

LA APORTACIÓN DE LA ESTADÍSTICA EXPLORATORIA AL ANÁLISIS DE DATOS CUALITATIVOS

THE CONTRIBUTION OF EXPLORATORY STATISTICS TO THE ANALYSIS OF QUALITATIVE DATA

Günter L. Huber (*)
Leo Gürtler
Universidad de Tübingen
Alemania

Samuel Gento
Universidad Nacional de Educación a Distancia
España

Resumen

Aplicar métodos estadísticos a la investigación cualitativa en las ciencias sociales había sido considerado como mezcla ilícita. Sin embargo, desde hace dos décadas se recomienda combinar métodos cualitativos y cuantitativos según el enfoque de métodos mixtos. Una combinación muy directa es la conversión de datos cualitativos en datos cuantitativos o vice versa, lo que se aplica regularmente tanto en investigaciones cualitativas como cuantitativas sin causar enfrentamientos paradigmáticos. El presente artículo trata de diferenciar entre tres modos de conversión de datos y describir cómo la estadística exploratoria puede servir en estos casos y cómo podría facilitar que los investigadores encuentren patrones de significado en sus datos. Además, ilustra la utilidad y los límites del análisis exploratorio de datos con dos trabajos de los autores: el análisis de los debates de TV entre Clinton y Trump y el análisis de las dimensiones del liderazgo educativo.

Palabras clave: Análisis exploratorio de datos, diseños de conversión de datos, métodos estadísticos, enfoque cualitativo, enfoque cuantitativo, métodos mixtos, liderazgo educativo.

Abstract

Applying statistical methods in qualitative social research was, in the past, sometimes considered an illicit mixing. Around two decades ago, however, combining qualitative and quantitative methods is recommended, according to the “mixed methods” approach. A very direct combination is to convert qualitative data into quantitative data and vice versa, which is done regularly, both in qualitative and quantitative investigations without causing paradigmatic confrontation. This article differentiates between three modes of data conversion and describes how exploratory statistics may serve in these cases and assist the investigators in finding meaning patterns in their data. Additionally, it illustrates the usefulness and the limits of this approach with two examples taken from the authors’ studies: an analysis of the TV-debates between Clinton and Trump and an analysis of the dimensions of educational leadership.

(*)Autor para correspondencia:

Günter L. Huber.
Universidad de Tübingen.
Facultad de Economía y Ciencias
Sociales.
Departamento de Ciencias de
Educación.
Institut für Erziehungswissenschaft,
Münzgasse 11, 72070 Tübingen,
Alemania.
Correo de contacto:
huber.paedpsy@uni-tuebingen.de

©2010, Perspectiva Educacional
<http://www.perspectivaeducacional.cl>

Keywords: Exploratory data analysis, data conversion designs, statistical methods, qualitative approach, quantitative approach, mixed methods, educational leadership.

RECIBIDO: 5 de diciembre de 2016
ACEPTADO: 01 de octubre de 2017
DOI: 10.4151/07189729-Vol.57-Iss.1-Art.611

1. Introducción

Ya en los comienzos del tratamiento de textos en las ciencias sociales, se diferenciaron los análisis cualitativos y cuantitativos y se consideraron ambos como enfoques incompatibles. En 1952, dos publicaciones determinaron la tendencia a seguir en el futuro: Berelson (1952) y Kracauer (1952). El primer trabajo describió el análisis de contenido como un método para procesar sistemáticamente, objetivamente y cuantitativamente el contenido manifiesto en los textos. Según Kracauer (1952), un análisis cualitativo de contenido permite revelar las categorías latentes, ocultas del significado en un texto e independiente de su contenido manifiesto.

Las disputas basadas en esta distinción nunca fueron meras discusiones metodológicas, sino una combinatoria que incluía la pregunta del *método correcto* y los objetivos de las investigaciones. Una serie de publicaciones en la revista "Educational Researcher" de la American Educational Research Association (AERA) durante los años 80 del siglo pasado representa un buen ejemplo de este debate: la comunidad científica se dedicó a argumentar sobre la polémica cualidad versus cantidad. Representantes del campo cualitativo denominaban a sus oponentes como *number crunchers* (trituradores de números), mientras que representantes del campo cuantitativo hablaban de *story tellers* (narradores de historias; p.ej. Smith, 1983). Los artículos de Smith (1983) y de Smith y Heshusius (1986) tratan de hacer hincapié en una contradicción insuperable entre cualidad y cantidad en las investigaciones de las ciencias sociales. Tales autores advierten las consecuencias serias epistemológicas cuando los investigadores aceptan la propuesta de poder aplicar métodos cualitativos y cuantitativos como herramientas complementarias. Smith y Heshusius (1986) manifiestan que desde el punto de vista epistemológico es imposible que investigadores del enfoque cuantitativo o cualitativo puedan discutir sobre sus problemas: "... if a quantitative inquirer disagrees with a qualitative inquirer, is it even possible for them to talk to each other? The answer, for the present anyway, is a qualified no" (Smith & Heshusius, 1986, p. 11).

Esta posición supone que ambos enfoques no tienen nada en común: mientras que investigaciones cuantitativas tratan de *comprobar* hipótesis por medio de cuantificar características seleccionadas de su muestra, investigaciones cualitativas tratan de *construir* hipótesis por medio de comparar características de los casos analizados. Lo anterior expresa un punto de vista de exclusión mutua y manifiesta la necesidad de decidir entre recordar las acciones, opiniones o actitudes de la gente de la manera más objetiva posible o tratar de explicar por qué las personas actúan, piensan o hablan de la manera observada. Al describir este enfrentamiento algunos autores hablan de *enfrentamientos paradigmáticos* (Gage, 1989; Bryman,

2008).

