



ORIGINAL

Validación y adaptación cultural al español de la Rúbrica de retroalimentación del desempeño quirúrgico (SPR)

Jaime Andrés Leal^{a,b,*}, Daniela Gutiérrez Zúñiga^{a,c}, Stephanie Renza^a y Felipe Valbuena^{a,c}^a Departamento de Ortopedia, Hospital Universitario de la Samaritana, Bogotá, Colombia^b Departamento de Ortopedia, Hospital Universitario Mayor-Corporación Méderi, Bogotá, Colombia^c Departamento de Ortopedia, Pontificia Universidad Javeriana, Bogotá, Colombia

Recibido el 23 de junio de 2024; aceptado el 22 de noviembre de 2024

Disponible en Internet el 24 de enero de 2025

PALABRAS CLAVE

Retroalimentación;
Evaluación
educacional;
Rúbrica de
Retroalimentación del
Desempeño Quirúrgico
(SPR);
Educación médica;
Estudio de validación

Resumen

Introducción: la formación de profesionales en medicina ha evolucionado hacia un modelo basado en competencias. En este contexto, la evaluación de competencias quirúrgicas es fundamental para asegurar estándares de calidad y competencia. Sin embargo, en nuestro entorno aún falta una herramienta estandarizada y validada que permita una evaluación efectiva. La Rúbrica de Retroalimentación del Desempeño Quirúrgico (SPR) fue desarrollada para este propósito, pero no ha sido adaptada al español. Este estudio busca traducir, validar y adaptar culturalmente la SPR para su uso en programas de especialidades quirúrgicas en Hispanoamérica.

Materiales y métodos: se llevó a cabo un estudio prospectivo, observacional, de adaptación transcultural y validación psicométrica de la SPR de mayo de 2020 a septiembre de 2021. La validación lingüística incluyó la traducción directa y reversa, revisión por un comité de expertos y pruebas de comprensión. La validación psicométrica evaluó la validez de apariencia, contenido, convergente, consistencia interna y reproducibilidad.

Resultados: se realizaron 160 evaluaciones con la SPR, de las cuales 129 fueron aptas para el análisis y cumplieron los criterios de inclusión y exclusión. El alfa de Cronbach fue de 0,9; indicando alta fiabilidad. El análisis factorial reveló 3 factores latentes relacionados con la atención prequirúrgica, la colaboración y comunicación, y las técnicas quirúrgicas. No se encontraron diferencias significativas en las evaluaciones según el año de residencia o el rol quirúrgico.

Conclusiones: la versión en español de la SPR es una herramienta válida y confiable para la evaluación formativa de residentes en cirugía. Su implementación contribuye a estandarizar la evaluación de competencias quirúrgicas en contextos de habla hispana, facilitando una retroalimentación efectiva y precisa en la formación quirúrgica.

* Autor para correspondencia.

Correo electrónico: Jaime.leal@javeriana.edu.co (J.A. Leal).

KEYWORDS

Feedback;
Educational
measurement;
Surgical Procedure
Feedback Rubric (SPR);
Medical education;
Validation study

Validation and cultural adaptation to Spanish of the surgical performance feedback rubric (SPR)

Abstract

Introduction: The training of medical professionals has shifted towards a competency-based model. In this context, the evaluation of surgical competencies is crucial to ensure standards of quality and competence. However, there is still a lack of standardized and validated tools for effective evaluation in our setting. The "Surgical Procedure Feedback Rubric" (SPR) was developed for this purpose but has not been adapted to Spanish. This study aims to translate, validate, and culturally adapt the SPR for use in surgical specialty programs in Latin America.

Methods: A prospective observational study of cross-cultural adaptation and psychometric validation of the SPR was conducted from May 2020 to September 2021. The linguistic validation included forward and backward translation, review by an expert committee, and comprehension testing. The psychometric validation assessed face, content, convergent validity, internal consistency, and reproducibility.

Results: A total of 160 evaluations using the SPR were conducted, of which 129 were suitable for analysis. Cronbach's alpha was 0,9; indicating high reliability. Factor analysis revealed three latent factors related to pre-surgical care, collaboration and communication, and surgical techniques. No significant differences were found in evaluations based on the year of residency or surgical role.

Conclusions: The Spanish adaptation of the SPR is a valid and reliable tool for the formative assessment of surgical residents. Its implementation supports the standardization of surgical competency evaluation in Spanish-speaking settings, facilitating effective and precise feedback in surgical training.

