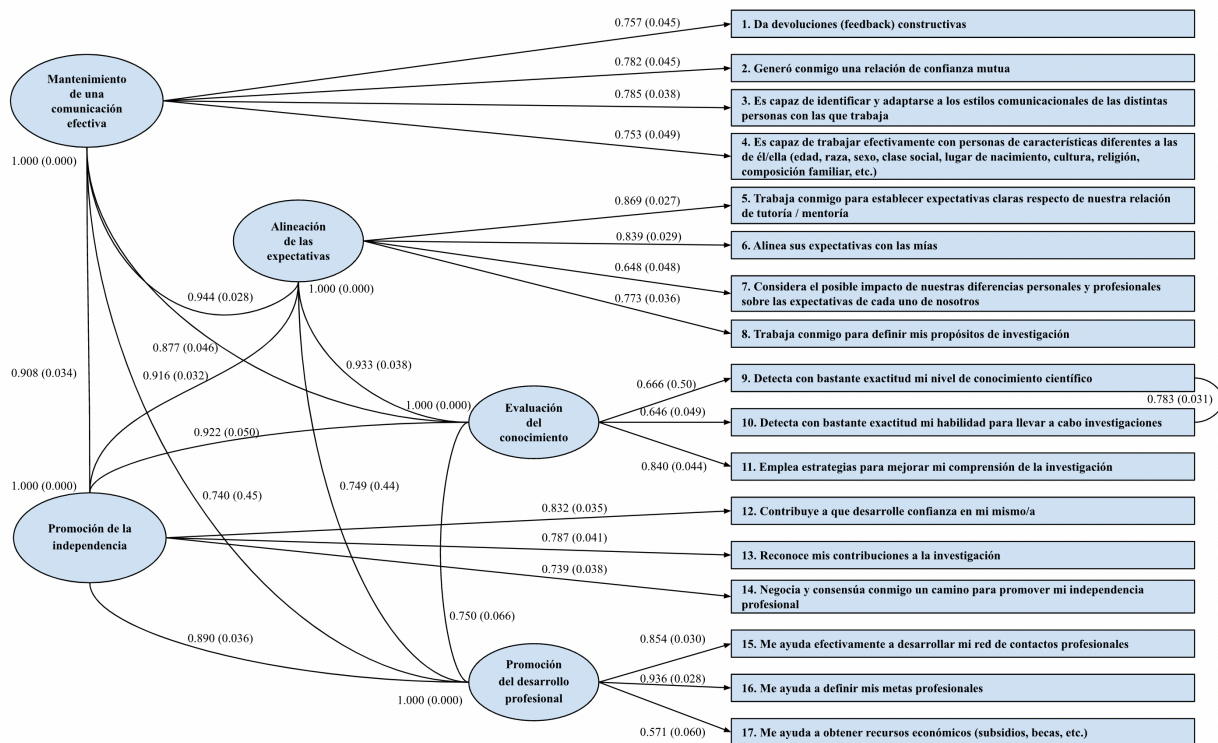


Tabla 2. Principales indicadores de la versión final abreviada (17 ítems) del cuestionario *Evaluación de las Competencias de Mentoría*

Dominio	Número de ítem	Carga factorial: coeficiente estandarizado	Carga factorial: error estándar	Valor t	Fiabilidad compuesta	Varianza media extractada
Mantenimiento de una comunicación efectiva	1	0,757	0,045	>1,96	0,853	0,592
	2	0,782	0,045	>1,96		
	3	0,785	0,038	>1,96		
	4	0,753	0,049	>1,96		
Alineación de las expectativas	5	0,869	0,027	>1,96	0,865	0,619
	6	0,839	0,029	>1,96		
	7	0,648	0,048	>1,96		
	8	0,773	0,036	>1,96		
Evaluación del conocimiento	9	0,666	0,050	>1,96	0,764	0,522
	10	0,646	0,049	>1,96		
	11	0,840	0,044	>1,96		
Promoción de la independencia	12	0,832	0,035	>1,96	0,830	0,619
	13	0,787	0,041	>1,96		
	14	0,739	0,038	>1,96		
Promoción del desarrollo profesional	15	0,854	0,030	>1,96	0,839	0,644
	16	0,936	0,208	>1,96		
	17	0,571	0,060	>1,96		

Tabla 3. Evaluación de la validez discriminante o divergente del cuestionario a través de la comparación entre la varianza media extractada de cada factor (valores de la diagonal de la tabla en letra negrita) y el cuadrado de las correlaciones de este factor y cada uno del resto de los factores (resto de los valores numéricos de la tabla). Notas: VME: varianza media extractada. De acuerdo a lo propuesto por Fornell²⁵, habría sido deseable que los valores de varianza media extractada hubieran sido mayores que el resto de los valores numéricos de la tabla.