Además, la posición sugiere al mismo tiempo que investigadores cuantitativos no dependan de la interpretar sus datos e investigadores cualitativos eviten cuantificaciones de cualquier tipo como fuente para concluir el análisis. Sin embargo, cada cantidad puede ser considerada un aporte en caso que los investigadores puedan interpretarla en relación con suposiciones teóricas; interpretación que forma parte de un proceso de índole cualitativa. Por otro lado, cada informe sobre comparaciones cualitativas depende de aplicar al menos cuantificaciones mínimas, por ejemplo, las expresiones como “Los entrevistados por lo general / raras veces / nunca mencionaban...”. No es preciso discutir los esfuerzos interpretativos y cualitativos necesarios en el proceso de diseñar una investigación social, cualquiera sea su tipo. Sin embargo, en ambos enfoques raramente se menciona el hecho de que utilizan conceptos o técnicas del otro campo.

En la actualidad es posible constatar que debates serios y extensos como los planteados por Smith (1983), condujeron al desarrollo de un enfoque complementario de utilización de *métodos mixtos*, además de nuevas discusiones críticas (Wiggins, 2011). El planteamiento de métodos mixtos fue documentado en el manual metodológico de Tashakkori y Teddlie (2003). Un grupo de trabajo de la American Psychological Association (2006) confirmó en general la necesidad y el valor de basar la práctica en hallazgos científicos logrados con diversos métodos, incluyendo entre ellos a los de corte cuantitativo y cualitativo. Teddlie y Tashakkori (2006) concibieron una taxonomía de combinaciones metodológicas, donde se encuentra un diseño de investigación que utiliza el mismo conjunto de datos para análisis cualitativos y cuantitativos. Esto permitiría afirmar que para realizar este tipo de diseño se tienen que convertir los datos cualitativos en datos cuantitativos y viceversa. Los autores incluso lo describen como un diseño de conversión y como una característica única de la investigación con métodos mixtos que utilizan los términos *conversión de datos*, *cuantificar* y *cualificar* (Teddlie & Tashakkori, 2006, p. 17) para mostrar cómo podría aportar la estadística exploratoria al análisis de datos cualitativos. Todo lo anterior se vuelve un antecedente relevante para el presente artículo, debido a que se ha seleccionado a los diseños de conversión con la finalidad de diferenciar entre tres modos de aplicar la conversión de datos a investigaciones empíricas.

2. La aportación potencial del análisis exploratorio de datos

La posición que se ha manifestado sobre la incompatibilidad de métodos cualitativos y cuantitativos en los años 80 no ha tenido en cuenta el desarrollo que se ha producido desde hace algunos años acerca de la utilización de métodos cuantitativos. Tukey (1977) es la obra fundamental en relación con el Análisis Exploratorio de Datos (AED), desarrollado para maximizar la comprensión de un conjunto de datos, revelar estructuras ocultas dentro de los datos, extraer variables importantes, descubrir casos atípicos o anomalías, comprobar suposiciones, desarrollar modelos ahorradores y determinar composiciones óptimas de factores (National Institute of Standards and Technology, NIST). El NIST explica, además, que “the real insight and “feel” for a data set comes as the analyst judiciously probes and explores the various subtleties of the data” (párrafo 1.1.4). Es más importante para este objetivo “to draw on our own human pattern-recognition and comparative abilities in the context of a series of judicious graphical techniques applied to the data” (párrafo 1.1.4).

De todas maneras, el AED incluye un conjunto de estrategias estadísticas que permiten encontrar regularidades o muestras específicas en los datos antes de aplicar métodos más sofisticados. En el contexto de investigaciones cualitativas debería ser más interesante la aportación del AED al objetivo central del enfoque cualitativo, a saber, generar hipótesis. Alea, Jiménez, Muñoz, Torrelles y Viladomiu (2014) subrayan en su guía que: “Mediante AED será posible identificar las medidas de síntesis adecuadas para resumir la información contenida en la muestra y formular hipótesis sobre la población” (p. 63). De la misma manera describe Inzunsa (2014) la utilidad del AED en su estudio empírico asociado al razonamiento estadístico de estudiantes universitarios:

El análisis exploratorio de datos puede ser una herramienta de utilidad en la generación de hipótesis, conjeturas y preguntas de investigación acerca de los fenómenos de donde los datos fueron obtenidos, y por ello se propone como medio para desarrollar una comprensión global de los datos. (Inzunsa, 2014, p. 1264).

La cita anterior permitiría aseverar que esta herramienta puede promover y lograr perfectamente el objetivo esencial de investigaciones cualitativas. En las palabras de Flick (2009, p.12): “The goal of your research then is less to test what is already known (e.g., theories already formulated in advance), but to discover and develop the new and to develop empirically grounded theories”.

Los métodos de AED ofrecen, generalmente, una variedad de representaciones gráficas de

los resultados de operaciones cuantitativas, por lo que evidenciarían más fácilmente patrones o estructuras más o menos ocultas en los datos. El proverbio afirma que una imagen vale más que mil palabras y, en verdad, los gráficos producidos por el AED ofrecen, en el sentido más literal, una visión de la estructura de los datos. Los diagramas producidos por métodos estadísticos permiten que el investigador descubra fácilmente relaciones, similitudes y diferencias entre sus casos y que los lectores de su informe comprendan mejor las conclusiones.

En el campo cuantitativo puede ser ventajoso complementar los procedimientos clásicos con conclusiones interpretativas basadas en representaciones gráficas del análisis exploratorio de datos. Por ejemplo, con un gráfico sencillo puede contestarse la pregunta de si una interacción entre los predictores de un modelo lineal impide la interpretación de predictores individuales o no. Sería complicado establecer la misma conclusión si solamente se consideraran los coeficientes del modelo lineal. Claro está que en estudios cualitativos se necesitan en primer lugar datos cuantitativos para aplicar métodos del análisis exploratorio y producir representaciones gráficas. El presente estudio se enfoca con más detalle en la recomendación de convertir datos, más aún en la posibilidad de cuantificar datos cualitativos.

3. Modos de conversión de datos

A continuación, se diferenciarán los tres modos de conversión de datos. El primer modo es casi ubicuo en las investigaciones sociales de manera que usualmente no se da cuenta del hecho de conversiones cualitativas-cuantitativas o viceversa. El segundo modo se basa en el análisis del mismo conjunto de datos tanto con métodos cualitativos como cuantitativos. El tercer modo por fin combina un análisis cualitativo completo y un análisis exploratorio de datos como acto seguido.

