

## Mapeo científico como técnica de investigación: puntos de corte en pruebas de evaluación educativa referidas a criterios como campo de conocimiento

Melissa Villalobos García – Universidad de Valladolid  
 José María Marbán Prieto – Universidad de Valladolid  
 Rocío Anguita – Universidad de Valladolid

 0000-0001-7812-3500  
 0000-0002-6561-6784  
 0000-0002-2533-8871

Recepción: 24.08.2021 | Aceptado: 16.10.2021

Correspondencia a través de **ORCID**: Melissa Villalobos

 **0000-0001-7812-3500**

Citar: Villalobos, M, Marbán, JM y Anguita, R (2021). Mapeo científico como técnica de investigación: puntos de corte en pruebas de evaluación educativa referidas a criterios como campo de conocimiento. *REIDOCREA*, 10(33), 1-18.

Agradecimiento: A la Universidad de Valladolid por facilitar el acceso a las bases de datos requeridas para la realización de este estudio.

Estudio de investigación de Trabajo Fin de Máster: Máster en Investigación aplicada a la Educación, Facultad de Educación, Universidad de Valladolid, España.

Área o categoría del conocimiento: Multidisciplinar.

**Resumen:** A través de una estrategia de revisión creativa de la literatura conocida como mapeo científico o Science Mapping, se lleva a cabo una caracterización de la configuración intelectual del campo de conocimiento del establecimiento de puntos de corte en pruebas de evaluación referidas a criterios. Partiendo del análisis de 601 artículos científicos se identifican las tendencias de evolución del campo de conocimiento y sus principales líneas de publicación a partir de los autores, citas y la interacción entre éstas. Además, a partir de la identificación de nodos, se observa que es un tema desarrollado desde hace varias décadas con mayor crecimiento en volumen de producción científica a partir del año 2000. Este campo de conocimiento muestra una progresión desde aspectos vinculados a metodología hacia la especificación en campos de evaluación de conocimiento como el de educación y el de la salud.

**Palabra clave:** Mapeo Científico

***Science Mapping as a research approach: cut-off points in criteria-referenced educational evaluation tests as field of knowledge***

**Abstract:** Through a creative literature review strategy known as Science Mapping, a characterization of the intellectual configuration of the field of knowledge about cut-off points in evaluation tests referred to criteria was carried out. Based on the analysis of 601 scientific articles, trends in the evolution of the field of knowledge and its main lines of publication was identified based on the authors, citations, and the interaction between them. In addition, from the identification of nodes, it was observed that it is a topic developed for several decades with greater growth in volume of scientific production from the year 2000. This field of knowledge shows a progression from aspects related to methodology towards specification in fields of knowledge evaluation such as education and health.

**Keyword:** Science mapping

### Introducción

#### Mapeo científico o Science Mapping como técnica de investigación

La producción científica en múltiples disciplinas ha experimentado un crecimiento y evolución exponencial en los últimos 15 años (Scharnhorst, Börner, & Van den Besselaar, 2012). Por esa razón, al explorar un campo de conocimiento, por ejemplo, en el ámbito de la investigación educativa se requiere del uso de técnicas de análisis que permitan evidenciar de forma eficaz las dinámicas, las estructuras y la evolución de dicho campo de conocimiento. Esto conlleva la necesidad de emplear una serie de habilidades cognitivas de orden superior para elaborar búsquedas de información que sean oportunas, competentes y, principalmente, que faciliten la exploración del campo de conocimiento (Hart, 1998).

La perspectiva de análisis descrita anteriormente se refiere al concepto de *revisión creativa de la literatura*, desarrollado por Hart (1998), quien construye este concepto alrededor de la necesidad de profundizar para tener una visión amplia de un tema a partir de la posibilidad de explorar diferentes ideas para buscar relaciones no evidentes entre los datos disponibles. Para llevarlo a cabo, el autor define una serie de estrategias oportunas a partir de las cuales se podría abordar un proceso de investigación bibliográfica orientada a la caracterización de la configuración intelectual de un campo de conocimiento.

Dentro de estas estrategias, destacan las de argumentación y análisis, las de organización y expresión de ideas y las de mapeo científico o, como su autor las define en su idioma original, *science mapping strategies* (Hart, 1998). Estas últimas se consideran útiles para explorar un campo de conocimiento desde una perspectiva estructural y dinámica al permitir examinar su evolución en el tiempo, mapear la literatura disponible y trabajar directamente con conjuntos de términos compartidos entre los documentos publicados (Rodríguez, Alcaide y Cobo, 2018). Asimismo, ofrecen la oportunidad de “mapear las relaciones intelectuales que existen en la literatura sobre un tema determinado y así revelar los patrones implícitos, entre los orígenes de una idea, su desarrollo y su implementación” [traducción libre] (Hart, 1998: 161). De esta manera, al resumir los temas de investigación principales de un campo de conocimiento estudiado es posible identificar otros aspectos de interés como brechas de investigación existentes y, a partir de ello, proponer líneas de investigación futuras en un campo de estudio (Ruoyu et al., 2019).

Los avances tecnológicos y las oportunidades que la web 2.0 ofrece en la actualidad han permitido generar diferentes herramientas de apoyo que facilitan este tipo de análisis, favoreciendo la navegación y exploración de los patrones y las tendencias dentro de un conjunto de publicaciones científicas (Chen, 2017). Así, por ejemplo, se han desarrollado diferentes herramientas con funcionalidades y limitaciones específicas, según el objetivo investigativo, como lo son *VOSViewer* (Van Eck & Waltman, 2018), *CiteSpace* (Chen, 2018) y otras como *HistCite*, *SciMAT*, y *Sci2*.

En este contexto, las estrategias de investigación de mapeo científico o *Science Mapping* se constituyen cada vez más como una vía eficaz y oportuna para caracterizar la evolución de un campo de conocimiento (Díaz-Iglesias, Blanco-González y Orden-Cruz, 2019). De acuerdo con Vargas-Quesada, Chinchilla-Rodríguez y Perianes-Rodríguez (2018, p.973):

*Science mapping* ha sido utilizado para mostrar la columna vertebral de la ciencia a nivel mundial, representar y analizar la estructura de grandes bases de datos científicas y, visualizar la estructura intelectual [...], principales líneas de investigación y evolución mundial, entre otras importantes iniciativas.

De esta manera, en el presente artículo se muestra un ejemplo de aplicación de esta estrategia de investigación, con la finalidad de generar reflexión en torno al aprovechamiento de este tipo de análisis dentro del ámbito de la investigación educativa. Para ello, se elige como campo de conocimiento la temática relacionada con el establecimiento de puntos de corte en pruebas de evaluación educativa referidas a criterios.

### **Pruebas de evaluación educativa referidas a criterios**

Desde principios de la década de los 80, Popham (1983) hacía referencia a la diferenciación entre las pruebas de evaluación referidas a criterios y las pruebas

referidas a la norma, al explicar que estas últimas “están destinadas a determinar la posición de un sujeto examinado en relación con el rendimiento de un grupo de otros sujetos que hayan hecho ese mismo test” (Popham, 1983, p.47), mientras que el test basado en criterios “se emplea para determinar la posición de un individuo con respecto a un dominio de la conducta perfectamente definido” (Popham, 1983, p.134). La diferencia descrita anteriormente se considera como base para la identificación de distinciones más específicas, por ejemplo, en cuanto al objetivo que se busca con cada tipo de prueba. De acuerdo con Barrios y Cosculluela (2013), en el caso de las pruebas referidas a la norma el principal objetivo es la medición de una variable psicológica o un rasgo, con lo cual tienden a ser más empleadas en ámbitos de evaluación de personalidad o de actitudes, mientras que las pruebas referidas a criterios tienen como principal objetivo medir un conjunto de conocimientos o competencias, por lo que tienden a ser más empleadas en el ámbito de la educación.