© 2024 The Author(s). Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introducción

El objetivo principal de la formación de profesionales en medicina se basa en el entrenamiento y el desarrollo de habilidades para poder responder a las necesidades de una sociedad. En los últimos años, y a nivel mundial, el modelo de educación médica ha migrado desde una enseñanza tradicional hacia el modelo de educación basado en competencias¹. Este modelo de enseñanza cambia el enfoque centrado en la duración de un programa académico hacia el logro de hitos fundamentales de aprendizaje, orientados a alcanzar un desempeño profesional competente. Se basa en las Actividades Profesionales Confiabiles (APROC) y en la delegación gradual de responsabilidades².

La evaluación de la adquisición de competencias quirúrgicas, tal como señalan James et al.³, tiene como objetivos fundamentales ofrecer una plataforma de aprendizaje consistente a través de la retroalimentación dirigida al estudiante, valorar la evolución y progresión del aprendizaje a lo largo del entrenamiento quirúrgico, y, en última instancia, asegurar el mantenimiento de los estándares profesionales de calidad y competencia en la práctica quirúrgica. No obstante, y hasta la fecha actual, en nuestro contexto aún se carece de una herramienta estandarizada y formalmente validada que permita evaluar

de manera efectiva y llevar a cabo este proceso de retroalimentación en las especialidades quirúrgicas, con especial énfasis en la valoración del desempeño dentro del quirófano. Esta ausencia subraya la necesidad imperante de desarrollar y validar un instrumento que pueda ser aplicado de manera uniforme, proporcionando así una base sólida para la evaluación y el avance en la formación quirúrgica.

La Rúbrica de Retroalimentación del Desempeño Quirúrgico, (*Surgical Procedure Feedback Rubric [SPR]*), en su idioma original, fue creada por Toprak et al. en 2016⁴. Esta herramienta innovadora está diseñada para evaluar aspectos clave, como la preparación del quirófano, las habilidades técnicas, las competencias intrínsecas y habilidades blandas. Lo que distingue a la SPR es su capacidad para proporcionar no solo una evaluación objetiva del desempeño de los residentes, sino también ofrecer una retroalimentación oportuna por parte del docente respecto al procedimiento quirúrgico. Esto permite una guía precisa en el proceso de aprendizaje del residente, enmarcado dentro de un modelo basado en competencias.

Durante el proceso de desarrollo de esta herramienta y su aplicación en residentes de ortopedia y cirugía general, la SPR demostró poseer una sólida validez de constructo y una capacidad notable para discriminar entre los distintos años de residencia⁴. Este hecho subraya su utilidad práctica y relevancia en el contexto de la formación quirúrgica. Sin

embargo, a pesar de su comprobada eficacia, hasta el momento, el instrumento no ha sido adaptado ni validado para su uso en el idioma español. Esta brecha representa una oportunidad significativa para ampliar la accesibilidad y aplicabilidad de la SPR en entornos de habla hispana, facilitando así una evaluación y retroalimentación aún más inclusivas y efectivas en el ámbito quirúrgico.

Con base en las recomendaciones de Kondo et al. enfatizamos la necesidad de adaptar y validar rúbricas en contextos locales, utilizando un lenguaje claro y acorde con las competencias a evaluar para lograr evaluaciones significativas⁵.

El objetivo de este estudio fue traducir, validar y adaptar sistemáticamente la SPR para su uso en programas de posgrado de especialidades quirúrgicas en Hispanoamérica, y en Colombia en particular, con el propósito de facilitar una evaluación objetiva de los residentes y mejorar los modelos educativos médicos basados en competencias. Este proceso incluyó una traducción rigurosa, validación por expertos y adaptación cultural para asegurar la relevancia y aplicabilidad en nuestro contexto, contribuyendo así a la optimización de la educación quirúrgica en español.

Materiales y métodos

Todos los procedimientos de investigación empleados en este estudio se ajustaron a la Declaración de Helsinki y recibieron la aprobación del Comité de Ética de Investigación del Hospital Universitario de La Samaritana según el Acta de reunión: 03-2020 del 25 marzo de 2020. Este comité clasificó el estudio como una investigación sin riesgo, siguiendo la Resolución N° 008430 de 1993 del Ministerio de Salud de Colombia. Por tanto, no se requirió la solicitud de consentimiento informado.

El estudio se llevó a cabo durante los meses de mayo de 2020 a septiembre de 2021 con un diseño prospectivo observacional de adaptación transcultural y validación psicométrica de la SPR.

El estudio se llevó a cabo mediante las siguientes fases:

Validación lingüística

Basado en las recomendaciones del Instituto MAPI y del consenso ISPOR⁶, la validación lingüística comprendió:

Comité de revisión

Se constituyó un comité de revisión integrado por un experto en idiomas, un especialista en el área quirúrgica y un residente. Este equipo, en colaboración con el grupo de investigadores y antes de proceder con la primera traducción, anticipó posibles desafíos en el proceso de traducción y validación de la rúbrica.

