



Validación de contenido de una rúbrica para evaluar la competencia oral de una lengua extranjera desde la socioformación

Validation of content of a rubric to evaluate the oral competence of a foreign language from the socioformative approach

Ismaila Mounkoro,¹ María Guadalupe Veytia Bucheli²

Centro Universitario Cife, Mexico¹
Universidad Autónoma del Estado de Hidalgo, Mexico²

Orcid ID: orcid.org/0000-0003-4198-9583¹

Orcid ID : orcid.org/0000-0002-1395-1644²

Recibido: 24 de diciembre de 2018

Aceptado: 27 de junio de 2019

Resumen

Esta investigación surge de la carencia de herramientas para calificar la competencia oral de una lengua extranjera que sea pertinente a las apuestas de la sociedad del conocimiento sujeto al enfoque de la socioformación. El propósito fue realizar la validación de contenido de una herramienta para calificar la competencia oral de una lengua extranjera en el nivel B1. Para esto se empleó un estudio instrumental según las siguientes fases: 1) elaboración de una herramienta a través de un esquema tipo socioformativo; 2) revisión y mejora de la herramienta por parte de tres expertos; 3) validación de contenido por un jurado de 16 expertos en el campo. Los resultados obtenidos indican que la herramienta es adecuada para calificar la competencia oral de una lengua extranjera, con una V de Aiken superior a V 0.75. La Confiabilidad inicial fue de 0.842. Conclusiones: la rúbrica es adecuada; sin embargo, hay que implementarla a un público más amplio, con el fin de definir la validez del constructo.

Palabras clave: evaluación socioformativa, lengua extranjera, idiomas, rúbrica socioformativa, validez, educación

¹**Correspondencia del autor:**

E-mail: mounkorosoumi@hotmail.com

Abstract

This research arises from the lack of tools to qualify the oral competence of a foreign language. That is relevant to the bets of the knowledge society subject to the socioformative approach. The purpose was to validate the content of a tool to qualify the oral competence of a foreign language at level B1. An instrumental study was used according to the following stages: 1) development of a tool through a socioformative type scheme; 2) review and improvement of the tool by three experts; 3) validation of content by a jury of 16 experts in the field. The results obtained indicate that the tool is suitable for qualifying the oral competence of a foreign language, with an Aiken V higher than V 0.75. The initial reliability was 0.842. Conclusions: the rubric is adequate; however, it must be implemented to a wider public in order to define the construct validity

Key words: socioformative evaluation, foreign language, Languages, socioformative rubric, validity, education

Introducción

Considerada tradicionalmente como la meta final del ciclo de aprendizaje, la evaluación es un procesamiento para decidir respecto a las informaciones recogidas sobre el desempeño de los alumnos (Arias & Maturana, 2005). Es como la referencia en el que se apoya el proceso de enseñanza-aprendizaje, su objetivo es analizar la cantidad de adquisición de las metas didácticas que el maestro espera que los alumnos alcancen (Bores & Camacho, 2014), sin embargo, según el marco común europeo para las lenguas (2002) evaluar va más allá de solo medir o valorar. Según la socioformación, el concepto de evaluación se considera como un proceso que pretende la mejoría de forma permanente, identificando logros y puntos de mejora en la resolución de problemas en un contexto determinado (Tobón, 2013b).

Según Martínez (2016), diversos autores consideran relevante la evaluación del desempeño; sin embargo, Tobón, Guzmán, & Tobón (2018), mencionan que frecuentemente, la evaluación se queda solo en el análisis de los trabajos, y se deja de lado el rendimiento académico. Además, la evaluación está poco orientada hacia la mejora, continúan Tobón, Guzmán, & Tobón, (2018). Esas dificultades de evaluación se extienden a todas las áreas del proceso enseñanza-aprendizaje, incluso en los idiomas. La falta de coherencia en la evaluación, las insuficiencias de los instrumentos a evaluar adecuadamente y las confusiones que presenten son unos ejemplos significativos de esas dificultades (Barquero & Ureña, 2015).

El marco común europeo para las lenguas (2012) fue en su tiempo la solución adecuada a esas insuficiencias. Ofrecían ciertas ventajas como: contribuir a definir las estructuras oficiales de evaluación, facilitar el reconocimiento de las competencias en el mundo del trabajo, apoyar en la elaboración de estrategias de aprendizaje, incitar a la autoevaluación y a la autonomía. Gracias a esas ventajas que ofrecía, fue rápidamente adoptado por la comunidad educativa y también por todas las instituciones y gobiernos que sea en Europa, Estados Unidos o Asia.