Partiendo de esto puede afirmarse que las pruebas de evaluación referidas a criterios constituyen una herramienta importante dentro del ámbito de la investigación educativa. De hecho, en las últimas décadas, se han desarrollado numerosas estrategias para llevar a cabo el establecimiento de los puntos de corte en este tipo de pruebas, es decir, de estrategias útiles para establecer cómo se lleva a cabo la definición de los niveles mediante los cuales se pretende clasificar o ubicar el logro de los participantes en la prueba (Cizek & Bunch, 2007). Ante ello, uno de los principales interrogantes que han surgido dentro de este campo de conocimiento ha sido cómo identificar cuál estrategia podría ser más pertinente para definir dichos puntos de corte. García, Abad, Olea y Aguado (2012) explican que, de acuerdo con la Asociación Americana de Investigación Educativa (*American Educational Research Association –AERA*, 1999) no es posible definir un único procedimiento como el más adecuado, ya que un mismo método puede ser aplicado de diferentes maneras, según los requerimientos contextuales y de la propia prueba. Lo que sí parece estar claro es que cualquier controversia en relación con la selección de una u otra estrategia podría minimizarse implementándola fielmente y reuniendo evidencia sólida sobre la validez del proceso y de los resultados obtenidos (Cizek, Bunch y Koons, 2004).

En esta misma línea, desde hace varias décadas se han desarrollado numerosas estrategias que pretenden responder a las demandas de definición de puntos de corte en diferentes contextos, con lo cual existen diversas opciones para generarlos, así como clasificaciones de las estrategias disponibles. Dentro de los modelos de clasificación de estas estrategias destaca la de Barrios y Cosculluela (2013, p.127) y respaldada por Muñiz (2018), en la que se definen tres categorías:

1. **Estrategias centradas en el test (*Test Centered*)**. Basadas en la valoración que un grupo de expertos hace de manera detallada sobre los ítems de una prueba de evaluación. Dentro de este grupo destacan estrategias como las de Nedelsky (1954), la de Angoff (1984) y las de Sireci, Hambleton y Pitoniak (2004).
2. **Estrategias centradas en los examinados (*Examinee Centered*)**. Basadas en la valoración que un grupo de expertos hace sobre la competencia o desempeño de los sujetos evaluados. En este caso, destacan estrategias como las de *Grupos de contraste* de Berk (1976) y las de *Grupos límite* de Zieky y Livingston (1977). Cabe mencionar que Cizek y Bunch (2007) también denominan este tipo de estrategias como *métodos holísticos*.
3. **Métodos de compromiso (*Norm referenced*)**. Equilibran la emisión de los juicios con la comparación del sujeto con su grupo normativo, es decir, con datos empíricos. Se incluyen dentro de este grupo técnicas como las de Hofstee (1983) y la Beuk (1984) y otras desarrolladas más recientemente como es el caso del método *Bookmark* (Jornet y Gonzales, 2009).

Partiendo de los aspectos planteados, se evidencia que este campo de conocimiento ha sido ampliamente estudiado desde hace varias décadas (Cizek y Bunch, 2007). Por una parte, se cuenta con estudios que se han enfocado a valorar un solo método de definición de puntos de corte, con el fin de aportar validez empírica y de probar el funcionamiento de algunas variaciones propuestas (Margolis y Clauser, 2014; Tiffin-Richards & Anand-Pant, 2013). También se cuenta con amplia bibliografía alrededor de la comparación de resultados entre las diferentes estrategias disponibles (Aranguren y Hoszowski 2017; Buckendahl, Smith, Impara y Plake, 2002; Cetin & Gelbal, 2013; Hamme & Shulz, 2011; Kaufman, Mann, Muijtjens & Van der Vleuten, 2000; Reckase, 2006; Sanju, Sayeed & Femi, 2006; Van Niljen y Janssen, 2008). Estos planteamientos demuestran que se trata de un campo de conocimiento complejo, con lo cual surge el interrogante siguiente: ¿Cuáles son las principales tendencias que caracterizan la configuración intelectual del campo de conocimiento sobre la definición de puntos de corte en pruebas de evaluación referidas a criterios?

### **Objetivo del estudio**

El interrogante y los aspectos anteriormente planteados justifican la aproximación a partir de la cual se aborda el objeto de estudio en el presente estudio, correspondiente a la estrategia de investigación y revisión creativa de la literatura definida como mapeo científico o Science Mapping. Concretamente, el objetivo de este estudio se orienta a: Identificar las principales tendencias que caracterizan la configuración intelectual del campo de conocimiento sobre la definición de puntos de corte en pruebas de evaluación educativa referidas a criterios mediante una estrategia de Science mapping.

### **Método**

#### **Tipo de estudio**

El presente estudio corresponde a una investigación teórica de tipo bibliométrica con un diseño descriptivo-retrospectivo (Montero y León, 2007). Este tipo de investigación se caracteriza por explorar la configuración intelectual de un campo de conocimiento sin aportar datos empíricos originales, sino que se basa en el análisis de trabajos elaborados previamente para identificar tendencias. Hoyos (2010) explica que este tipo de investigación se corresponde con un procedimiento científico cuyo objetivo es lograr un conocimiento crítico sobre un tema de interés que, en este caso, corresponde al ámbito psicométrico de la definición de puntos de corte en pruebas de evaluación educativa referidas a criterios.

El carácter descriptivo-retrospectivo del estudio está dado por el análisis de relaciones entre variables en una situación pasada, siendo las unidades de análisis objetos y no personas y que, en este caso, corresponden a los documentos o publicaciones analizadas (Montero y León, 2007). A partir de la identificación, selección y organización del material a analizar se amplía y profundiza en el tema abordado, pero tomando en consideración una serie de elementos que permitan comprender de mejor manera la configuración intelectual del campo de conocimiento como base para el análisis crítico (Bernal, Martínez, Parra y Jiménez, 2015).

La revisión de la literatura dentro de este diseño se lleva a cabo con importante rigor metodológico con el fin de que las inferencias y las relaciones conceptuales que se identifiquen puedan, de manera confiable, dar cuenta del saber acumulado en el campo de conocimiento de interés (Bernal, Martínez, Parra y Jiménez, 2015). De esta manera, con el fin de emitir inferencias confiables en el presente estudio se utilizaron los criterios

de rigurosidad científica desarrollados por Barbosa, Barbosa y Rodríguez (2013), los cuales corresponden a:

1. **Criterio de finalidad.** *Representa el compromiso por establecer objetivos de investigación previos.* En el caso del presente estudio, la revisión creativa de la literatura se elabora partiendo de un objetivo concreto orientado a la caracterización de un campo de conocimiento de alto interés para el sector de la investigación educativa, como lo es el establecimiento de puntos de corte en pruebas de evaluación educativa referidas a criterios.
2. **Criterio de coherencia.** *Refiere a la posibilidad de contar con unidad interna en materia de fases, actividades y datos.* El estudio se desarrolla siguiendo un hilo investigativo de revisión creativa de la literatura, el cual determina los procedimientos a desarrollar en coherencia con el objetivo planteado.
3. **Criterio de fidelidad.** *Alude a un respaldo en materia de recolección y transcripción de la información.* Durante el desarrollo del presente estudio este criterio se aplicó a partir del proceso de respaldo las búsquedas de información realizadas, así como los diferentes tipos de análisis desarrollados. El material se encuentra debidamente almacenado y disponible para cualquier consulta o análisis posterior que se requiera.
4. **Criterio de comprensión.** *Se traduce en el favorecimiento de la construcción teórica sobre el objeto de estudio.* El diseño de investigación del presente estudio aporta evidencias para la comprensión de la configuración intelectual del campo de conocimiento definido, siendo la finalidad apoyar la toma de decisiones de los actores involucrados en la definición de puntos de corte en pruebas de evaluación en el ámbito educativo.