Traducción – traducción reversa

Una vez autorizado el inicio de la traducción, se seleccionaron 2 traductores oficiales para el estudio, quienes eran bilingües y pertenecían a la cultura objetivo de traducción y/o

retraducción. Se llevó a cabo traducción directa al español, resultando en un documento por cada traductor. Estos documentos fueron inicialmente evaluados por el comité de revisión, y se solicitó la opinión de un tercer traductor, quien confirmó su similitud de forma. Posteriormente, se intercambiaron los documentos entre los traductores para su revisión cruzada y se realizó una retraducción al inglés para verificar la fidelidad de la traducción.

Equivalencia

El texto resultante fue revisado nuevamente por el comité de revisión, el cual evaluó su concordancia con la rúbrica original.

Pilotaje o prueba de comprensión

Se analizó la claridad del lenguaje y la facilidad de comprensión de los ítems del cuestionario. Para ello, se seleccionó una muestra de 10 residentes del hospital, a quienes se les aplicó la rúbrica adaptada. Posteriormente, se llevó a cabo entrevistas en profundidad con el objetivo de evaluar la comprensión de las premisas, identificar tópicos que presentaran dificultades de entendimiento y detectar cualquier ambigüedad en el instrumento.

Estos aspectos se consideraron para valorar el tiempo requerido para completar la rúbrica, determinar la necesidad de entrenamiento previo para su uso, examinar las características del formato del instrumento y juzgar la facilidad, tanto en tiempo como en forma, para asignar el puntaje final.

Revisión semántica

Se solicitó a un grupo de 3 docentes de especialidades quirúrgicas, cuya lengua materna es el español pero que poseen un dominio competente del inglés, que examinaran la terminología utilizada en la rúbrica. El objetivo era asegurar que el significado de las distintas frases se mantuviera coherente entre la versión traducida y la original.

Correcciones y versión final del documento

Se realizaron revisiones para identificar errores de contenido y aplicación en el instrumento. La versión final de la rúbrica fue evaluada por los mismos 3 docentes y 10 residentes que participaron en las fases previas, con el propósito de determinar su utilidad. Durante esta evaluación, se registró el tiempo requerido para responder y calificar utilizando la rúbrica, así como las condiciones de su aplicación.

Aplicación del cuestionario

La rúbrica se aplicó a los participantes que cumplían con los criterios de inclusión descritos más adelante; luego de proporcionarles una explicación detallada del procedimiento y obtener su consentimiento verbal informado. Para el *retest*, la rúbrica se volvió a aplicar entre 3 y 14 días después de la primera evaluación.

Validación psicométrica

Para evaluar la validez y confiabilidad de una rúbrica destinada a medir competencias quirúrgicas, se diseñó un proceso de validación estructurado en varias etapas, manteniendo la idea central de garantizar la precisión y la eficacia de la herramienta de evaluación.

Validez de apariencia

En la primera fase, se convocó a un grupo de 5 docentes con especialización en diversas áreas quirúrgicas para realizar una evaluación preliminar del cuestionario. El objetivo era determinar si, desde un análisis superficial o primera impresión, la rúbrica parecía ser capaz de evaluar efectivamente las competencias que pretende medir.

Validez de contenido

Posteriormente, se pidió al mismo grupo de docentes que profundizaran en su análisis para valorar si la rúbrica abarca todos los aspectos esenciales del constructo que se desea medir, para confirmar que no se habían omitido componentes críticos en la construcción de la herramienta y que esta reflejaba adecuadamente el espectro completo de competencias quirúrgicas relevantes.

Consistencia interna

Para cuantificar la estabilidad y coherencia interna del instrumento, se propuso calcular el alfa de Cronbach para la totalidad de la muestra y valorar hasta qué punto los diferentes ítems de la rúbrica miden de manera consistente el mismo constructo, proporcionando una estimación de su fiabilidad.

Validez (convergente)

Se propuso comparar los resultados obtenidos mediante la rúbrica validada con otros instrumentos establecidos y confiables, utilizando el coeficiente de correlación de Spearman. Este análisis de validez convergente permitiría verificar si la rúbrica mide el constructo de interés de manera similar a otras herramientas validadas. Sin embargo, la revisión de la literatura no arrojó resultados sobre herramientas o instrumentos validados en español que evaluaran el mismo constructo.