3.1. Conversión 1.0.

Este modo de conversión de datos es muy frecuente en las investigaciones de las ciencias sociales, aunque al emplear técnicas cualitativas y/o cuantitativas suscita oposición (Teddlie & Tashakkori, 2006, p. 19). Además, convertir datos según este modo es algo muy trivial y por eso los investigadores ni siquiera discuten la mezcla metodológica. Tal como uno de los evaluadores de este artículo confirmó, se realiza la conversión 1.0 en muchos estudios para agrupar casos, sujetos, contextos, entre otros datos. A ello se le añade el interés por comparar segmentos del conjunto de datos. A saber, en la mayoría de las investigaciones sociales los datos socio-demográficos incluyen datos métricos, como por ejemplo la edad o los años

de experiencia profesional de los encuestados mediante un cuestionario. Generalmente, se agrupan los datos y convierten en datos nominales. Así, se podrían identificar, por ejemplo, las diferencias entre tres grupos de encuestados con experiencia alta/media/baja. Por otro lado, existen también datos nominales sobre informaciones socio-demográficas (género, profesión, entre otros). Es irrelevante que se etiqueten los grupos resultantes (edad, experiencia, tipo de profesión, etc.) con números (grupo 1, grupo 2, etc.) o con palabras (experiencia alta, experiencia media, etc.), lo relevante es el nivel de los datos y cuáles operaciones matemáticas pueden llevarse a cabo al nivel dado, lo que implicaría establecer o no relaciones entre variables, casos, etc.

La segmentación de la base de datos disponible por medio de conversión 1.0 de datos no sirve solamente para contestar preguntas particulares de investigación, tales como ¿hay diferencias entre los votos de electores jóvenes, de mediana edad o mayores? Más bien, lo que se identifica con un aporte es el hecho de permitir agrupar los datos disponibles por medio de varias variables, es decir, en caso del análisis electoral establecer características individuales como en el caso de investigaciones electorales la información asociada a la educación, al número de habitantes, la región, entre otros elementos. Esto puede generar nuevas preguntas de investigación, además de identificar una nueva recopilación de datos útiles para una futura investigación.

3.2. Conversión 2.0.

El diseño de conversión 2.0 también se basa en el análisis del mismo conjunto de datos tanto con métodos cualitativos como cuantitativos. Ejemplos típicos y bien conocidos son las técnicas del análisis cuantitativo de textos (Andréu, 2011) o del análisis lexicométrico (Beaudouin, 2016; Benzécri, 1992; Tristl, Müller & Bachmann, 2015). Para interpretar textos particulares se identifican primero las palabras clave vinculadas con la pregunta de investigación y justificadas por orientaciones teóricas. Después, en el segundo paso del estudio, se cuantifican estas palabras clave, es decir, se contabiliza su frecuencia, acción que permite complementar el análisis con métodos estadísticos. En este contexto, un paso cualitativo previo del análisis cuantitativo de textos prepara y convierte los datos para el segundo paso, el paso cuantitativo del estudio. La aportación del análisis exploratorio de datos en este caso lo vamos a ilustrar posteriormente con un ejemplo sencillo de un análisis temático preparado para este artículo.

Mientras que la conversión de datos en el caso del diseño de conversión 1.0 es sencillamente un modo del manejo de datos para agrupar casos, el enfoque del diseño de conversión 2.0 combina elementos esenciales del paradigma de codificación del ámbito cualitativo con

el paradigma de frecuencias de la estadística. Para tal efecto, se formula una hipótesis respecto a aspectos particulares de un texto y se determinan las palabras clave como unidades de significado, las que van a permitir indicar la presencia o ausencia de los aspectos hipotéticos.

Así, si un investigador tratara de obtener información sobre la motivación de los solicitantes para una determinada posición, a partir de sus solicitudes podría compilar una lista de palabras o formulaciones críticas que expresen hechos vinculados con la motivación. En definitiva, podría recoger hechos o aspectos deseados y no deseados que aparecieran en tales solicitudes, hecho que se vincula con procedimientos cualitativos. Tras el recuento de las frecuencias de tales palabras o formulaciones se pueden categorizar y comparar las solicitudes (basándose en las frecuencias), analizar las características de la distribución de frecuencias y determinar probabilidades. Por supuesto, la utilización de la estadística exploratoria puede ayudar también a descubrir patrones de motivación en los documentos de texto.

Casi la totalidad de los software para el análisis de datos cualitativos ofrece en la actualidad un sin número de herramientas para el análisis de textos según el modelo de conversión 2.0. Al mismo tiempo, hay programas especializados en el análisis lingüístico, por ejemplo, el paquete de software AntConc (Anthony, 2011).

Un problema posible del diseño de conversión 2.0 lo constituye el hecho de que códigos o palabras clave separados no ofrecen mucha información sobre las relaciones más profundas entre casos, códigos y sus interacciones. El diseño de conversión 3.0 trata de eliminar esta desventaja.

3.3. Conversión 3.0.

La utilización del diseño de conversión 3.0 permite combinar un análisis cualitativo completo y un análisis exploratorio de datos como acto seguido. Este enfoque parece constituir la estrategia adecuada para analizar el contenido de textos con el fin de determinar cualitativamente las palabras clave, contarlas, analizarlas y comparar sus frecuencias. Pero, en este caso, el primer paso representa un estudio cualitativo completo con resultados basados en un sistema de interpretación/codificación bien justificado. Es decir, los resultados ofrecen no solamente frecuencias y perfiles de códigos, sino también vinculaciones típicas entre los códigos y las estructuras de significado reconstruidas. Si después se identifican códigos críticos (relacionados con la pregunta de investigación) y se les convierten en datos cuantitativos (es decir, en frecuencias), puede realizarse uno o varios análisis estadísticos que

ayuden a poner de manifiesto más estructuras o patrones de significado en el conjunto de datos. Obviamente, cada análisis estadístico se basa en códigos y casos de amplio significado. En consecuencia, en este paso no se cuentan ya sencillamente frecuencias de códigos, sino de patrones estructurales.