## Muestra

Dentro de las metodologías de estudios bibliométricos, los documentos a analizar corresponden a la muestra del estudio y la unidad de análisis corresponde al campo de conocimiento explorado (Angarita, 2014). En el caso del presente estudio, los documentos a analizar se obtuvieron de la base de datos *ISI Web of Science* (en adelante WOS) del paquete *ISI Web of Knowledge* de la empresa *Thomson Reuters*. La selección de esta base de datos responde a que, de acuerdo con Cortés (2008), se trata de una de las bases de mayor presencia científica y académica a nivel internacional.

La WOS únicamente indexa producción científica de revistas arbitradas y, según datos de *Thomson Reuters* (citado en Cortés, 2008), está compuesta por diferentes índices que, en este caso, constituyen el universo muestral del estudio (por ejemplo, el *Science Citation Index Expanded* y el *Social Sciences Citation Index*, entre otros). Dentro de las limitaciones que supone una base de datos de este tipo se encuentra que es de pago, por lo que es importante aclarar que, para el desarrollo de la presente investigación, se accedió a ella gracias a los servicios estudiantiles brindados por el sistema de bibliotecas de la Universidad de Valladolid (de la Comunidad Autónoma de Castilla y León, España). Como parte de la estrategia de selección de la muestra, se definieron los siguientes criterios de inclusión:

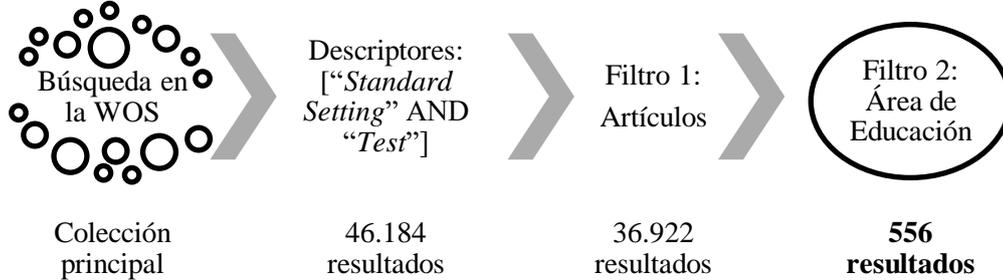
1. Estar redactado en formato de artículo científico.
2. Estar indexado en el *ISI Web of Science*.
3. Responder a los descriptores [*“Standard Setting” AND “Test”*] OR [*“Standard Setting” AND “Cut scores”*].
4. Corresponder al ámbito de la educación.

Partiendo de los criterios de selección se llevó a cabo la primera búsqueda de información en la WOS, utilizando los descriptores mencionados. De esta primera búsqueda se recuperó un total de 556 fuentes de información pertinentes. En la **Tabla 1** se muestra el diagrama del procedimiento llevado a cabo:

**Tabla 1.** Procedimiento para la primera búsqueda de información en la WOS

**Fecha de la búsqueda:** 28 de marzo del 2018.

**Procedimiento:**



**Script ejecutado:**

**Tema:** (standard setting) AND **Tema:** (tests)

**Refinado por:** Tipos de documento: (ARTICLE) AND **Categorías de Web of Science:** (EDUCATION EDUCATIONAL RESEARCH OR EDUCATION SCIENTIFIC DISCIPLINES )

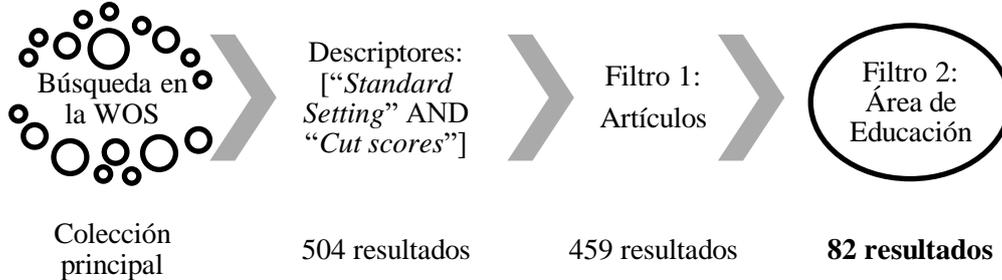
Índices=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Período de tiempo=Todos los años

Posteriormente, se ejecutó una segunda búsqueda de información utilizando el otro conjunto de descriptores seleccionados para aproximar el tema de interés. En este caso, se recuperaron 82 fuentes de información más. En la **Tabla 2** muestra el procedimiento llevado a cabo:

**Tabla 2.** Procedimiento para la segunda búsqueda de información en la WOS

**Fecha de la búsqueda:** 28 de marzo del 2018.

**Procedimiento:**



**Script ejecutado:**

**Tema:** (standard setting) AND **Tema:** (cut scores)

**Refinado por:** Tipos de documento: (ARTICLE) AND **Categorías de Web of Science:** (PSYCHOLOGY EDUCATIONAL OR EDUCATION EDUCATIONAL RESEARCH OR EDUCATION SCIENTIFIC DISCIPLINES OR EDUCATION SPECIAL)

Índices=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Período de tiempo=Todos los años

Como resultado de ambas búsquedas se contó con un total de 638 artículos científicos. Sin embargo, al descartar aquellos que se repitieron como resultado de ambos procesos de búsqueda, la muestra final para el desarrollo del estudio quedó conformada por un total de **601 artículos científicos**, que cumplieron con los criterios de inclusión definidos para la investigación.

## Procedimiento de recolección y análisis de los datos

Para la definición de las fases que guiaron el desarrollo del estudio empleando la estrategia de *Science Mapping*, se tomaron como base los modelos de Aria y Cuccurullo (2017) y Hoyos (2010) y se establecieron tres fases para su desarrollo:

1. **Fase preparatoria.** *Recolección de datos, selección de las estrategias de análisis y depuración de la base de datos.* Se llevó a cabo la definición de los criterios de inclusión para seleccionar las fuentes de información a analizar. Posteriormente, se llevó a cabo la recolección de la información desde la WOS, siguiendo la estrategia de filtrado según los criterios establecidos. Dentro de esta fase se preparó la base de datos, lo que implicó realizar un proceso de depuración y unificación de la información disponible según los requerimientos de formato establecidos por las herramientas de análisis seleccionadas. Además, se preparó el *script* para ejecutar el análisis global de la base de datos, así como la definición de los parámetros y aspectos a considerar (véase *script* de *Bibliometrix* ejecutado en *R* en **Anexo 1**). De esta manera, la fase preparatoria incluyó los procesos de recuperación, filtrado y pre-procesado de la base de datos para el estudio, así como la definición de las estrategias y las herramientas de análisis a emplear.
2. **Fase descriptiva.** *Caracterización de la configuración intelectual del campo de conocimiento y visualización de las principales tendencias.* Dentro de esta fase se inició el análisis de la información utilizando la herramienta de software *R* en su versión 3.4.4 (*R Core Team*, 2018). Específicamente, se empleó una herramienta útil para el análisis de tipo *Science Mapping*, desarrollada por Aria y Cuccurullo (2017) que permite realizar análisis estructurados y caracterizaciones generales de amplios volúmenes de información científica mediante el software *R*. Para el desarrollo de la estrategia de visualización de tendencias se empleó la herramienta *VOSViewer*, con el objetivo de visualizar y extraer las redes de nodos o mapas de tendencias propios del conjunto de datos.
3. **Fase analítica.** *Análisis e interpretación de los resultados obtenidos.* El análisis se llevó a cabo a partir de la definición de variables de interés como la conformación conceptual y la semántica del campo de conocimiento, la identificación de los autores clave, la caracterización de las tendencias de citación entre autores y los principales referentes por espacio geográfico, entre otros. Adicionalmente, dentro de esta fase se llevó a cabo una interpretación y discusión de los resultados obtenidos, con lo cual se resaltaron las principales características del campo de conocimiento en estudio, se identificaron referentes clave a considerar en esta temática. Finalmente, a partir de la información analizada se generaron posibles líneas de investigación futura con recomendaciones pertinentes para que se puedan extender los alcances de la presente investigación.

## Resultados

### Descripción del conjunto de datos

Para la caracterización del conjunto de datos se ejecutaron en dos momentos los análisis bibliométricos en *R*. Primero se analizó el *output* o conjunto de datos de manera exploratoria al momento de obtenerlo desde la WOS (Véase columna *Output Inicial* en la **Tabla 3**). Con posterioridad a la depuración de la base de datos (Véase columna *Output Depurado* en la **Tabla 3**), se ejecutaron de nuevo los análisis. A este punto, es importante mencionar que la comparación entre los datos del análisis inicial y el depurado se muestran únicamente en la **Tabla 3** para efectos de comparar los

resultados del proceso de depuración. Sin embargo, en los análisis y tablas posteriores únicamente se muestran los datos depurados.

**Tabla 3.** Información principal del conjunto de datos (n=601 artículos)

Conjunto de datos	Ouput Inicial	Ouput Depurado
<b>Artículos</b>	601	601
Fuentes (Revistas, libros, etc.)	179	179
Palabras clave Plus (ID)	985	985
Palabras clave del autor (DE)	1441	1441
Periodo	1976 - 2018	1976 - 2018
Promedio de citas por artículo	14,77	14,77
<b>Autores</b>	1717	1685
Apariciones del autor	1952	1952
Autores de artículos de autor único	95	91
Autores de artículos de varios autores	1622	1594
Artículos por autor	0,35	0,36
Autores por artículo	2,86	2,80
Índice de colaboración	3,3	3,24

Fuente: Traducción libre de salida de datos de *Bibliometrix* de R.

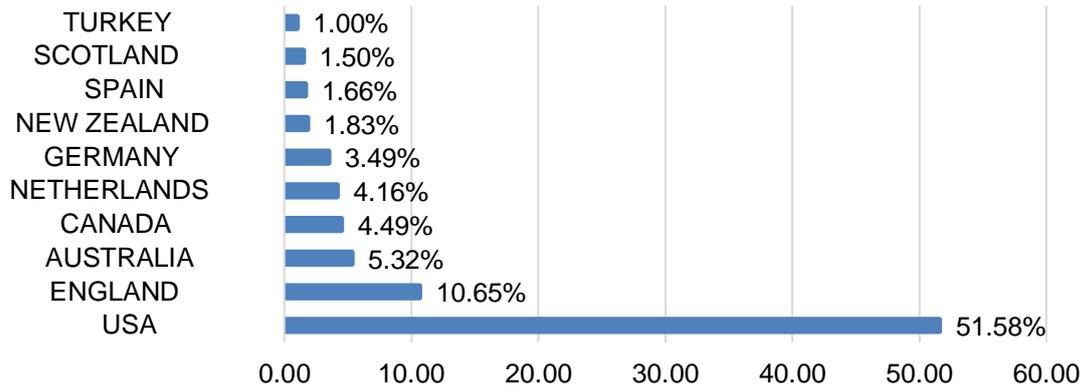
Como se observa, se trata de un conjunto de 601 artículos de 179 revistas científicas, cuyo rango de años de publicación está comprendido desde 1976 hasta el 2018. Además, se identifica la participación de al menos 1685 autores. Los datos sobre la autoría de los documentos sugieren que, en la mayoría de los casos, se han realizado las publicaciones en conjunto con otros autores y no de manera individual. Lo anterior se confirma con el hecho de que se cuenta con un promedio de 0,36 artículos por autor, es decir, que el promedio de autores por artículo es cercano a 3 (concretamente 2,80). Pese a esto, como se muestra en la **Tabla 4**, algunos de los autores tienen altas tasas de productividad, como es el caso de *Van del Vleutenc*, quien se coloca como el autor más productivo dentro del conjunto de datos, con un registro de al menos 20 artículos en los que figura como primer autor.

**Tabla 4.** Autores más productivos dentro del conjunto de datos (n=601 artículos)

	Autores	Artículos
1	VAN,DERVLEUTENC	20
2	MUIJTJENS,A	10
3	SCHERPBIER,AJJA	9
4	NORCINI,JJ	8
5	PLAKE,BS	7
6	WYSE,AE	7

Fuente: Traducción libre de la salida de datos del paquete *Bibliometrix* de R.

El análisis de los autores más productivos dentro del campo de conocimiento también permite ubicar los países de donde provienen dichos autores, con lo cual es posible generar un ranking de los más productivos dentro del conjunto de datos. En este sentido, a partir del análisis de la información y como se observa en la **Figura 1** se destaca Estados Unidos como el país con mayor cantidad de publicaciones dentro del conjunto de datos (310 publicaciones, correspondientes al 51,08% del total que conforma la muestra). Adicionalmente, en la **Tabla 5** es posible identificar que este país se ha citado en al menos 5.231 ocasiones, lo cual equivale a un promedio de citación de 16,87 artículos.



**Figura 1.** Porcentaje de aporte por país al campo de conocimiento estudiado (n=601 artículos). Fuente: Elaboración propia a partir del conjunto de datos.

**Tabla 5.** Total de citaciones por país (n=601 artículos)

País	Total Citaciones	Promedio de citaciones por artículo
1 USA	5231	16,87
2 ENGLAND	897	14,02
3 NETHERLANDS	735	24,40
4 CANADA	428	15,85
5 AUSTRALIA	331	10,34
6 NEW ZEALAND	166	15,09

Fuente: Traducción libre de la salida de datos del paquete *Bibliometrix* de R

Por otra parte, como se observa en la **Tabla 6**, algunos de los artículos que conforman la base de datos cuentan con altos índices de citación, es decir, son importantes referentes dentro del campo de conocimiento del establecimiento de puntos de corte en pruebas educativas de evaluación referida a criterios. De esta manera, destaca el caso del artículo [ANDERSON JR; CORBETT AT; KOEDINGER KR; PELLETIER R, (1995), J. LEARN. SCI], que ha sido citado en al menos 642 ocasiones en los otros artículos del conjunto de datos.