Confiabilidad (reproducibilidad, *test – retest*)

Para evaluar la reproducibilidad de la rúbrica, se realizaron mediciones hasta 14 días después de la primera evaluación. Utilizando estos datos, se examinó la correlación entre las 2 puntuaciones mediante el coeficiente de correlación de Spearman, lo que permitiría calcular el coeficiente de correlación intraclase y, por ende, la confiabilidad de la rúbrica a lo largo del tiempo.

Concordancia entre observadores

Finalmente, para evaluar la correlación y consistencia entre diferentes evaluadores, se compararon los resultados obtenidos por diferentes observadores utilizando nuevamente el coeficiente de correlación de Spearman. Este análisis era fundamental para determinar si la rúbrica ofrecía resultados consistentes independientemente del evaluador, asegurando así su objetividad y confiabilidad.

Población, tamaño de la muestra y criterios de inclusión/exclusión

Durante el periodo de mayo de 2020 a septiembre de 2021, se incluyeron en el estudio residentes quirúrgicos de primer a cuarto año que estuvieran cumpliendo sus rotaciones en nuestro hospital y que participasen en procedimientos quirúrgicos bajo la supervisión de un especialista. Para la investigación, se calculó un tamaño de muestra de 116 evaluaciones, basándose en la expectativa de alcanzar un coeficiente de correlación de Pearson de 0,9; con un nivel de significación (error alfa) de 0,05 y un poder estadístico (1 - error beta) de 0,8; asumiendo una hipótesis de diferenciación clara en la habilidad evaluada.

Los criterios de inclusión especificados fueron que los residentes estuviesen rotando activamente en el hospital y participasen directamente en alguna intervención quirúrgica supervisada. Se establecieron criterios de exclusión para garantizar la calidad y consistencia de los datos recolectados, excluyendo aquellas evaluaciones de rúbrica no realizadas dentro de las primeras 6 horas después del procedimiento, *retests* aplicados antes de los 3 días o después de los 14 días del procedimiento inicial y rúbricas incompletas o sin los datos adecuadamente diligenciados.

Plan de análisis

En el plan de análisis diseñado para evaluar las calificaciones obtenidas a través de la rúbrica, se asignaron puntajes a las opciones de calificación de la siguiente manera: 0 puntos para respuestas marcadas como «No evaluado o No aplica», 1 punto para «Requiere atención», 2 puntos para «En desarrollo» y 3 puntos para «Logra (cirujano independiente)». Este esquema de calificación permitió una cuantificación objetiva de las respuestas.

Para el análisis de los datos recopilados, se utilizó el software estadístico SPSS en su versión 26 académica, estableciendo un valor de significación alfa de 0,5. Para explorar la estructura subyacente de la rúbrica, se llevó a cabo un análisis factorial exploratorio, aplicando la técnica de factorización del eje principal acompañada de una rotación promax, lo cual facilitó la identificación de patrones de correlación entre las distintas variables evaluadas.

La determinación del número de factores a considerar se basó en el análisis de los valores propios iniciales y la inspección del diagrama de desviación (*scree plot*), lo que permitió establecer una estructura factorial final. Esta

estructura fue posteriormente evaluada en términos de consistencia interna, utilizando el alfa de Cronbach, y se estableció un valor umbral de 0,7 para indicar una consistencia interna aceptable en cada uno de los factores identificados.

Además, para examinar si la rúbrica era capaz de diferenciar entre las actuaciones intraoperatorias de estudiantes en distintos niveles de formación y roles quirúrgicos, se analizó la influencia del año de formación posgrado y el papel desempeñado durante la cirugía en las calificaciones obtenidas. Las puntuaciones factoriales se obtuvieron promediando las calificaciones de los ítems correspondientes a cada factor. Empleando estas puntuaciones factoriales como variable dependiente, se procedió a realizar un análisis de varianza ANOVA de 3 vías, permitiendo una comprensión más profunda de las capacidades y competencias evaluadas por la rúbrica en cuestión.

Resultados

En el transcurso de este estudio se realizaron 160 evaluaciones mediante la rúbrica diseñada; no obstante, se descartaron 31 de ellas por incompletitud o porque el *retest* no se ajustó al intervalo de tiempo previamente definido. Esto dejó un total de 129 rúbricas aptas para el análisis, superando el tamaño de muestra proyectado inicialmente. A través de tablas de frecuencia, se efectuó un análisis descriptivo para las variables cualitativas, revelando que el 48% de las intervenciones correspondieron a reducción abierta y fijación interna (RAFI); el 41% de los residentes evaluados estaban en su tercer año de formación y el 43,4% de los participantes asumieron el rol de cirujano principal (tabla 1).