Sin embargo, el éxito del marco común europeo para las lenguas (2012) fue de muy poca duración, y empezó a generar problemas y polémicas, y la evaluación fue uno de los ámbitos que ha generado más polémicas. De esas polémicas se pueden mencionar la falta de contextualización, la complejidad que representa poner en práctica sus propuestas y la carencia de un instrumento de evaluación adaptado a la sociedad del conocimiento (Figueras, 2008). Aparte de esas polémicas, la evaluación de la competencia oral está considerada como un proceso complejo porque intervienen una serie de situaciones que no dependen precisamente de los alumnos: el nerviosismo, el cansancio o estado de ánimo de los examinadores (Bores, & Camacho, 2014).

De allí surge la necesidad de crear un nuevo instrumento válido, fiable, auténtico, y de impacto. Uno de ellos es la rúbrica socioformativa, definida como la guía precisa para valorar los aprendizajes y las producciones hechas por los estudiantes. Además, ella permite la clara identificación de la importancia de los contenidos y metas de los trabajos (Gómez & Tobón, 2018). Gatica & Uribarren (2013) mencionan que con estas rúbricas los campos considerados como subjetivos o carentes de precisión son fácilmente valorados, eso gracias a la cualificación progresiva de los logros por medio de criterios en él; asimismo, las rúbricas valoran desde un nivel principiante hasta un nivel avanzando de aprendizaje (Bores & Camacho, 2014).

Tomando en cuenta lo antes mencionado, este trabajo se basó en los objetivos sucesivos:

- 1) Elaborar una herramienta adecuada y pragmática para calificar la expresión oral de una lengua extranjera, considerando las apuestas de la sociedad del conocimiento y una visión socioformativa;

- 2) Elaborar la validación de contenido de la herramienta por expertos, para definir su nivel de importancia y congruencia teórica.

Se consideró importante abordar la competencia oral y el nivel B puesto que de las competencias básicas de la lingüística aplicada (expresión oral, comprensión oral, expresión escrita y comprensión escrita), las competencias orales son las que presentan mayores desafíos en cuanto a su adquisición y desarrollo (Ellis, 2003; Nakatani, 2010) y el nivel B1 es vital porque corresponde a usuarios independientes con el idioma, es decir, que cuentan con la fluidez necesaria para comunicarse sin esfuerzo con hablantes nativos (Consejo europeo, 2002).

Etapas

El estudio de validez de la herramienta se implementó de las siguientes maneras:

1. Creación de la herramienta y consideración del jurado experto

Anteriormente a la elaboración se hizo una investigación de herramientas que calificaran la expresión oral de un idioma extranjero en sitios tales como IRESIE, Google Scholar, Redalyc, WoS, SCIELO y SCOPUS. Como resultado, se ha encontrado una cuantiosa cantidad de rúbricas que evalúan la competencia oral de lenguas extranjeras, las más comunes son las rubricas para evaluar el Diploma de Estudio de Lengua francesa, rubricas para el Test de Evaluación de Francés y las rubricas para calificar el test Cambridge; sin embargo, no se encontró ninguna información sobre sus procesos de validación.

Ya finalizada la investigación, se elaboró la herramienta tomando como referencias lo siguiente: 1) metodología de las rubricas desde la socioformación (Hernández, Tobón, 2017a, 2017b; Tobón et al, 2017); taxonomía socioformativa (Tobón, 2017a); criterios para calificar la expresión oral y escrita de una lengua extranjera (Bores & Camacho, 2014). Acto seguido se elaboró la siguiente herramienta con un contenido de 5 ítems y una sola dimensión: la producción oral, y después se pidió la ayuda de un jurado experto con el fin de mejorar la herramienta. Estos tres expertos son referentes en la elaboración de criterios para calificar los idiomas extranjeros con más de una década de trabajo en dicho campo. Ellos revisaron la rúbrica mediante la opción de formularios de Google, se pidió la plantilla a CIFE y se mandó el link a los expertos; la meta fue mejorar la herramienta de manera significativa

y definitiva para ser aplicada. Además, se llenó una encuesta de Factores Sociodemográficos creado por CIFE (2015), teniendo como objetivo recopilar información sociodemográfica de los expertos en cuanto a la edad, sexo, nivel académico y sueldos. En este estudio se consideraron los datos sociodemográficos y la experiencia de formación académica.