**Tabla 6.** Manuscritos principalmente citados dentro del conjunto de datos (n=601 artículos)

Artículos	Total de Citaciones (TC)	TC por año
1 ANDERSON JR; CORBETT AT; KOEDINGER KR; PELLETIER R, (1995), J. LEARN. SCI.	642	27.91
2 HELLER P; HOLLABAUGH M, (1992), AM. J. PHYS.	238	9.15
3 RETHANS JJ; NORCINI JJ; BARON-MALDONADO M; BLACKMORE D; JOLLY B; LADUCA T; LEW S; PAGE GG; SOUTHGATE L, (2002), MED. EDUC.	189	11.81
4 CAPELLA J; SMITH S; PHILP A; PUTNAM T; GILBERT C; FRY W; HARVEY E; WRIGHT A; HENDERSON K; BAKER D; RANSON S; REMINE S, (2010), J. SURG. EDUC.	147	18.38
5 NORCINI JJ, (2003), MED. EDUC.	138	9.20
6 MOSS PA, (1992), REV. EDUC. RES.	135	5.19

Fuente: Traducción libre de la salida de datos del paquete *Bibliometrix* de R.

Al analizar en detalle el contenido de este artículo posicionado como un referente clave dentro del campo de conocimiento de interés, se observa que fue publicado en el año 1995 por el *Journal of the Learning Sciences* y se trata de una revisión de las publicaciones realizadas en los últimos 10 años sobre el desarrollo de tutores computarizados para apoyar el aprendizaje de las matemáticas de los estudiantes. Considerando que el desarrollo de este tipo de apoyo computarizado requiere de la definición de sistemas de evaluación que permitan identificar un nivel de competencia por parte del aprendiz, se encuentra coherencia en cuanto al peso que tiene este artículo dentro de este campo de conocimiento. Otros artículos que destacan dentro de los más citados son el de [HELLER P; HOLLABAUGH M, (1992), AM. J. PHYS.], con un total 238 citaciones y el de [RETHANS JJ; NORCINI JJ; BARON-MALDONADO M; BLACKMORE D; JOLLY B; LADUCA T; LEW S; PAGE GG; SOUTHGATE L, (2002), MED. EDUC.], con un total de 189 citaciones. En ambos casos, se trata de artículos relacionados con la evaluación de habilidades. El primero, aborda el tema de la evaluación de habilidades de resolución de problemas dentro del ámbito educativo, mientras que el segundo, está orientado a la diferenciación entre los modelos de medición basados en competencias y los basados en desempeños, en evaluaciones desarrolladas dentro de la medicina. A partir del análisis de tendencias en citaciones y publicaciones, es posible identificar otras referencias clave que no son necesariamente artículos científicos, pero que igualmente tienen un importante peso dentro de la muestra del estudio al ser referentes frecuentemente citados dentro del conjunto de datos. En este caso, se destacan las referencias que se muestran en la **Tabla 7**:

**Tabla 7.** Otras referencias frecuentemente citadas dentro del conjunto de datos (n=601 artículos)

Fuente citada	Frecuencia	Referencia completa
[ANGOFF W H, 1971, ED MEASUREMENT, P508]	83	<b>Libro:</b> Angoff, W.H. (1971). <i>Scales, norms and equivalent scores. Educational Measurements.</i> Washington, DC: American Council on Education.
[NORCINI JJ, 2003, MED EDUC, V37, P464, DOI 101046/J1365-2923200301495X]	30	<b>Artículo:</b> Norcini, J. (2003). Setting standards on educational tests. <i>Medical Education</i> , 37(5), pp. 464-469.
[CIZEK G J, 2007, STANDARD SETTING GUI]	29	<b>Libro:</b> Cizek, G. & Bunch, M. (2007). <i>Standard Setting: A Guide to Establishing and Evaluating Performance Standards on Tests.</i> Thousand Oaks, CA: Sage Publications Ltd.

Fuente: Traducción libre de la salida de datos del paquete *Bibliometrix* de R.

## Análisis de redes de nodos y mapas de tendencias del conjunto de datos

Como primera línea de investigación dentro mapeo científico o *Science Mapping* se generó un análisis de la semántica del conjunto de datos que, de acuerdo con Aria y Cucurullo (2017), es útil para visualizar la estructura cognitiva del campo de conocimiento de interés. En primer lugar, se llevó a cabo un análisis de las palabras clave más frecuentemente utilizadas, tanto por parte de los autores de los artículos que conforman el conjunto de datos (Véase columna *Palabras clave del autor* en la **Tabla 8**), como por parte de la propia base de datos al momento de indexar cada artículo (Véase columna *Palabras clave- Plus* en la **Tabla 8**).

**Tabla 8.** Palabras clave más relevantes utilizadas en el conjunto de datos (n=601 artículos)

Palabras clave del autor [Authors Keywords]	Artículos	Palabras clave-Plus [Keywords-Plus]	Artículos
ASSESSMENT	41	EDUCATION	61
STANDARD SETTING	39	PERFORMANCE	53
STANDARDS	27	STUDENTS	37
EDUCATION	20	TESTS	32
CURRICULUM	15	STANDARDS	27
EDUCATIONAL MEASUREMENT	14	SCORES	22
CLINICAL COMPETENCE	12	COMPETENCE	20

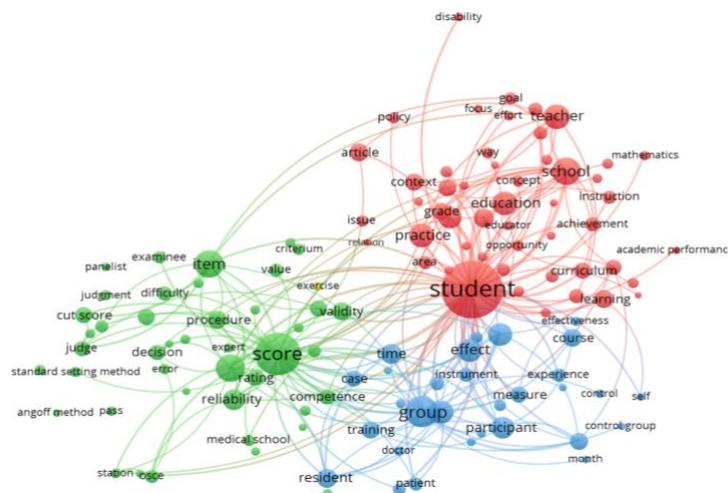
Fuente: Traducción libre de la salida de datos del paquete *Bibliometrix* de R.



Para ampliar la caracterización de la semántica vinculada al campo de conocimiento, se llevó a cabo un segundo análisis de co-ocurrencia de los términos empleados, esta vez, tomando en cuenta las palabras de las otras partes del cuerpo de los artículos. En este caso, se definieron como parámetros los siguientes:

1. **Utilizar los términos disponibles en el título y el resumen de los artículos.** En este caso, entraron en el análisis los 601 títulos de cada uno de los artículos que conforman el conjunto de datos y los resúmenes correspondientes a cada uno de ellos.
2. **Tomar en cuenta todos los términos identificados (*full counting*).** Se identificaron un total de 14.246 términos en el título y resumen de cada uno de los artículos que conforman el conjunto de datos.
3. **Utilizar en la red aquellos términos que aparecieran mínimo 25 veces en el conjunto de datos.** Pese a que el programa sugiere considerar aquellos que aparecen al menos 10 veces, se refinó la búsqueda bajo este parámetro con el fin de identificar los términos que tuvieran un mayor peso dentro del conjunto de datos. Así, se redujo a 178 la cantidad de términos que cumplían con este parámetro.
4. **Visualizar la red utilizando el 70% de los términos que cumplieran con los parámetros establecidos.** En este caso, dicho porcentaje correspondió a un total de 125 términos. Como valor adicional, se definió visualizar un máximo de 500 líneas de asociación entre los términos identificados para favorecer la generación de una gráfica de tendencia más limpia a la vista.