Posteriormente, se elaboró una tabla de frecuencia para las variables, comparando los puntajes obtenidos en el *pretest* y el *retest* para cada ítem (tabla 2).

La comparación de los puntajes totales antes y después (fig. 1) indicó una diferencia significativa, mostrando un cambio notable en los resultados ($p = 0,037$) (tabla 3). Para evaluar la fiabilidad de la rúbrica en 2 momentos distintos, se calculó el alfa de Cronbach, obteniendo un resultado de 0,9. Este valor elevado sugiere que la escala posee un alto grado de fiabilidad, indicando precisión en el *test* y un error de medida reducido.

Además, para examinar la validez de la herramienta, se efectuó un análisis factorial exploratorio, aplicando el método de extracción por máxima verosimilitud y rotación varimax. El objetivo era identificar los factores que capturan la mayor variabilidad y revelan la asociación entre los 14 atributos. Se detectó una correlación adecuada, con puntajes superiores a 0,7; entre los ítems 1 y 2; 3 y 4; 4 y 6; 5 y 6; 5 y 7; 6 y 7; 6 y 8; 7 y 8; 8 y 9; y 12 y 13.

Posteriormente, la prueba de esfericidad de Bartlett (tabla 4), que examina la hipótesis nula de que las variables en estudio no están correlacionadas en la muestra; mostró valores elevados junto a valores pequeños de significatividad, que sugieren la posibilidad de rechazar la hipótesis nula y concluir que las variables en la muestra están correlacionadas lo suficiente como para proceder al análisis factorial. La significación asociada a la prueba de

Tabla 1 Análisis descriptivo de la rúbrica

		Total	Porcentaje
		<i>n</i>	(%)
<i>Procedimiento</i>	Amputaciones	10	7,8
	Artroplastias	8	6,2
	Artroscopias	2	1,6
	Desbridamientos	23	17,8
	Clavos endomedulares	17	13,2
	Fijadores externos	3	2,3
	Osteosíntesis	63	48,8
	Tejidos blandos, injertos y colgajos	3	2,3
	Total	129	100
<i>Año de residencia</i>	1. ^{er} año	19	14,7
	2. ^{do} año	20	15,5
	3. ^{er} año	54	41,9
	4. ^{to} año	36	27,9
	Total	129	100
<i>Rol en el procedimiento</i>	1. ^{er} ayudante	44	34,1
	2. ^{do} ayudante	25	19,4
	Cirujano principal	56	43,4
	Sin medición	4	3,1
	Total	129	100

esfericidad de Bartlett fue de 0,000; permitiendo descartar la hipótesis nula de no correlación entre variables. El KMO (prueba de Kaiser-Meyer-Olkin) arrojó un valor superior a 0,80; indicando que la matriz de datos es adecuada para el análisis factorial.

La matriz de patrón demostró la contribución única de cada variable al factor (tabla 5). La solución factorial rotada oblicuamente reveló la existencia de 3 factores latentes que agrupan todas las variables. Un primer factor incluyó variables de atención prequirúrgica, plan posquirúrgico y enseñanza; un segundo factor englobó colaboración y comunicación; y un tercer factor abarcó técnicas quirúrgicas y aplicación del procedimiento, concordando con la estructura de la rúbrica original (tabla 6).

Con las variables identificadas para cada factor, se calcularon los puntajes promedio y se utilizaron como variables dependientes en un análisis de varianza ANOVA de 3 vías, investigando si el tipo de procedimiento, el año de residencia y el rol influían en el puntaje del factor. La prueba F no arrojó resultados significativos para la mayoría de las categorías, rechazando la hipótesis nula de igualdad de puntajes promedio. Solo el ANOVA fue significativo para el factor 2 (ítems 12, 13 y 11), indicando que el tipo de procedimiento sí afectaba el puntaje promedio.

Este análisis minucioso, apoyado por el análisis factorial exploratorio, confirma que la rúbrica mide eficazmente 3 factores cruciales. No obstante, no demostró una capacidad discriminatoria por año de residencia, clasificación y rol en el procedimiento, lo que sugiere que, aunque la rúbrica es valiosa para evaluar aspectos específicos como la preparación prequirúrgica y la técnica quirúrgica, podría necesitar ajustes para evaluar efectivamente las diferencias entre los niveles de formación o roles quirúrgicos de manera más precisa.