Tabla N° 1

Datos sociodemográficos de los expertos

Expertos 3	Características	Datos
	Función	Docentes:0.03%
		Directivos:0.06 %
	Ultimo grado académico:	Maestría: 0.06%
		Doctorado: 0.03%
	Campo de experiencia profesional:	Educación/evaluación
	de años de experiencia profesional:	entre 10 a 3 años
	artículos publicados en el campo	1 a 2
	ponencias presentadas en el área	entre 2 a 10
	Experiencia en la revisión, diseño y/o validación de un determinado instrumento de investigación	más de 10 años

Elaboración propia

2. Estudio de la validez de contenido

Se llevó a cabo por medio de 15 jueces evaluadores (ver Tabla 2). El estudio de la calificación cuantitativa se determinó con el coeficiente de validez de contenido de la V de Aiken (Aiken, 1980) tomando 0.75 como valor mínimo. Se eligió el jurado a partir de estos requisitos: mínimo tener una maestría y/o un doctorado; experiencia en campo de investigación, mínima 10 años; cantidad de publicaciones, y experiencia en la elaboración y/o validación de herramientas de investigación.

Tabla N° 2

Datos de la competencia de los jueces

	Características	Datos
Jueces	Sexo	70% hombres
		30% mujeres
	Role	Docentes:65%
		calidad de vida:5%
		Formación 22%
		proceso organizacionales 8%
		Directivos:7%
	Último nivel de estudio:	Maestría: 85%
		Doctorado:15%
	Áreas de experiencia profesional:	Docencias, directivos, diseño de instrumentos :
	Número de años de experiencia profesional:	5 a 38
	Número de artículos publicados en el área:	0 a 20
	Número de libros publicados en el área:	0 a 7
	Número de ponencias presentadas en el área	5 a 100
Experiencia en la revisión,	más de 10 años	
diseño y/o validación de un determinado instrumento de Investigación	100%	

Elaboración propia

3. Práctica de la prueba con un grupo piloto

Una vez que el jurado experto evaluó la herramienta, inicio la práctica a un grupo piloto de 17 personas, todos ellos docentes de idiomas con más de 1 año de experiencia y en su mayoría examinadores de certificaciones oficiales en lenguas extranjeras y que tengan como mínimo la licenciatura. Se evaluó a estudiantes del nivel b1 y después respondieron a una encuesta de satisfacción del instrumento sobre redacción y claridad del mismo. Se definió la confiabilidad de la rúbrica por medio del coeficiente de Alfa de Cronbach. En la Tabla 3 se describen las características de esta muestra.

Tabla N° 3

Datos sociodemográficos del grupo piloto

Grupo piloto 17	Características	Datos
	Sexo	70 % hombres
		30% mujeres
	Estado civil	96 % soltero /soltera
		2% unión libre
		1% Divorciado/a
		2% casado/casada
	Ocupación	100 Docente

Elaboración propia

Aspectos éticos

Según las normas de formato APA (6ª edición), toda investigación subyace en 3 principios éticos y legales, uno de ellos es la protección de los derechos y el bienestar de los participantes en la investigación (Código de Ética APA; APA, 2002). Todos los participantes en nuestro estudio lo hicieron de forma voluntaria y se les aseguró la confidencialidad y el anonimato respectivo.

Resultados

Se presentan en dos etapas: primero, los resultados e interpretación estadística del jurado y luego los del grupo piloto.

1.- Juicio de expertos

Considerando las opiniones del jurado se determinó que los indicadores de la herramienta son adecuados. A pesar de ello, hubo comentarios en cuanto a la redacción, tales como la ortografía, presentar una redacción alternativa, y suprimir algunas palabras. Para resumir el análisis cuantitativo como cualitativa a través de la V de Aiken, se aprobó la validez de los indicadores.

Tabla N° 4

Resultados del análisis cuantitativo de la V de Aiken

	Media	V de Aiken
--	-------	------------

	Pertinencia	Redacción	Pertinencia	Redacción
Ítem 1	3.8	3.27	0.933	0.755
Ítem 2	3.67	3.6	0.888	0.866
Ítem 3	3.8	3.6	0.933	0.866
Ítem 4	3.74	3.6	0.911	0.866
Ítem 5	3.89	3.87	0.955	0.955

Elaboración propia

2. Grupo piloto

En esa etapa se calificó la satisfacción con la herramienta y el nivel de aprehensión de las indicaciones y de los ítems. El nivel de aprehensión de las indicaciones, preguntas y criterios fue bueno en general, lo mismo en cuanto al nivel de satisfacción con la herramienta (Tabla 5). Además, se comprobó una alta consistencia interna de la herramienta en este grupo (Alfa de Cronbach 0.842).