Como se observa en la **Figura 3**, dentro del conjunto de datos se identificó una red semántica compuesta por tres clústeres o subgrupos que representan las áreas temáticas en las que los términos identificados tienden a estar más relacionados. Al visualizar en detalle la red generada, se observa que el primer clúster (representado con el color rojo en la Figura 3) se agrupa principalmente en torno al concepto [*Student*] o *estudiante* y está vinculado a un total de 51 términos relacionados con temas educativos como son: *aprendizaje*, *escuela*, *maestro* y *logro académico*, entre otros. Este subconjunto de la red muestra que del total de 601 artículos que conforman la base de datos destaca un importante grupo que tiene que ver con el ámbito de la educación, lo cual es coherente con los descriptores y filtros empleados para llevar a cabo la búsqueda.



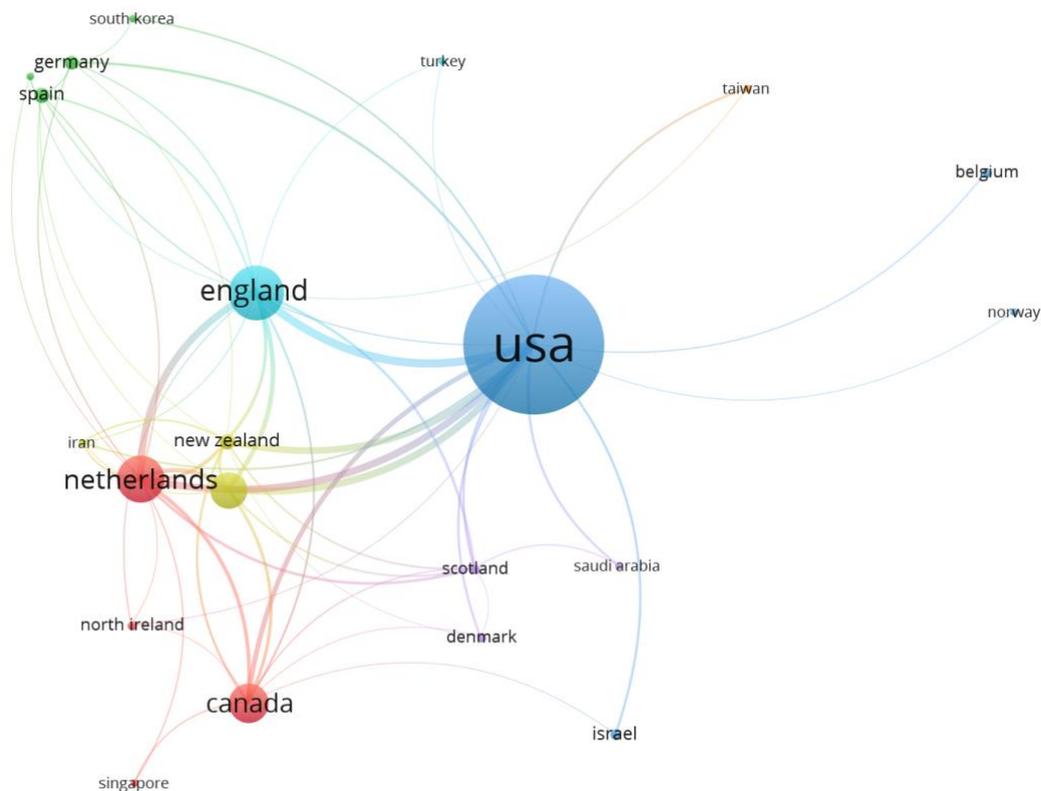
**Figura 3.** Análisis de co-ocurrencia de términos del título y resumen de los artículos (Min. Strength = 100, Max. Lines = 500, n= 601 artículos, 14.246 conceptos,  $\geq 25$  repeticiones, 70% representados en la red).



La segunda línea de investigación explorada fue la tendencia de citación por región geográfica, con el objetivo de identificar aquellos países que parecen estar determinando de manera más importante los avances en el campo de conocimiento estudiado. Para ello, en primer lugar, se realizó un análisis de co-citación (*co-citation*) entre los países con el objetivo de identificar los principales referentes y la manera en la que se vinculan las regiones. Para ello, se definieron como parámetros de análisis:

1. **Tomar en cuenta todos los países citados en cada artículo.** Del total de 601 artículos que componen el conjunto de datos se tuvo registro de 51 países involucrados en las publicaciones.
2. **Visualizar en la red aquellos países que tuvieran un mínimo de tres artículos asociados.** De esta manera, se incluyeron en la visualización de la red un total de 27 países que cumplían dicha condición.

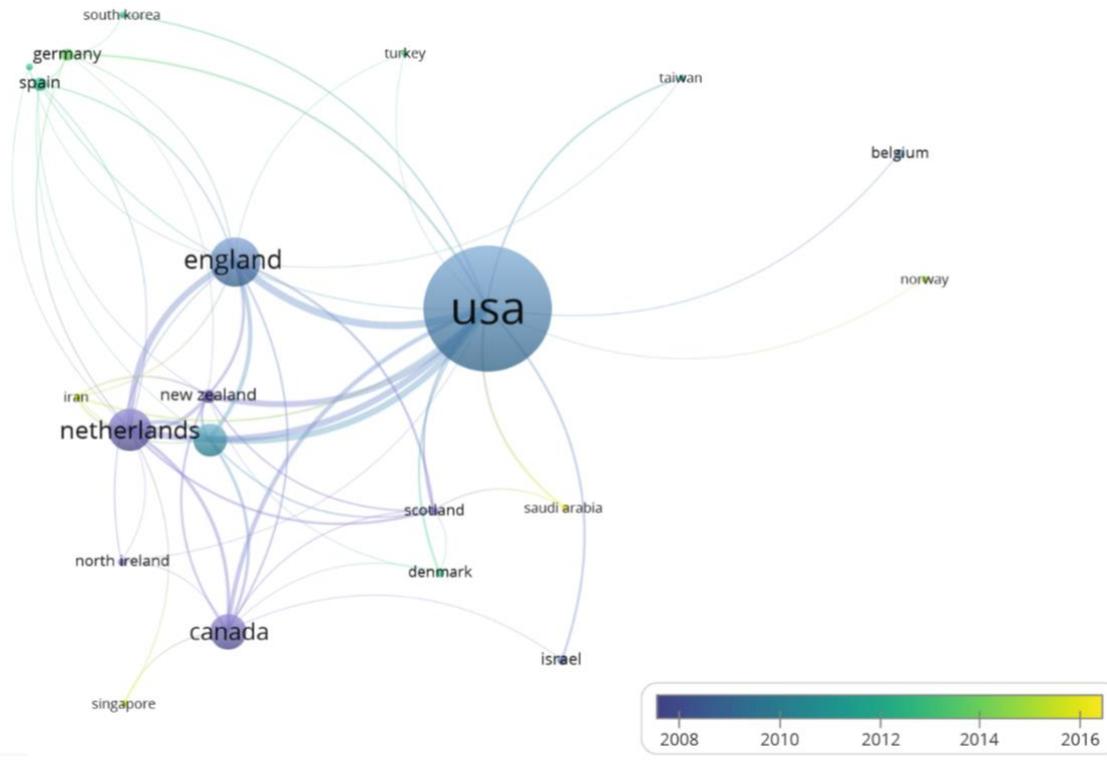
En la **Figura 5**, es posible visualizar la tendencia identificada, en la que se destaca Estados Unidos como el principal referente citado en las diferentes publicaciones, lo cual concuerda con lo observado en la descripción general del conjunto de datos, donde este país se identificó como el más productivo. La tendencia de citación de este país como referente muestra que tiene relación con países como Inglaterra, Holanda, Canadá y Australia. En segundo lugar, se identifica una relación menor con otros países europeos como Alemania, España, Escocia y Bélgica.