Tabla 2 Tabla de frecuencia de variables *pretest* y *retest*

			<i>Pretest</i>		<i>Retest</i>	
			Total	Porcentaje	Total	Porcentaje
			<i>n</i>	%	<i>n</i>	%
<i>Preparación prequirúrgica</i>	Historia clínica e imágenes	No evaluado	9	7	10	7,8
		Requiere atención	3	2,3	2	1,6
		En desarrollo	27	20,9	31	24
		Logra	90	69,8	86	66,7
		Total	129	100	129	100
	Conocimiento del procedimiento	No evaluado	9	7	10	7,8
		Requiere atención	1	0,8	2	1,6
		En desarrollo	32	24,8	35	27,1
		Logra	87	67,4	82	63,6
		Total	129	100	129	100
<i>Razonamiento quirúrgico</i>	Incisión	Requiere atención	2	1,6	1	0,8
		En desarrollo	77	59,7	76	58,9
		Logra	50	38,8	52	40,3
		Total	129	100	129	100
	Plan quirúrgico	Requiere atención	7	5,4	4	3,1
		En desarrollo	64	49,6	68	52,7
		Logra	58	45	57	44,2
		Total	129	100	129	100
	<i>Técnica quirúrgica y aplicación al procedimiento</i>	Requiere atención	2	1,6	4	3,1
		En desarrollo	43	33,3	52	40,3
		Logra	84	65,1	73	56,6
		Total	129	100	129	100
<i>Plan posquirúrgico</i>	Planos tisulares	Requiere atención	0	0	4	3,1
		En desarrollo	57	44,2	57	44,2
		Logra	72	55,8	68	52,7
		Total	129	100	129	100
	Manejo de tejidos	Requiere atención	3	2,3	5	3,9
		En desarrollo	44	34,1	48	37,2
		Logra	82	63,6	76	58,9
		Total	129	100	129	100
	Sangrado	Requiere atención	1	0,8	5	3,9
		En desarrollo	63	48,9	62	48,1
		Logra	65	50,4	61	47,3
		Total	129	100	129	100
	Técnica quirúrgica	Requiere atención	5	3,9	8	6,2
		En desarrollo	66	51,2	67	51,9
		Logra	58	44,9	56	41,9
		Total	129	100	129	100
	Plan posquirúrgico	No evaluado	10	7,7	7	5,4
		Requiere atención	1	0,8	1	0,8
		En desarrollo	34	26,4	39	30,2
		Logra	84	65,1	82	63,6
		Total	129	100	129	100
<i>Comunicación</i>	Comunicación	Requiere atención	1	0,8	0	0
		En desarrollo	17	13,2	23	17,8
		Logra	111	86	105	81,4
		Perdidos	0	0	1	0,8
		Total	129	100	129	100
<i>Colaboración</i>	Retroalimentación	Requiere atención	1	0,8	1	0,8
		En desarrollo	22	17,1	29	22,5
		Logra	104	80,6	98	76
		Perdidos	2	1,6	1	0,8
		Total	129	100	129	100

Tabla 2 (continuación)

			Pretest		Retest	
			Total	Porcentaje	Total	Porcentaje
			n	%	n	%
Enseñanza y tutoría	Respeto	Requiere atención	1	0,8	1	0,8
		En desarrollo	19	14,7	23	17,8
		Logra	109	84,5	105	81,4
		Total	129	100	129	100
	Enseñanza y tutoría	Requiere atención	5	3,9	7	5,4
		En desarrollo	37	28,7	41	31,8
		Logra	54	41,8	48	37,2
		No evaluado	33	25,6	33	25,6
		Total	129	100	129	100

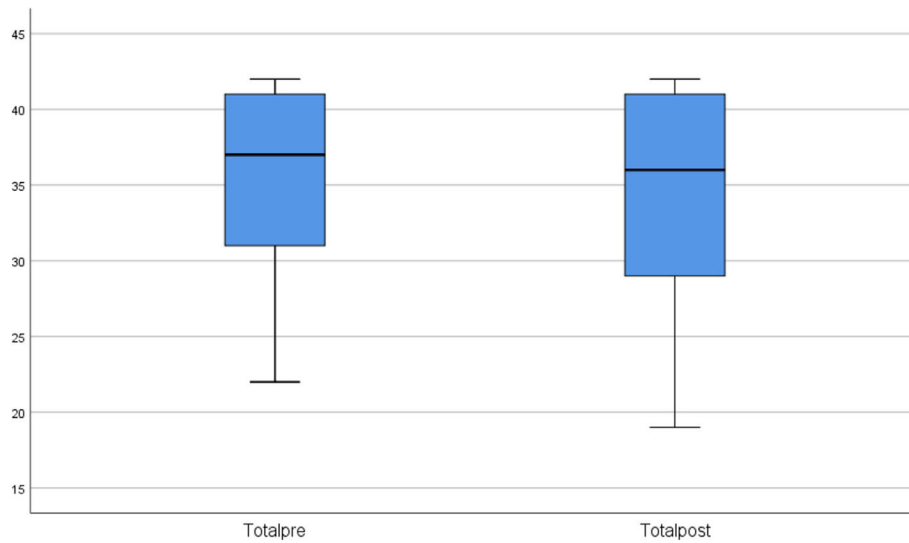


Figura 1 Diagrama de cajas y bigotes de la diferencia del puntaje total entre la rúbrica pre- y pos.