	Mantenimiento de una comunicación efectiva	Alineación de las expectativas	Evaluación del conocimiento	Promoción de la independencia	Promoción del desarrollo profesional
Mantenimiento de una comunicación efectiva	VME: 0,592	0,891	0,769	0,824	0,548
Alineación de las expectativas	0,891	VME: 0,619	0,870	0,839	0,561
Evaluación del conocimiento	0,769	0,870	VME: 0,522	0,850	0,563
Promoción de la independencia	0,824	0,839	0,850	VME: 0,619	0,792
Promoción del desarrollo profesional	0,548	0,561	0,563	0,792	VME: 0,644

Tabla 4. Comparación del ajuste de tres modelos según los indicadores de buen ajuste. Notas: RMSEA: error de aproximación cuadrático medio; CFI: índice de ajuste comparado; TLI: índice de Tucker-Lewis; WRMR: residuo cuadrático medio ponderado; ^aSegún la bibliografía, el Chi² no es un indicador apropiado cuando hay más de 12 variables ya que puede presentar un resultado significativo aún en presencia de un buen ajuste del modelo³⁴. ^bSegún la bibliografía, no debe ser mayor de 0,08 y cuanto más pequeño, mejor³⁴. ^{c y d}Según la bibliografía, ambos deben ser mayores de 0,9 y cuánto más altos, mejor. ^eSegún la bibliografía, debe ser menor que 1 y cuanto más pequeño, mejor²⁴.

Modelos	Chi ² ^a	RMSEA ^b	CFI ^c	TLI ^d	WRMR ^e
Un factor	<0,0001	0,080	0,973	0,969	0,807
Dos factores	0,0004	0,056	0,987	0,985	0,659
Cinco factores	0,0134	0,045	0,992	0,990	0,564

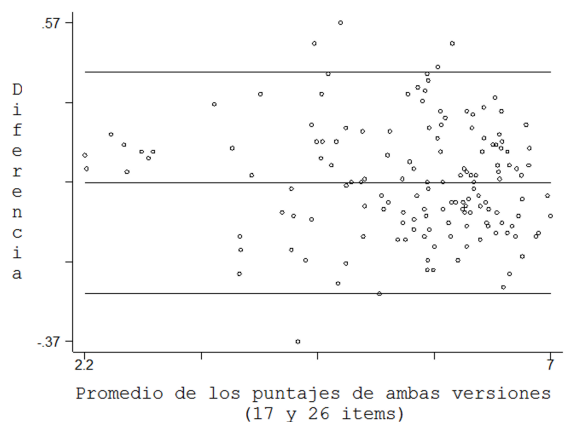
Validez de criterio externo concurrente

La correlación entre el puntaje total del cuestionario y la pregunta de validación de criterio externo convergente *¿Recomendaría a esta persona como tutor o mentor?* fue positiva y alta (coeficiente de correlación de Pearson: 0,847; p<0,001).

Además, documentamos una casi perfecta correlación y concordancia entre el puntaje global de la versión extensa (26 ítems) y el de la versión abreviada (17 ítems) del cuestionario. La correlación (coeficiente de Pearson: 0,988; p<0,001) fue muy alta y tal como afirma Streiner²⁶, podemos considerar que ambos cuestionarios (la versión extensa y la versión abreviada) se comportan como si fueran la misma herramienta de medición, corroborándose desde el punto de vista matemático la irrelevancia de los ítems eliminados.

La exploración gráfica de la concordancia de ambas versiones a través de un gráfico de Bland y Altman (Figura 2) permite observar que la nube de puntos no tiene una estructura particular, lo que refleja que la discordancia entre ambas escalas no predomina en ningún rango particular de su puntuación. La diferencia media entre los puntajes de ambas escalas fue 0,098 puntos (intervalo de confianza [IC] del 95% 0,072 a 0,124) y los límites del 95% del acuerdo se encontraron entre -0,229 y 0,426, diferencias poco relevantes dado el rango de puntuación de la escala (1 a 7).

Figura 2. Concordancia entre la versión extensa (26 ítems) y la versión abreviada (17 ítems) a través de un gráfico de Bland y Altman.