**Figura 5.** Análisis de citación entre países dentro del conjunto de datos (n= 601 artículos, Max. countries per document= 51, Min. documents per country= 3).

En este análisis también se exploró la evolución temporal de las interacciones entre países y, como se muestra en la **Figura 6**, los primeros países que empezaron a ser referentes en este campo parecen ser Holanda, Canadá, Nueva Zelanda y Escocia. Adicionalmente, se observa que los dos países que tienden a ser más citados en la

actualidad (Estados Unidos e Inglaterra), empezaron a ser citados de manera más reciente, lo cual sugiere que se identificaron como referentes en los últimos años. Asimismo, es posible visualizar que países como España, Alemania, Singapur e Irán se han incorporado de manera aún más reciente al campo de conocimiento.



**Figura 6.** Evolución temporal de la citación entre países dentro del conjunto de datos (n= 601 artículos, Max. countries per document= 51, Min. documents per country= 3).

La utilidad de visualizar la conformación de esta red y su evolución temporal se relaciona con la posibilidad de visualizar qué países o regiones podrían tener mayor experiencia en el campo de conocimiento estudiado, que en este caso parecen ser Europa y Estados Unidos. Adicionalmente, una de las brechas identificadas a partir de este análisis es que, dentro del conjunto de datos, no se visualizaron referentes de peso que pertenezcan a la región latinoamericana.

## Discusión

El análisis de la configuración intelectual del campo de conocimiento de la definición de puntos de corte en pruebas de evaluación educativa referidas a criterios permitió realizar un mapeo de sus principales características, conformación y evolución temporal de los principales referentes conceptuales en el campo. Dentro de los principales hallazgos destacó que, si bien se trata de un tema desarrollado desde hace varias décadas (finales de los 70), ha tenido su mayor crecimiento en cuanto a volumen de producción científica a partir del año 2000.

Al valorar la evolución temporal del campo de conocimiento se evidenció una progresión de las publicaciones que iba desde aspectos vinculados a la metodología (en los primeros años) hacia la especificación en campos de evaluación de conocimiento concretos como el de educación y el de la salud. Asimismo, los resultados permitieron evidenciar las principales tendencias en cuanto a autores y países más influyentes

dentro del campo de conocimiento, dentro de lo que se destaca Estados Unidos como el país más productivo. Lo anterior es coherente con las tendencias actuales en cuanto los países que son potencias en temas de investigación en educación y llama la atención la ausencia de países latinoamericanos dentro del conjunto de referentes clave, lo cual sugiere que el campo podría ser aún incipiente dentro de dicho contexto, al menos en cuanto a producción o publicación científica.

Otro hallazgo relevante fue que la semántica del conjunto de datos es coherente con la selección de descriptores utilizados en la búsqueda de información en la WOS. Esto se refuerza con el contenido de los tres subgrupos temáticos en torno a los cuales interactuaron las publicaciones, correspondientes a temas educativos y metodológicos vinculados con las evaluaciones de desempeño en poblaciones estudiantiles. Este hallazgo evidencia la importancia que tiene el uso adecuado de los metadatos en las publicaciones científicas, lo cual se considera indispensable para las comunidades científicas al estar, actualmente, inmersas en una sociedad de conocimiento en la que las fuentes de información se multiplican a diario de maneras incalculables.

En este sentido, el desarrollo del procedimiento metodológico para alcanzar los resultados descritos con anterioridad requirió de un detallado y riguroso trabajo, en el que el establecimiento de unos criterios de calidad claros y la definición de parámetros concretos para llevar a cabo el análisis fueron clave para delimitar de manera oportuna las líneas de investigación. El establecimiento de dichos criterios y pautas para realizar el análisis se consideraron pertinentes para alcanzar los resultados del estudio, pero también se consideran útiles para conducir líneas de investigación similares, e incluso con mayor nivel de profundidad, por ejemplo, con la aplicación de parámetros de análisis y de visualización aún más complejos dentro de la estrategia de mapeo científico.

Por otra parte, es importante destacar la contribución que este estudio significa para el ámbito de la investigación educativa. Siendo la utilización de la estrategia de mapeo científico o *Science Mapping* un hilo orientador efectivo y pertinente para la exploración de una temática concreta. En este sentido, se considera que el diseño metodológico planteado en el presente estudio representa un ejemplo del potencial aprovechamiento de las bases de datos científicas y de las diferentes herramientas de software que se han desarrollado en los últimos años para apoyar su uso, de manera que cada vez más se realicen caracterizaciones de la configuración intelectual de un campo de conocimiento con un alcance mayor.

Vinculado con lo anterior, se considera que el presente estudio contribuye con evidencia empírica sobre cómo los avances tecnológicos pueden ser utilizados de manera oportuna en el campo de la investigación aplicada a la educación, lo cual constituye uno de los principales y más importantes aportes del estudio al campo de conocimiento. Partiendo de ello, se recomienda que futuras investigaciones puedan ampliar las posibilidades de profundización en la caracterización de la configuración intelectual de un campo de conocimiento a partir, por ejemplo, de la exploración de parámetros más complejos que actualmente están disponibles dentro del repertorio de los análisis bibliométricos.

En la misma línea se considera interesante realizar una comparación de los alcances y limitaciones de las principales herramientas de software que actualmente se encuentran disponibles para llevar a cabo análisis de este tipo. De esta manera, se podría generar mayor orientación en cuanto a criterios o pautas útiles a tomar en cuenta al seleccionar una herramienta tecnológica para conducir procesos de investigación en educación similares al desarrollado en el presente estudio.

Finalmente, a partir de la publicación de este artículo, se invita a la exploración de otras formas y estrategias metodológicas oportunas para llevar a cabo un aprovechamiento del potencial de las herramientas tecnológicas dentro del contexto específico del ámbito de la investigación educativa dado que, si bien muchas de las herramientas disponibles en la actualidad están al servicio de este ámbito de la investigación, las experiencias en torno a su exploración no suelen a ser ampliamente documentadas.