4. Ejemplos de estudios con conversión de datos

A continuación, esbozamos dos estudios con diseños de conversión como ejemplos de la aplicación del AED al análisis de datos cualitativos. El primero de los ejemplos aplica el modelo 2.0, mientras que el segundo se encuentra bajo las consideraciones del modelo 3.0. El ejercicio se asocia al objetivo de ilustrar cómo las posibilidades gráficas del AED pueden contribuir en el proceso de combinación de técnicas cualitativas y cuantitativas.

4.1. Ejemplo de un diseño de conversión 2.0: Análisis cuantitativo de texto.

En un taller sobre la aplicación de software para el análisis cualitativo, Huber (2004) utilizó la transcripción del primer debate de TV en los EE.UU. en la campaña electoral entre los candidatos George H. W. Bush, Bill Clinton y Ross Perot. Esperábamos que el candidato ya en el cargo y sus oponentes se refiriesen de distinto modo a sus logros, intenciones, programas, entre otras temáticas. Además, se esperaba que trataran de animar a sus electores de manera diferente. Por eso, Huber y Gürtler (2015) decidieron, desde lo cualitativo-interpretativo, escoger palabras clave como los pronombres *I* (yo) y *we* (nosotros) combinados con verbos como *I'll* (quiero), *we'll* (queremos), *I've* (he), *we've* (hemos), entre otros. Sin embargo, mientras que el candidato en el cargo (Bush) utilizaba referencias a sí mismo (por ende, empleaba el pronombre *I*) 1,22 veces más frecuentemente que formulaciones con *we*; el oponente Clinton hablaba 0,79 veces menos de sí mismo que de *we* (refiriéndose al conjunto de electores potenciales, incluido él mismo). En el caso de Perot, no se encontró diferencia alguna digna de mención.

En 2016, Gürtler y Huber aplicaron la misma lista de palabras clave en un análisis de las transcripciones de los tres debates de TV entre Hillary Clinton y Donald Trump. Debido a la naturaleza de las disputas, los autores ampliaron la lista de palabras clave por referencias a la otra persona: *she* (ella), *he* (él), *Hillary*, *Donald* y las combinaciones correspondientes con los verbos mencionados anteriormente. Como instrumento de investigación se utilizó el software AQUAD Siete (Huber, 2015), que contiene una parte estadística de Gürtler y Huber (2015) compuesta por algoritmos adaptados del paquete estadístico R (R Core Team,

2015). En sitios de internet se pueden encontrar sin dificultad herramientas digitales e incluso otros programas libres que cumplen las mismas finalidades, entre ellos se encuentran: Atlas ti, Cassandre, freeQDA, HyperResearch, MAXQDA, Nvivo, RQDA.

Las preferencias diferentes de los hablantes Clinton y Trump respecto a expresiones con I (c_t_i), we (c_t_w) y cada vez la otra persona (c_t_o) son representadas en Figura 1, un gráfico mosaico ('mosaic plot') producido por los R-componentes. Como puede verse claramente en este gráfico, Clinton se refirió más que Trump a las dimensiones que aludía a los pronombres en primera persona singular y plural. Trump, en cambio, obtuvo más frecuentemente asociadas a lo que había dicho o hecho (c_t_o) su oponente.

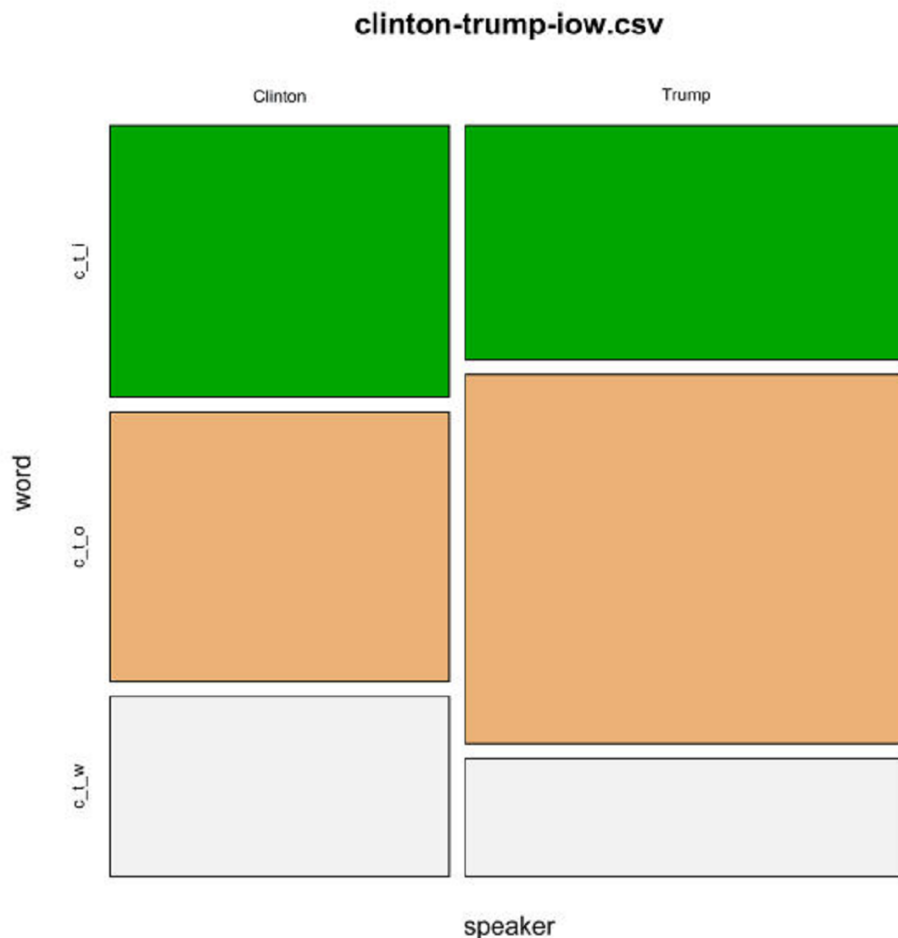


Figura 1. Gráfico mosaico de frecuencias vinculadas

Con respecto a los mencionados hallazgos, se puede concluir que la diferencia entre Clinton-Trump y sus modos de presentar sus argumentos es tan grande como cabría esperar de antemano. El gráfico mosaico muestra visiblemente las tendencias de las expresiones de