DISCUSIÓN

A través de esta investigación hemos validado una versión abreviada (17 ítems), autoadministrada y adaptada transcultural-

mente al español del cuestionario multidimensional *Evaluación de las competencias de la mentoría* (en inglés *Mentoring Competency Assessment* o *MCA*) publicado por Fleming et al.⁸

Nuestra tasa de respuesta fue de 35 %, similar a la comunicada por la bibliografía especializada respecto de las encuestas realizadas a través de Internet que tienen como destinatarios a profesionales de la salud³⁵⁻³⁸. Logramos alcanzar el 93 % (159/171) del tamaño muestral planificado. Por esta razón, decidimos realizar el análisis factorial confirmatorio mediante la estimación robusta de mínimos cuadrados ponderados, que cuenta con evidencia de que en presencia de tamaños muestrales de un mínimo de cien observaciones con distribución normal y no normal, es capaz de realizar estimaciones con una precisión aceptable²³.

Entre las limitaciones de nuestro diseño mencionamos que, al igual de Aikens et al.³⁹, habíamos establecido como criterio de inclusión el haber participado de un vínculo académico de un mínimo de seis meses de duración con un/a mentor/a, lo que puede ser considerado breve, dada la perspectiva temporal en la que puede transcurrir una relación de mentoría.

Entre las limitaciones de nuestros resultados, mencionamos que si bien documentamos que el instrumento resultante tiene buena capacidad para poder realizar en forma global una clasificación adecuada entre personas con diferentes competencias para ejercer mentoría en investigación, su validez divergente o discriminante podría llegar a no ser suficiente. Esto implica que, salvo en lo que se refiere a la promoción del desarrollo profesional, nuestro instrumento no es tan confiable a la hora de diferenciar el nivel de competencia en las diferentes áreas que hacen a una buena mentoría en investigación, representadas en cada uno de los dominios del cuestionario. Dicho de otra forma, entre dos personas con una puntuación global semejante, nuestro cuestionario no permite discriminar si entre ellas hay diferente competencia en alguno de los dominios específicos. Por esta razón recomendamos que a la hora de realizar evaluaciones de profesionales que ejercen mentoría en investigación, los resultados de este cuestionario sean principalmente interpretados en forma global, utilizando los puntajes parciales por dominio, a modo orientativo. Queda para futuras investigaciones la posibilidad de modificar el cuestionario con el fin de mejorar su validez divergente.

Aclaramos que, si bien la validez divergente o discriminante no había sido comunicada en la publicación del cuestionario en inglés que dio origen a nuestro trabajo⁸, era posible anticipar este problema en base a las altas correlaciones entre algunos de sus dominios (varias de ellas superiores a 0,8, e incluso una de ellas, superior a 0,9) comunicadas por sus autores.

CONCLUSIONES

En la actualidad la escala *MAC* es utilizada como herramienta de evaluación de diversos programas tendientes al desarrollo de habilidades en mentoría^{13,14} en investigación.

Consideramos que contar con la primera herramienta en español para evaluar de manera cualitativa y en forma global las com-

petencias en mentoría en investigación será de mucha utilidad para nuestra comunidad científica hispanoparlante. Este instrumento será de ayuda para que los programas de formación y entrenamiento de investigadores –p. ej. doctorados, maestrías y/o licenciaturas de carreras que involucren componentes de investigación–, y las agencias que otorgan becas y subsidios de investigación a equipos integrados por investigadores/as noveles supervisados por investigadores experimentados realicen evaluaciones periódicas del desempeño de quienes ejercen los roles de mentoría. En este sentido destacamos que, como sostiene Fernández Fastuca⁴⁰, no es común que en Argentina el desempeño de los directores de tesis (p. ej. doctorados en ciencias biológicas o sociales) sea evaluado de manera formal y/o sistemática. Es en este contexto que esperamos que contar con la versión adaptada transculturalmente y validada en español del *MAC* que estamos presentando –un instrumento breve, autoadministrado y de sencilla aplicación–, ayude a motivar a los equipos responsables de la gestión de educativa de este tipo de programa de formación de jóvenes profesionales en competencias de investigación, a realizar evaluaciones de quienes ejercen mentoría en investigación, lo que permitirá implementar intervenciones de retroalimentación para mejorar dichos procesos.

Consideramos que, al igual que en las experiencias mencionadas^{13,14}, también contribuirá a evaluar el impacto de intervenciones educativas y de entrenamiento en las competencias requeridas para ejercer esta tarea profesional.