Tabla 3 Prueba de muestras emparejadas de puntaje total rúbrica pre- y pos

		Media		N	Desv. Desviación		Desv. Error promedio		
Par 1	Total pre	35,36		129	5684		0,5		
	Total pos	34,93		129	6095		0,537		
		Diferencias emparejadas					t	gl	Sig. (bilateral)
		Media	Desv. Desviación	Desv. Error promedio	95% de intervalo de confianza de la diferencia				
						Inferior	Superior		
Par 1	Total pre	0,426	2294	0,202	0,027	0,826	2111	128	0,037
	Total pos								

Tabla 4 Test de esfericidad de Barlett y prueba de Kaiser-Meyer-Olkin (KMO)

Prueba de KMO y Bartlett		
Medida Kaiser-Meyer-Olkin de adecuación de muestreo		0,816
Prueba de esfericidad de Bartlett	Aprox. chi-cuadrado	1332,366
	gl	91
	Sig.	0

Discusión

En este estudio presentamos la validación al español de la SPR, desarrollada originalmente por Toprak et al. en el Departamento de Cirugía de la Universidad de Queens⁴. El proceso de validación se llevó a cabo por un equipo compuesto por 2 docentes, una epidemióloga y una residente del área quirúrgica. La rúbrica final en español se presenta en el Anexo 1.

Este estudio presenta la adaptación y validación al español de la SPR, una herramienta diseñada específicamente para evaluar competencias quirúrgicas en el contexto intraoperatorio. La validación de la SPR en español es un avance importante, dado que en Hispanoamérica no se cuenta con instrumentos previamente adaptados que evalúen de manera estandarizada y objetiva el desempeño quirúrgico de los residentes.

Los resultados de este estudio sugieren que la SPR adaptada mantiene una alta consistencia interna (alfa de Cronbach = 0,9), y su análisis factorial exploratorio confirmó la agrupación en 3 dimensiones esenciales del desempeño quirúrgico: atención prequirúrgica, colaboración y comunicación, y técnicas quirúrgicas. Este análisis también mostró que, aunque la SPR es adecuada para la evaluación formativa, tiene limitaciones para diferenciar entre niveles de formación o roles específicos en el equipo quirúrgico, lo

Tabla 5 Matriz de patrón factorial

	Factor		
	1	2	3
Prequirúrgico – historia clínica	1037		
Prequirúrgico - conocimiento	1011		
Plan posquirúrgico	0,303		0,178
Enseñanza y tutoría	0,137		
Colaboración - recepción		1034	
Colaboración - trato		0,915	
Comunicación		0,515	0,181
Técnica quirúrgica – manejo de tejidos	–0,103		0,946
Técnica quirúrgica - planos tisulares			0,913
Técnica quirúrgica – control sangrado			0,827
Técnica quirúrgica - preparación			0,781
Técnica quirúrgica - direccionamiento			0,756
Razonamiento quirúrgico – plan quirúrgico			0,735
Razonamiento quirúrgico – incisión			0,574

Método de extracción: máxima verosimilitud.

Método de rotación: Oblimin con normalización Kaiser.

Tabla 6 Matriz de correlaciones factorial

Factor	1	2	3
1	1	0,32	0,49
2	0,32	1	0,47
3	0,49	0,47	1

Método de extracción: máxima verosimilitud.

Método de rotación: Oblimin con normalización Kaiser.

cual podría indicar una utilidad más centrada en el seguimiento del progreso individual en habilidades prácticas.

Es importante señalar que una limitación de este estudio es que todos los procedimientos evaluados con la SPR se realizaron en el área de ortopedia, y que casi la mitad fueron cirugías de traumatología. Aunque no consideramos que esto afecte la aplicabilidad de la SPR en otras especialidades quirúrgicas, es un aspecto relevante a tener en cuenta. Se recomienda que futuras investigaciones evalúen el desempeño de la SPR en otras áreas quirúrgicas para confirmar su validez y utilidad generalizadas.