Clinton y Trump, particularmente la tendencia de Trump a hablar sobre todo sobre su oponente. Tales resultados (frecuencias, gráfico mosaico) pueden, al menos, justificar en cierto modo el análisis del estilo de comunicación de ganadores y perdedores en campañas electorales en tiempos de debates post-factuales. Es preciso mantener presente que en el análisis cualitativo no cuenta el mero tamaño de los datos, sino que son relevantes los eventos raros o efectos pequeños, ya que pueden contribuir notablemente a explicar un determinado fenómeno. Una tarea a realizar, posteriormente, sería investigar cómo Clinton y Trump prepararon precisamente sus argumentos, que términos utilizaron junto a los pronombres y cómo trataron de animar a sus potenciales electores.

En caso de que parezca necesario, es cierto que se podría continuar por calcular la significancia estadística de las diferencias, pero conviene repetir que, generalmente, la significancia estadística no se considera importante en el contexto del AED. Tukey (en Jones 1986, p. 794) declaró que “Exploratory data analysis [...] does not need probability, significance or confidence” y explicó (p. 815) “If we need a short suggestion of what exploratory data analysis is, I would suggest that: 1. it is an attitude, AND 2. a flexibility, AND 3. some graph paper (or transparencies, or both)”. Además, la significancia estadística como criterio no es fidedigno para interpretaciones posteriores, un aspecto que muchos investigadores como Gigerenzer (2004) han explicado con anterioridad.

4.2. Ejemplo de un diseño de conversión 3.0: Análisis cualitativo de contenido.

En su estudio de la calidad de las instituciones educativas y el liderazgo educativo en España, Letonia y América-Latina, Gento, Huber, González, Palomares y Orden (2015) incluyen un cuestionario sobre el liderazgo educativo (Gento & Cortés, 1995). Esto tiene por finalidad determinar la importancia y la evidencia de los componentes y descriptores importantes del liderazgo (ambos en una escala de nueve puntos desde 1=poco hasta 9=mucho). Gento (2001, 2002) describe ocho dimensiones como características críticas del liderazgo educativo a saber; estas son dimensiones: carismática, afectiva, anticipadora, profesional, participativa, cultural, formativa y administrativa. Las respuestas al cuestionario recogidas hasta enero de 2014 corresponden a respuestas de personas que mayoritariamente se encontraban en España y en Letonia; aunque también se aplicaron algunos de los cuestionarios en ocho países de Latino-América. Por término medio se evaluó la importancia de las ocho dimensiones de liderazgo entre promedios de 7.31 a 7.66, mientras que los valores de evidencia fueron un poco más bajos, a saber, de 6.82 a 7.33 (Gento, 2014). Los valores más altos atribuidos a la importancia de las dimensiones (aunque con pequeñas variaciones)

ponen de manifiesto la importancia atribuida a las dimensiones de tal liderazgo educativo; pero, al mismo tiempo, justo los promedios similares dificultan el intento de descubrir algunos patrones ocultos en los datos y de proponer sugerencias para promover los modos de liderazgo.

Por este motivo, se diseñaron, adicionalmente, entrevistas semi-estructuradas para recoger datos cualitativos referidos al liderazgo educativo y a su repercusión sobre la calidad de las instituciones educativas. Con este propósito, se entrevistó a educadores con cargos directivos que antes habían sido evaluados con el mismo instrumento cuestionario. Esto tuvo por objetivo obtener información de primera mano acerca de las fortalezas y debilidades de liderazgo, así como en relación a su influencia en la educación, las instituciones educativas, situaciones críticas-relevantes de ejercicio de liderazgo y modos de actuar de los líderes educativos. Los procedimientos detallados pueden sintetizarse en una aplicación que consideró un diseño secuencial cuantitativo-cualitativo para obtener información más profunda sobre el tema de investigación.

Gento y Huber (2013) realizaron, transcribieron e interpretaron 32 entrevistas. Para realizar todos estos procedimientos, recurrieron al paradigma de codificación del enfoque cualitativo de investigación. Los hallazgos de estudio confirmaron los resultados del cuestionario cuantitativo; además, los autores añaden algunos detalles interesantes que sobrepasan las valoraciones de la importancia y la evidencia. Sin embargo, aún con tales resultados, aún era difícil la percepción de relaciones o estructuras ocultas en los textos transcritos de las entrevistas. En esta situación, los investigadores decidieron llevar a cabo un análisis según las ideas del diseño de conversión 3.0, es decir, complementar la información ganada mediante el análisis cualitativo-interpretativo de contenido con la información procedente de la aplicación de AED. Además, analizaron adicionalmente el mismo conjunto de datos perteneciente a los códigos cualitativos elaborados por análisis de contenido, mediante técnicas cuantitativas de AED. Para precisar el trabajo realizado, a continuación, se presentan algunos códigos conceptuales que se refieren a cada una de las ocho dimensiones de liderazgo educativo. De este modo, se construyeron ocho meta-códigos, cuyo nombre pertenece a cada una de las dimensiones correspondientes. Un recuento de las frecuencias de este meta-código convirtió los datos cualitativos y produjo los datos cuantitativos para el AED.