Material suplementario

1. Material suplementario 1: Estrategia de búsqueda bibliográfica.
2. Material suplementario 2: Comparación de las instrucciones de llenado del cuestionario y de sus ítems originales con el texto final resultante del proceso de la traducción, retro-traducción y conciliación de las diversas versiones obtenidas, luego de incorporar las sugerencias recolectadas a través de entrevistas cognitivas con personas potencialmente usuarias del cuestionario.
3. Material suplementario 3: Cargas factoriales de la versión original del Cuestionario *Evaluación de las Competencias de Mentoría* publicadas por Fleming et al.⁸, y descripción de las puntuaciones de nuestra muestra de validación.
4. Material suplementario 4: Versión abreviada del Cuestionario *Evaluación de las Competencias de Mentoría*.

Fuentes de financiamiento / Conflicto de interés de los autores: Fundación MF para el Desarrollo de la Medicina Familiar y Departamento de Investigación del Hospital Italiano de Buenos Aires. Los autores declararon no poseer conflictos de interés.

Recibido el 28/06/2022 Aceptado el 09/07/2022
y Publicado el 23/07/2022.

Nota de la editora

Este trabajo de investigación constituye una tesis doctoral, aprobada en el marco del Doctorado en Ciencias de la Salud del Instituto Universitario Hospital Italiano de Buenos Aires. Dado el proceso de evaluación previo realizado por el jurado conformado por Lorena Fernández Fastuca, Viviana Mancovsky e Ingris Peláez, se consideró que el proceso de revisión de pares había sido completado al presentarse el manuscrito para su publicación en la revista.

Referencias

1. Roch GR. Much ado about mentors. *Harv Bus Rev.* 1979;57(1):14–20.
2. Sambunjak D, Straus SE, Marušić A. Mentoring in academic medicine: a systematic review. *JAMA.* 2006;296(9):1103–1118. Available from: 10.1001/jama.296.9.1103.
3. Eby LT, Allen TD, Evans SC, et al. Does Mentoring Matter? A Multidisciplinary Meta-Analysis Comparing Mentored and Non-Mentored Individuals. *J Vocat Behav.* 2008;72(2):254–67. Available from: 10.1016/j.jvb.2007.04.005.