Tabla N° 5

Nivel de satisfacción de la herramienta

Pregunta	Bajo grado	Aceptable grado	Buen grado	Excelente grado
1. ¿El nivel de comprensión de las indicaciones de la herramienta es?	-	-	95%	5%
2. ¿El nivel de aprehensión de las preguntas o ítems es?	-	71%	12%	17%
3. ¿Cuál fue el nivel de satisfacción con la herramienta?	6%	6%	70%	18%
4. ¿Cuál es el nivel de importancia de las preguntas?	-	11%	59%	30%

Elaboración propia

Discusión

Se puede afirmar, luego de esta investigación, que la herramienta es adecuada para calificar la expresión oral de un idioma extranjero en el nivel B1 considerando el juicio de expertos, llevado a cabo por investigadores académicos expertos en el campo. Cierta número de trabajos científicos respaldan la relevancia del juicio de expertos para determinar la adecuación de las herramientas (de Arquer, 2011; Gorina, 2014), puesto que a menudo se elaboran herramientas poco adecuadas (Garrote, Rojas, & Carmen, 2015; Alarcón, Aidé, Trápaga, Arturo, & Rubén, 2017).

Además, se puede determinar que la herramienta está elaborada correctamente y es entendible para los posibles usuarios, considerando los resultados positivos que se reflejaron tanto en el grupo piloto como el jurado experto. Un número significativo de científicos enfatizan una atenta lectura de las preguntas y las posibles respuestas para obtener una herramienta con alta validez y confiabilidad (Lloret, Ferreres, Hernández, y Tomás, 2014). A pesar que dicho trabajo es relevante para la creación de una herramienta altamente confiable para calificar la expresión oral desde un punto de vista socioformativo—y de la sociedad del conocimiento—cabe mencionar la necesidad de promoción de otros trabajos de validez de constructo y confiabilidad con una encuesta dirigida a una población más amplia, incluyendo todos los niveles y competencias de las lenguas extranjeras.

Referencias

- American Psychological Association, American Educational Research Association and National Council on Measurement in Education (1954,1985). *Standards for Educational and Psychological Testing*. Washington, DC: Author.
- Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. *Educational and Psychological Measurement* 40, 955–959. Recuperado de: goo.gl/X1GUK1
- Alarcón, G., Aidé, L., Trápaga, B., Arturo, J., & Rubén, E, N (2017). Validez de contenido por juicio de expertos: propuesta de una herramienta virtual. *Apertura*, 9 (2), 42-53. <http://dx.doi.org/10.18381/Ap.v9n2.993>
- Arias, C-I. & Maturana, L- M. (2005). Evaluación en lenguas extranjeras: discursos y prácticas. *IKALA, Revista de Lenguaje y Cultura*, 10 (16), 63-91. Recuperado de: <http://goo.gl/rDriUo>
- Barquero, M & Ureña, E. (2015). Rúbricas para evaluar la competencia oral en un segundo idioma: Un estudio de caso. *InterSedes*, XVI, 34. Recuperado de: goo.gl/uuHgyd

- Bores, M. & Camacho, L. (2016): “Criterios para evaluar la expresión oral y escrita en la clase de español”. LI Congreso Internacional AEPE. Palencia, julio 25-29, 2016. Recuperado de: <http://goo.gl/JCZAfU>
- Cabero Almenara, J. y Llorente Cejudo, M. C. (2013), La aplicación del juicio de experto como técnica de evaluación de las tecnologías de la información (TIC). *Revista de Tecnología de Información y Comunicación en Educación*, 7 (2), 11-22. Recuperado de: goo.gl/gcYkcC
- Celina H. y Campo A. (2005). “Aproximación al uso del coeficiente alfa de Cronbach”, *Revista colombiana de psiquiatría*, vol. XXXIV, número 004, 572 – 580. Extraído desde goo.gl/hv1PvD
- Cohen, R. & Swerdlik, M. (2001). Pruebas y evaluación psicológicas: Introducción a las pruebas y a la medición. (4ª. ed.). México: McGraw-Hill. Recuperado de: goo.gl/cszXmD
- Consejo de Europa. (2002). Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación. Madrid: Ministerio de Educación, Cultura y Deporte, Instituto Cervantes, Anaya, 2002. Recuperado de: goo.gl/e572MU
- Corral, Y. (2009). Validez y confiabilidad de los instrumentos de investigación para la recolección de datos. *Revista ciencias de la educación*, 19(33),228-247. Recuperado de: goo.gl/YoorvT
- Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment research & Evaluation*, 10(7), 1-9.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press. Haozhang,
- De Arquer, M. I. (2011). NTP 401: Fiabilidad humana: métodos de cuantificación, juicio de expertos. *Centro nacional de condiciones de trabajo, España*. Recuperado de http://www.oect.es/inshtweb/contenidos/documentacion/fichastecnicas/ntp/ficheros/401a500/ntp_401.pdf
- Figueras, N. (2008). El MCER, más allá de la polémica. *MarcoELE*, 7, 26-34. Recuperado de: goo.gl/A1d1Qf
- Gatica, F. & Uribarren, T. (2013). ¿Cómo elaborar una rúbrica? *Investigación en Educación Médica* 2(1):61-65. Recuperado de: goo.gl/cszXmD
- Gómez-González, J. A. & Tobón, S. (2018): “Diseño y validación de un Instrumento para evaluar el impacto de los proyectos formativos en educación básica en México”, *Revista Atlante: Cuadernos de Educación y Desarrollo*, mayo. Recuperado de: <http://goo.gl/GjMKZK>