Un gráfico representando los resultados del escalamiento multi-dimensional (EMD) de estas frecuencias expone una estructura interesante de las dimensiones de liderazgo. Según Guerrero y Ramírez (2012) el EMD podría caracterizarse como “una técnica multivariante que crea un gráfico aproximado a partir de las similitudes o preferencias de un conjunto de ob-

jetos” (p. 2), lo que para el estudio serían las relaciones entre las dimensiones de liderazgo. La Figura 2 ilustra las relaciones internas de tres grupos de dimensiones de liderazgo, lo que permite interpretar las distancias: las dimensiones 4 y 5 (profesional y participativa) aparecen separadas de las seis dimensiones restantes, pero vinculadas entre sí. Las dimensiones 1 y 2 (carismática y afectiva) se encuentran dentro de un grupo junto con las dimensiones 3, 6, 7 y 8 (anticipadora, cultural, formativa y administrativa), pero separadas claramente de estas últimas dimensiones. De esta manera, los datos proporcionan información altamente relevante, por lo que se pueden identificar patrones interesantes.

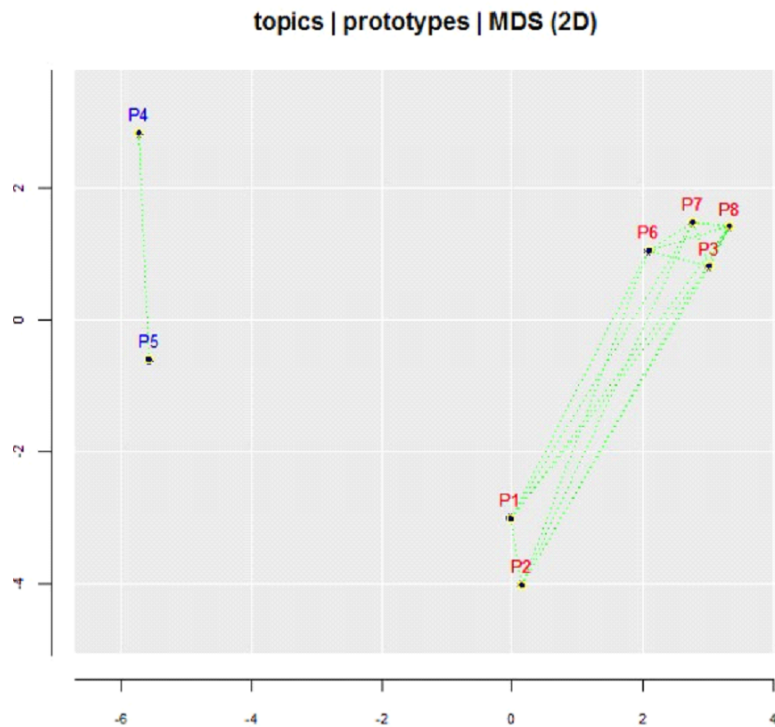


Figura 2. Escalamiento multi-dimensional de frecuencias de códigos

El grupo separado de las dimensiones 4 (profesional) y 5 (participativa), esquina superior izquierda del gráfico, realza el hecho de que los entrevistados están convencidos de que un estilo de liderazgo no debe repetir las estructuras jerárquicas muy comunes en las instituciones educativas en todo el mundo; por el contrario, parecen valorar como más convenientes maneras de liderazgo en colaboración con aquellos de los que se espera que actúen de modo coordinado con el líder. Los datos recomiendan el desarrollo de un estilo de liderazgo que permita a cada miembro responsable en una institución educativa participar en la responsabilidad de tomar y ejecutar decisiones, además de colaborar para lograr metas e intereses comunes. Este hallazgo invita

... to explore further possibilities for 'true democracy' through active participation. We need to create multiple relationships locally and globally with social spaces where we can engage with, and come to value and respect, one another and the group as a whole despite our differences in culture, religion, race, class, sexual orientation, age, life experience, endowment, and world view (Zuber-Skerrit, 2011, p. 4).

El resultado de las mencionadas aportaciones de implicados en el funcionamiento de instituciones educativas debería servir como fundamento sólido para desarrollar hipótesis correspondientes y diseños de estudios futuros en el ámbito de liderazgo educativo.

Idéntica consideración puede desprenderse de la proximidad de las dimensiones carismática y afectiva (1 y 2) sección media en la parte baja del gráfico. En el conjunto de datos de Gento (2014) subgrupos de determinados centros escolares aportan generalmente valores altos para estas dos dimensiones, así como diferencias entre la importancia y la evidencia atribuidas a estas dimensiones del liderazgo. Los datos muestran que la evidencia de la dimensión afectiva del liderazgo aparece más alta que la importancia. Los datos debieran conducir a posteriores estudios que se alineen con las evidencias contrastadas y que informen sobre la relación entre las dimensiones carismática y afectiva en el comportamiento de líderes educativos.

Se ha llevado a cabo también un EMD basado en los datos de los distintos entrevistados. La Figura 3 muestra los resultados de este EMD. Los números que corresponden a cada uno de los entrevistados representan a los individuos que expresaron opiniones durante la realización de las entrevistas. Sin embargo, pareciera que tres personas (D2, D5 y D12), cuyos números identificativos se sitúan en la parte inferior izquierda del gráfico, representan un pequeño grupo con puntos de vista de algún modo diferente respecto al liderazgo educativo o, al menos, a alguna de las dimensiones de tal liderazgo. Este resultado requiere un análisis de las opiniones exactas en este grupo, de la situación particular de sus miembros (tipo de su institución educativa, cargo profesional, etc.) y de la selección de más líderes en situación semejante para futuras entrevistas: hipotéticamente, podría ser factible encontrar relaciones interesantes entre sus situaciones profesionales o desafíos particulares y entre sus modos de entender las condiciones y tareas del liderazgo educativo.