## Referencias

- American Educational Research Association [AERA]. (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Anderson, J., Corbett, A. Koedinger, K. & Pelletier, R. (1995). Cognitive Tutors: Lessons Learned. *Journal of the Learning Sciences*, 4(2), 167-207.
- Angarita, L. (2014). Estudio bibliométrico sobre uso de métodos y técnicas cualitativas en investigación publicada en bases de datos de uso común entre el 2011-2013. *Revista Iberoamericana de Psicología: Ciencia y Tecnología*, 7(2), 67-76.
- Angoff, W. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), *Educational measurement* (pp. 508-600). Washington, DC: American Council on Education.
- Angoff, W. (1984). *Scales, norms, and equivalent scores*. Princeton, New Jersey: Educational Testing Service.
- Aranguren, M. y Hozowski, A. (2017). Aprender 2016. Serie de documentos técnicos 5: Bookmark, Establecimiento de puntos de corte. Ministerio de Educación y Deportes: Buenos Aires.
- Aria, M. & Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959-975.
- Barbosa, J., Barbosa J. y Rodríguez, M. (2013). Revisión y análisis documental para estado del arte: Una propuesta metodológica desde el contexto de la sistematización de experiencias educativas. *Investigación Bibliotecológica*, 27(61), 83-105.
- Barrios, M. y Cosculluela, A. (2013). Fiabilidad. En J. Meneses, M. Barrios, A. Bonillo, A. Cosculluela, L. Lozano, J. Turbany y S. Valero (Eds.). *Psicometría* (pp. 75-141). Barcelona: OUC.
- Berk, R. (1986). A consumer's guide to setting performance standards on criterion referenced tests. *Review of Educational Research*, 56, 137-172.
- Bernal, D., Martínez, L., Parra, A. y Jiménez, J. (2015). Investigación documental sobre calidad de la educación en instituciones educativas del contexto Iberoamericano. *Revista Entramados Educación y Sociedad*, 2(2), 107- 124.
- Beuk, C. (1984). A method for reaching a compromise between absolute and relative standards in examinations. *Journal of Educational Measurement*, 21, 147-152.
- Buckendahl, Ch., Smith, R., Impara, J. & Plake, B. (2002). A comparison of Angoff and Bookmark Standard Setting Methods. *Journal of Educational Measurement*, 34(3), 253-263.
- Capella, J., Smith, S. Philp, A., Putnam, T., Gilbert, C., Fry, W., Harvey, E., Wright, A., Henderson, K., Baker, D., Ranson, S. & Remine, S. (2010). Teamwork training improves the clinical care of trauma patients. *Journal of Surgical Education*, 67(6), 439-443.
- Cetin, S. & Gelbal, S. (2013). A comparison of Bookmark and Angoff Standard Setting Methods. *Educational Sciences: Theory & Practice*, 13(4), 2169-2175.
- Chen, Ch. (2017). Science Mapping: A Systematic Review of the Literature. *Journal of Data and Information Science*, 2(2), 1-40.
- Chen, Ch. (2018). How to Use CiteSpace. Lean Publishing.
- Cizek, G. & Bunch, M. (2007). *Standard setting: A guide to establish and evaluating performance standards on tests*. California: Sage Publications.
- Cizek, G., Bunch, M. & Koons, H. (2004). Setting performance standards: Contemporary methods. *Educational Measurement: Issues and Practice*, 23(4), 31-49.
- Cizek, G. & Bunch, M. (2007). *Standard setting: A guide to establish and evaluating performance standards on tests*. California: Sage Publications.
- Cortés, J. (2008). Web of Science: Termómetro de la producción internacional de conocimiento: Ventajas y limitaciones. *Culcity*, 5(29), 5-15.
- Díaz-Iglesias, S., Blanco-González, A. y Orden-Cruz, C. (2019). Science mapping analysis of change management. *Espacios*, 40(32), 17-29.
- García, P., Abad, F., Olea, J. y Aguado, D. (2012). Un nuevo método de standard setting basado en la TRI: aplicación a eCat-Listening. Informe de investigación eCAT 12-01. Madrid: Universidad Autónoma de Madrid.
- Hamme, C. & Shulz, M. (2011). Reliability and validity of Bookmark-based methods for standard setting: Comparisons to Angoff-based methods in the national assessment of Educational Progress. *Educational Measurement: Issues and Practice*, 30(2), 3-14.
- Hart, C. (1998). *Doing a literature review: Releasing the social science research imagination*. London: Sage Publications.
- Heller, P. & Hollabaugh, M. (1992). Teaching problem-solving through cooperative grouping. Designing problems and structuring groups. *American Journal of Physics*, 60(7), 637-644.
- Hofstee, B. (1983). The case for compromise in educational selection and grading. In S. B. Anderson & J. S. Helmick (Eds.), *on Educational testing* (pp. 109-127). San Francisco: Jossey-Bass.
- Hoyos, C. (2010). Un modelo para la investigación documental. Guía teórico-práctica sobre construcción de Estados del Arte con

importantes reflexiones sobre la investigación. Medellín: Señal Editora.

Jornet, J. y Gonzales, J. (2009). Evaluación criterial: determinación de estándares de interpretación (EE) para pruebas de rendimiento educativo. *Estudios sobre Educación*, 16, 103-123.

Kaufman, D., Mann, K., Muijtjens, A. & Van der Vleuten, C. (2000). A comparison of standard-setting procedures for an OSCE in undergraduate medical education. *Academic Medicine*, 75(3), 267-271.

Margolis, M. & Clauser, E. (2014). The impact of examinee performance information on judges' cut scores in modified Angoff standard-setting exercises. *Educational Measurement: Issues and Practice*, 33(1), 15–22.

Montero, I. y León, O. (2007). A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology*, 7(3), 847-862.

Moss, P. (1992). Shifting Conceptions of Validity in Educational Measurement: Implications for Performance Assessment. *Review of Educational Research*, 62(3), 229-258.

Muñiz, J. (2018). *Introducción a la Psicometría: Teoría Clásica y TRI*. Madrid: Pirámide.

Nedelsky, L. (1954). Absolute grading standards for objective tests. *Educational and Psychological Measurement*, 14, 3–19.

Norcini, J. (2003). Setting standards on educational tests. *Medical Education*, 37(5), 464-469.

Popham, J. (1983). *Evaluación basada en criterios*. Madrid: Magisterio Español.

R Core Team (2018). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Reckase, M. (2006). A conceptual framework for a psychometric theory for standard setting with examples of its use for evaluating the functioning of two standard setting methods. *Educational Measurement: Issues and Practice*, 5(2), 4-18.

Rethans, J., Norcini, J., Baron-Maldonado, M., Blackmore, D., Jolly, B., LaDuca, T., Lew, S., Page, G., Southgate, L. (2002). The relationship between competence and performance: implications for assessing practice performance. *Medical Education*, 36(10), 901-909.

Rodríguez, M., Alcaide, L. & Cobo, J. (2018). Analyzing the scientific evolution and impact of e-Participation research in JCR journals using science mapping. *International Journal of Information Management*, 40, 111–119.

Ruoyu, Z., Patrick, P., Poorang, W., Hannah, Y., Yan, L. & Han, Y. (2019). A science mapping approach based review of construction safety research. *Safety Science*, 113, 285-297.

Sanju, G., Sayeed, H. & Femi, O. (2006). Standard setting: Comparison of two methods. *BMC Medical Education*, 6(46), 1-6.

Schamhorst, A., Börner, K. & Van den Besselaar, P. (2012). *Models of science dynamics: Encounters between complexity theory and information sciences*. Dordrecht: Springer.

Sireci, S., Hambleton, R. y Pitoniak, M. (2004). Setting passing scores on licensure examinations using direct consensus. *CLEAR Exam Review*, 15(1), 21-25.

Tiffin-Richards, S. & Anand-Pant, H. (2013). Setting standards for English foreign language assessment: Methodology, validation, and a degree of arbitrariness. *Educational Measurement: Issues and Practice*, 32(2), 15–25.

Van Eck, N. & Waltman, L. (2018). *Manual for VOSviewer version 1.6.7*. Netherlands: Universiteit Leiden.

Van Nijlen, D. & Janssen, R. (2008). Modeling judgments in the Angoff and contrasting-groups method of standard setting. *Journal of Educational Measurement*, 45(1), 45-63.

Vargas-Quesada, B., Chinchilla-Rodríguez, Z. & Perianes-Rodríguez, A. (2018). Science mapping artificial intelligence. En Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA), I Workshop en Ciencia de Datos en Redes Sociales, Granada, 23-26 de octubre de 2018.

Zieky, M. & Livingston, S. (1977). *Manual for setting standards on the basic skills assessment tests*. Princeton, NJ: Educational Testing Service.