- Garrote, R., Rojas, P., & Carmen, M. (2015). La validación por juicio de expertos: dos investigaciones cualitativas en lingüística aplicada. *Revista Nebrija de Lingüística Aplicada*, núm. 18. Recuperado de https://www.nebrija.com/revista-linguistica/files/articulosPDF/articulo_55002aca89c37.pdf
- Gorina, S., A. (2014). La gestión de la información científica proporcionada por el criterio de expertos. *Ciencias de la Información*, 45(2), 39-47. Recuperado de: <https://www.redalyc.org/html/1814/181432443007/>
- Hernández, J. S., Tobón, S., González, L., y Guzmán, C. (2015). Evaluación socioformativa y rendimiento académico en un programa de posgrado en línea. *Paradigma*, XXXVI, 30-41. Recuperado de: <https://goo.gl/gWcYTx>
- Hernández, J. S., Tobón, S., Ortega, M. F., & Ramírez, A. M. (2018). Evaluación socioformativa en procesos de formación en línea mediante proyectos formativos. *Educación*, 54(1), 147-163.
- Hernández-Mosqueda, J., S., Tobón-Tobón, S., y Guerrero-Rosas, G. (2016). Hacia una evaluación integral del desempeño: las rúbricas socioformativas. *Revista Ra Ximhai*, 12, 359-376. Recuperado de: goo.gl/etWLxC
- Jose-Antonio Gomez-Gonzalez y Sergio Tobón Tobón (2018). Diseño y validación de un Instrumento para evaluar el impacto de los proyectos formativos en educación básica en México. *Revista Atlante: Cuadernos de Educación y Desarrollo* (mayo). Recuperado de: goo.gl/GjMKZK
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Anales de Psicología*, 30(3), 1151-1169.
- Martínez, J.G. (2008). Las rúbricas en la evaluación escolar: Su construcción y su uso. *Avances en Medición* 6: 129-134. Recuperado de: goo.gl/jyGfGt
- Martínez Rizo, F. (2016). *La evaluación de docentes de educación básica. Una revisión de la experiencia internacional*. México: INEE.
- Méndez, M., C., y Rondón, S., M., A. (2012). Introducción al análisis factorial exploratorio. *Revista Colombiana de Psiquiatría* 41(1), 197-207.
- Nakatani, Y. (2010). Identifying strategies that facilitate EFL learners' oral communication: A classroom study using multiple data collection procedures. *The Modern Language Journal*, 94(1), 346-8512.
- Tobón, S. (2013b). *La evaluación de las competencias en la educación básica* (2da. Ed.). México: Santillana.
- Tobón, S. (2014) Evaluación con rúbricas socioformativas (Mapas de aprendizaje). *Plataforma CIFE*. Recuperado de www.cife.edu.mx

- Tobón, S. (2017a). *Evaluación socioformativa. Estrategias e instrumentos*. Mount Dora: Kresearch.
- Tobón, S., Guzmán, E., Tobón, B. (2018). Evaluación del Desempeño Docente en México: Del Proyecto de Enseñanza al Proyecto Formativo. *Atenas* 1, 18 - 33. Recuperado de: goo.gl/arjBu
- Zabala, A., & Arnau, L. (2007). *Cómo aprender y enseñar competencias*. Barcelona: Graó. Recuperado de: <http://goo.gl/bXkiYg>