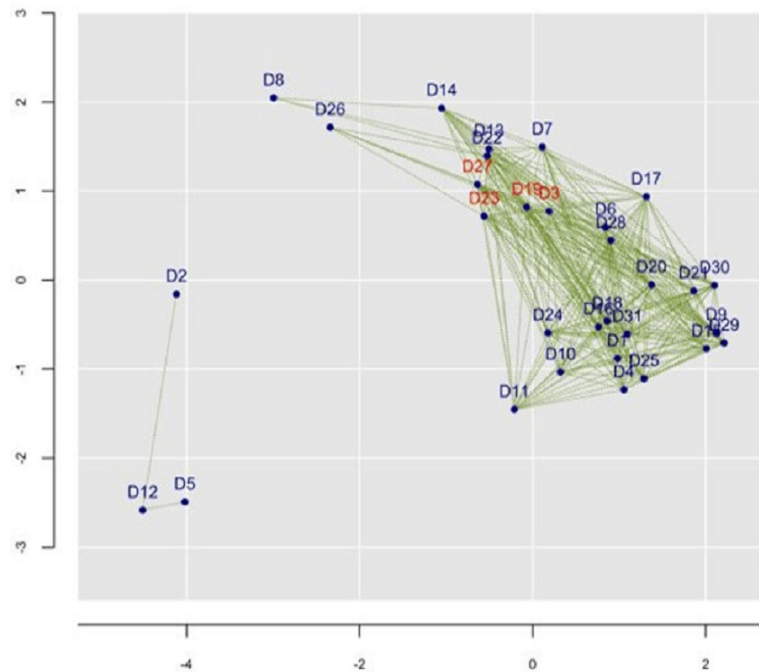


Figura 3. Escalamiento multidimensional de entrevistados

5. Conclusiones

Cabe esperar que los dos pequeños ejemplos descritos en este trabajo puedan contribuir a ilustrar, por un lado, las ventajas de diseños de conversión 3.0 para combinar técnicas cualitativas y cuantitativas en el análisis del mismo conjunto de datos. También se espera, por otro lado, que los ejemplos convengan a investigadores cualitativos de que alternar entre la interpretación de datos cualitativos y el análisis de frecuencias manifiestas en datos obtenidos con técnicas estadísticas exploratorias, puesto que este tipo de determinaciones no es una desviación pecaminosa de posiciones ortodoxas, sino una valiosa ampliación de los enfoques metodológicos. Valiosa en el sentido de que ayuda a los investigadores a explorar más profundamente sus datos con la finalidad de descubrir patrones implícitos o estructuras de datos no necesariamente visibles cuando se utilizan exclusivamente técnicas del análisis cualitativo-interpretativo. Las discusiones breves de los resultados de los ejemplos de la conversión 2.0 y sobre todo de la conversión 3.0, han mostrado cómo la combinación de enfoques interpretativos-cualitativos con técnicas de AED puede abrir el camino a más hipótesis.

En el caso de los debates de TV entre Clinton y Trump el análisis exploratorio del recuento de las palabras clave seleccionadas por medio de un proceder cualitativo es representado en el gráfico mosaico en la Figura 1. El gráfico muestra que Trump se refiere menos que Clinton

a sí mismo (c_t_l) y al público (c_t_w), sino que habla más frecuentemente de su oponente (c_t_o). Ahora sería necesario e interesante analizar el contenido y el sentido de estas declaraciones: ¿Critica a su oponente? ¿la ataca? ¿trata de establecer hechos alternativos?, ¿se presentan algunas otras razones para hechos específicos? ¿qué grupos de votantes reaccionan y de qué modo lo hicieron al momento de enfrentarse a estas declaraciones? Un primer análisis sencillo puede conducir a una serie de preguntas e hipótesis avanzadas. Lo mismo pasa con los resultados de los EMD en el caso del análisis de las dimensiones del liderazgo educativo. Los gráficos representan en un espacio multidimensional las distancias entre las dimensiones de liderazgo (Figura 2) y entre los entrevistados (Figura 3) basadas en la similitud de sus evaluaciones/opiniones. Al respecto se puede reconocer la proximidad entre las dimensiones y las personas, pero por medio del gráfico no es posible identificar la razón por la que aparecen estas similitudes o diferencias. Para dar respuesta a ello, es preciso regresar a los datos originales, desarrollar más hipótesis y continuar el análisis.

En la combinación complementaria de métodos de análisis las técnicas del AED invitan a *jugar con los datos* disponibles y a descubrir estructuras esperadas e inesperadas, incluso la ausencia de tales estructuras puede ser relevante. De este modo, las técnicas del AED pueden motivar a los investigadores a generar nuevas hipótesis y diseñar otros estudios centrados de mejor manera en la pregunta central de la investigación. Es decir, no se mezclan métodos de investigación sin una fundamentación clara, sino que se combinan métodos que complementan el uno al otro. En el ejemplo de la aplicación de la conversión 3.0, la codificación de las entrevistas producía el material para el AED y el resultado de la estructura de distancias mostrado en la Figura 2. En el caso de la triangulación, en un caso particular de métodos mixtos, se combinan perspectivas distintas para ganar una vista completa del caso. Por lo tanto, no se combinan solamente métodos, sino también enfoques para la recuperación de datos o fuentes de datos. En una investigación del fracaso o éxito escolar no se debería concentrar el análisis exclusivamente en los alumnos, sino complementar sus datos individuales (inteligencia, motivación, estilo de aprendizaje, entre otros elementos) con información sobre sus padres, sus profesores, el clima del aula, por ejemplo.

A pesar de todo lo anteriormente especifica, persiste el inconveniente del que Wiggins (2011) advierte. En todo detalle en su crítica de combinar ligeramente métodos de proveniencia cualitativa y cuantitativa, insiste en que no se pueden integrar los puntos de vista epistemológicos en aquellos estudios en que existe una base en los métodos mixtos. Este autor se refiere a los enfrentamientos paradigmáticos entre quienes se encuentran a favor de un punto de vista positivista, aquellos que buscan en sus datos cuantitativos explicaciones causales; mientras que de la perspectiva interpretativa se busca en los datos cualitati-

vos una serie de estructuras de sentido. Pero la realidad de las investigaciones sociales es más compleja. En la descripción de la conversión 1.0 de datos se aprecia que ambos campos suelen complementar el enfoque con ideas y técnicas de otro campo, es decir, aplicar datos cualitativos para la segmentación o las cuantificaciones de sus interpretaciones. Por supuesto, estas estrategias no resuelven el dilema epistemológico de los enfoques opuestos.