4. Terrasa S, Roni C, Rosli N, et al. Introducción a la mentoría en investigación y principales marcos teóricos para comprender y explicar sus procesos. *Evid Actual Práct Ambul.* 2019;22(1):e002004. Available from: 10.51987/EVIDENCIA.V22i1.4213.
5. Terrasa S, Durante E. Perfiles de mentoría en investigación y evidencia de sus beneficios. *Evid Actual Práct Ambul.* 2019;22(4):e002007. Available from: 10.51987/EVIDENCIA.V22i4.4254.
6. Terrasa S. Marcos teóricos para comprender y explicar los procesos de mentoría en investigación. *Evid Actual Práct Ambul.* 2022;25(2):e007017. Available from: 10.51987/EVIDENCIA.V25i3.7017.
7. Chen Y, Watson R, Hilton A. A review of mentorship measurement tools. *Nurse Educ Today.* 2016;40:20–8. Available from: 10.1016/j.nedt.2016.01.020.
8. Fleming M, House S, Hanson VS, et al. The Mentoring Competency Assessment: validation of a new instrument to evaluate skills of research mentors. *Acad Med.* 2013;88(7):1002–8. Available from: 10.1097/ACM.0b013e318295e298.
9. National Academies of Sciences, Engineering, and Medicine. *The Science of Effective Mentorship in STEMM.* Byars-Winston A, Lund-Dahlberg M, editors. Washington, DC: The National Academies Press; 2019. Available from: 10.17226/25568.
10. Pfund C, House SC, Asquith P, et al. Training Mentors of Clinical and Translational Research Scholars: A Randomized Controlled Trial. *Acad Med.* 2014;89(5):774–82. Available from: 10.1097/ACM.0000000000000218.
11. Johnson MO, Gandhi M. A mentor training program improves mentoring competency for researchers working with early-career investigators from underrepresented backgrounds. *Adv Health Sci Educ Theory Pract.* 2015;20(3):683–9. Available from: 10.1007/s10459-014-9555-z.
12. Camveren H, Vatan F. Psychometric testing of the Turkish version of the mentoring competency assessment scale for faculty. *J Pak Med Assoc.* 2018;68(12):1804–1808.
13. Rogers J, Branchaw J, Weber-Main AM, et al. How much is enough? The Impact of Training Dosage and Previous Mentoring Experience on the Effectiveness of a Research Mentor Training Intervention. *Understanding Interventions.* 2020;11(1):1–17.
14. Branchaw J, Guerrero L, Pfund C. Interventions to Optimize Mentoring Relationships for Diverse Biomedical Researchers. *Understanding Interventions.* 2020;11(1):1–10.
15. Ramada-Rodilla JM, Serra-Pujadas C, Gl DC. Adaptación cultural y validación de cuestionarios de salud: Revisión y recomendaciones metodológicas. *Salud Publica Mex.* 2013;55(1):57–66. Available from: 10.1590/s0036-36342013000100009.
16. Wild D, Grove A, Martin M, et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health.* 2005;8(2):94–104. Available from: 10.1111/j.1524-4733.2005.04054.x.
17. Méndez-Martínez C, Alonso-Rondón-Sepúlveda M. Introducción al análisis factorial exploratorio. *Rev Colomb Psiquiatr.* 2012;41(1):197–207. Available from: 10.1016/S0034-7450(14)60077-9.
18. Hair J, Anderson R, Tatham R, et al. *Multivariate Data Analysis.* Upper Saddle River, NJ, USA: Prentice-Hall, Inc; 1995.
19. Jöreskog K, Sörbom D. *PRELIS 2: User's reference guide.* Chicago: Scientific Software Internacional; 1996.
20. Sterne J, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. *BMJ.* 2009;338:b2393. Available from: 10.1136/bmj.b2393.
21. Useche L, Mesa D. Una introducción a la imputación de valores perdidos. *Terra.* 2006;p. 127–52.
22. Cheng-Hsien L. The performance of MLR, USLMV and WLSMV estimation in structural regression models with ordinal variables; 2014. Available from: <https://doi.org/doi:10.25335/M58979>.
23. Flora DB, Curran PJ. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods.* 2004;9(4):466–91. Available from: 10.1037/1082-989X.9.4.466.
24. Brown T. *Confirmatory Factor Analysis for Applied Research.* In: DA K, editor. *Methodology in the Social Sciences.* New York: The Guilford Press; 2006.
25. Fornell C, Larcker DF. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research.* 1981;18(1):39–39.
26. Streiner DL. Research methods in psychiatry: A checklist for evaluating the usefulness of rating scales. *Canadian Journal of Psychiatry.* 1993;38:140–148.
27. Bland JM, Altman DG. Applying the right statistics: Analyses of measurement studies. *Ultrasound Obstet Gynecol.* 2003;22(1):85–93. Available from: 10.1002/uog.122.
28. L MB. Consentimiento informado: una praxis dialógica para la investigación. *Rev Invest Clin.* 2009;61:73–82.
29. British Educational Research Association [BERA]. *Ethical Guidelines for Educational Research.* London; 2018. Available from: <https://www.bera.ac.uk/researchers-resources/publications/ethical-guidelines-for-educational-research-2018>.
30. Arroyo JLB, and. Los pronombres de tratamiento y la cortesía; 1994.
31. Ramos-González N, Alonso. La incidencia cultural en el uso de los pronombres de cortesía. In: and others, editor. *La enseñanza del español en tiempos de crisis;* 2009. p. 23–26.
32. Marín D. El uso de tú y usted en el español actual. *Hispania.* 1972;55(4):904–912.
33. *Cambridge U of. Cambridge Dictionary.* Cambridge, UK: Cambridge University Press; 2020.
34. Hair J, Babin B, Anderson R, Black W. *Multivariate Data Analysis* Cengage Learning EMEA. 2018;832.
35. Cho YI, Johnson TP, Vangeest JB. *Eval Health Prof.* 2013;36(3):382–407. Available from: 10.1177/0163278713496425.
36. Burke M, Hodgins M. Is Dear colleague enough? Improving response rates in surveys of healthcare professionals. *Nurse Researcher.* 2015;23(1):8–15. Available from: 10.7748/nr.23.1.8.e1339.
37. Scott A, Jeon SH, Joyce CM, et al. A randomised trial and economic evaluation of the effect of response mode on response rate, response bias, and item non-response in a survey of doctors. *BMC Med Res Methodol.* 2011;11:126. Available from: 10.1186/1471-2288-11-126.
38. Braithwaite D, Emery J, De Lusignan S, et al. Using the internet to conduct surveys of health professionals: A valid alternative? *Fam Pract.* 2003;20(5):545–51. Available from: 10.1093/fampra/cmg509.
39. Aikens ML, Sadselia S, Watkins K, et al. A social capital perspective on the mentoring of undergraduate life science researchers: An empirical study of Undergraduate-Postgraduate-Faculty triads. *CBE Life Sci Educ.* 2016;15(2):1–15. Available from: 10.1187/cbe.15-10-0208.
40. Fernandez-Fastuca L. *Pedagogía de la formación doctoral.* 1st ed. Ciudad Autónoma de Buenos Aires: Universidad Abierta Interamericana: Editorial Teseo; 2018.