La implementación de la SPR en entornos de habla hispana facilita el proceso de retroalimentación inmediata y precisa, esencial para el aprendizaje en quirófano. Su estructura proporciona un marco claro y accesible para evaluar habilidades técnicas y competencias de comunicación, promoviendo una interacción de retroalimentación dirigida entre docentes y residentes.

Dado que en la literatura no existen rúbricas validadas en español que evalúen competencias quirúrgicas de esta forma, la SPR se presenta como una opción viable y necesaria para estandarizar la evaluación en programas de formación quirúrgica. Además, su aplicación es intuitiva y requiere un tiempo mínimo de entrenamiento para los evaluadores, lo que podría favorecer su adopción generalizada en el ámbito de la educación médica quirúrgica.

En conclusión, la adaptación cultural de la SPR al español contribuye significativamente a reducir la brecha en la disponibilidad de herramientas de evaluación estandarizadas en nuestro idioma, especialmente en el ámbito quirúrgico. Este estudio no solo enriquece el repertorio de recursos en español para la educación médica de posgrado, sino que también proporciona una base sólida para implementar prácticas de evaluación más precisas y efectivas en el quirófano. Los resultados confirman que la versión en español de la SPR es una herramienta confiable y útil para la evaluación formativa de residentes en programas de cirugía, subrayando la importancia de contar con instrumentos validados culturalmente en el idioma local para optimizar las prácticas evaluativas y fortalecer la formación de competencias en el contexto quirúrgico.

Responsabilidades éticas

Todos los procedimientos de investigación empleados en este estudio se ajustaron a la Declaración de Helsinki y recibieron la aprobación del Comité de Ética de Investigación del Hospital Universitario de la Samaritana según el Acta de reunión: 03–2020 del 25 marzo de 2020.

Este comité clasificó el estudio como una investigación sin riesgo para los participantes, siguiendo la Resolución

N° 008430 de 1993 del Ministerio de Salud de Colombia. Por tanto, no se requirió la solicitud de consentimiento informado.

Financiación

La financiación de este estudio se realizó con recursos propios de los investigadores, no se recibió ni se requirió financiación externa.

Contribuciones

Concepción, desarrollo inicial y diseño de protocolo: **Jaime Andrés Leal, Daniela Gutiérrez Zúñiga, Stephanie Renza, Felipe Valbuena**. Recolección datos: **Jaime Andrés Leal, Daniela Gutiérrez Zúñiga y Felipe Valbuena**. Análisis estadístico: **Stephanie Renza**. Análisis e interpretación de resultados: **Jaime Andrés Leal, Daniela Gutiérrez Zúñiga, Stephanie Renza, Felipe Valbuena**. Escritura y aprobación de la versión final: **Jaime Andrés Leal, Daniela Gutiérrez Zúñiga, Stephanie Renza, Felipe Valbuena**. Asunción de la responsabilidad frente a todos los aspectos del manuscrito: **Jaime Andrés Leal, Daniela Gutiérrez Zúñiga, Stephanie Renza, Felipe Valbuena**.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses declarado para este estudio.

Declaración sobre el uso de IA generativa y tecnologías asistidas por IA en el proceso de redacción

Durante la preparación de este trabajo, los autores utilizaron OpenAI – ChatGPT 4o para mejorar la gramática.

Después de utilizar esta herramienta/servicio, los autores revisaron y editaron el contenido según fue necesario y asumen la plena responsabilidad por el contenido de la publicación.

Anexo A. Dato suplementario

Se puede consultar material adicional a este artículo en su versión electrónica disponible en <https://doi.org/10.1016/j.edumed.2024.101022>.

Bibliografía

1. Iobst WF, Sherbino J, Cate OT, Richardson DL, Dath D, Swing SR, et al. Competency-based medical education in postgraduate medical education. *Med Teach*. 2010;32(8):651–6.
2. Torres C, Goñy L, Muñoz N, Drago P. Entrustable professional activities: una propuesta innovadora para la evaluación de competencias médicas. *Rev Méd Chile*. 2018;146(9):1064–9.
3. James HK, Chapman AW, Pattison GTR, Fisher JD, Griffin DR. Analysis of tools used in assessing technical skills and operative competence in trauma and orthopaedic surgical training: a systematic review. *JBS Rev*. 2020;8(6) e19.00167-e19.00167.
4. Toprak A, Luhanga U, Jones S, Winthrop A, McEwen L. Validation of a novel intraoperative assessment tool: the surgical procedure feedback rubric. *Am J Surg*. 2016;211(2):369–76.
5. Kondo T, Nishigori H, van der Vleuten C. Locally adapting generic rubrics for the implementation of outcome-based medical education: a mixed-methods approach. *BMC Med Educ*. 2022;22(1):262.
6. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, 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.