Como consecuencia Wiggins (2011) teme que en lugar de diversidad de métodos se aplique solamente una selección limitada que además se encuentre dominada por la vista favorecida del investigador. Este aviso se refiere particularmente a los diseños secuenciales que ponen los diversos métodos dentro de pasos distintos de la investigación y determinan su papel en el estudio, tal como se ha descrito en los ejemplos del presente trabajo. Como aproximación a una solución sugerimos no poner en duda las contribuciones y la importancia de las *perspectivas distintas del mundo* (Wiggins, 2011), sino llamar pragmáticamente la atención acerca de lo que debería ser el foco de cada investigación, a saber, la pregunta de investigación.

Al contestar esta pregunta siempre se tiene que seleccionar y aplicar los métodos adecuados. Patton (2006, p. i) sostenía una posición clara respecto a métodos adecuados: “Methodological appropriateness means that designs should be judged on the extent to which they answer the inquiry question at hand, not whether they adhere to some preordinate standard.”. Entonces, en primer lugar depende de la pregunta de investigación si acaso los métodos particulares son aptos para el estudio. Obviamente la pregunta ¿Cuántos alumnos fracasan en nuestro sistema escolar? demanda otro método de investigación que la pregunta ¿Qué experimentan alumnos en caso de su fracaso escolar? O la pregunta ¿Cómo piensan los compañeros de clase sobre alumnos malos?

En este sentido, Tukey (1962, p. 13) en relación con la dominancia de la pregunta en el proceso de investigación e implícitamente al papel de métodos particulares: “Far better an approximate answer to the right question, which is often vague, than the exact answer to the wrong question, which can always be made precise”.

A lo largo de este trabajo se ha podido enseñar algunas aplicaciones en torno a algunas técnicas de AED y sus posibilidades de aproximarse a respuestas significativas por medio de gráficos, instrumentos que desvelan estructuras dentro de un conjunto de datos. El repertorio de AED es más amplio y ofrece varias técnicas estadísticas combinadas con una gran variedad de representaciones gráficas. Pero antes de detallar estas posibilidades es preci-

so retomar las advertencias de Wiggins (2011) e insistir que en los enfoques 2.0 y 3.0 de conversión de datos se hace una distinción clara entre fases cualitativas y cuantitativas. El modelo 3.0 aplica técnicas de AED a los resultados de un estudio completo de índole cualitativa, por eso los procedimientos estadísticos en los gráficos del AED no influyen o incluso determinan los procesos interpretativos del análisis cualitativo, tales como la codificación del contenido de entrevistas. Al contrario, en una segunda fase las técnicas de AED pueden contribuir en el descubrimiento de estructuras ocultas en el conjunto de interpretaciones de los códigos y conducir a hipótesis avanzadas.

Debido a que no es el objeto de este estudio precisar en las técnicas que ofrece el AED ni tampoco recurrir a más gráficos, sin embargo, en el presente trabajo se han esbozado las técnicas más frecuentes, muy útil para empezar y acercarse a una vista detallada de aplicación de estadísticas descriptivas a las frecuencias de datos originales (por ejemplo frecuencias de palabras clave) o de las interpretaciones (con frecuencias de códigos, entre otras opciones) y representar los resultados por medio de gráficos. En primer lugar, se puede mostrar la forma de la distribución de cada variable en histogramas, después se puede calcular medidas de posición tales como el promedio y la moda, y medidas de dispersión (desviación estándar, cuartiles) y representarlas p.ej. en gráficos de caja (box plots), diagramas de dispersión, scatterplots (de dos o tres dimensiones). Para analizar relaciones entre las variables y mostrarlas en representaciones gráficas, el AED ofrece modos de clasificación o agrupación como el análisis de cluster, la discriminación lineal y el escalamiento multidimensional. En los gráficos vinculados con estas técnicas se pueden apreciar similitudes y diferencias entre las variables expresadas por distancias espaciales.

Finalmente, pareciera que la combinación del enfoque de AED y las estrategias diversas del análisis de datos cualitativos resulta una medida investigativa sensata y útil. Por un lado, las técnicas de AED no interfieren con la interpretación inicial de los datos cualitativos, y por otro, la conversión de los datos cualitativos en frecuencias posibilita una serie de análisis exploratorios que contribuyen a aumentar la comprensión del conjunto de los datos cualitativos y su posterior interpretación. Esto se debería a que contribuyen a hacer visibles patrones de significado. De esta manera el AED complementa el análisis cualitativo sin dominarlo por pensamientos positivistas.

6. Referencias bibliográficas

- Alea, V., Jiménez, E., Muñoz, E., Torrelles, C., & Viladomiu, N. (2014). *Guía para el análisis estadístico con R Commander*. Barcelona: Publicacions i Edicions de la Universitat de Barcelona.
- Andréu, J. (2011). *Las técnicas de análisis de contenido: Una revisión actualizada*. Universidad de Granada. Recuperado de <http://public.centrodeestudiosandaluces.es/pdfs/S200103.pdf>
- Anthony, L. (2011). *AntConc* (Versión 3.2.3). Tokyo, Japan: Waseda University. Recuperado de <http://www.antlab.sci.waseda.ac.jp>
- APA Task Force on Evidence Based Practice (2006). Report of the 2005 presidential task force on evidence-based practice in psychology. *American Psychologist*, 61, 271-285.
- Beaudouin, V. (2016). Statistical analysis of textual data: Benzécri and the French school of data analysis. *Glottometrics*, 33, 56-72.
- Benzécri, J.P. (1992). *Correspondence analysis handbook*. New-York: Marcel Dekker, Inc.
- Berelson, B. (1952). *Content analysis in communication research*. Glencoe, Ill.: The Free Press.
- Bryman, A. (2008). The end of the paradigm wars? En P. Alasuutari, L. Bickman, & J. Brannen (Eds.), *The SAGE handbook of social research methods* (pp. 13-25). London: Sage Publications Ltd.
- Flick, U. (2009). *An introduction to qualitative research* (4th ed.). London: Sage Publications Ltd.
- Gage, N. L. (1989). The paradigm wars and their aftermath. A "historical" sketch of research on teaching since 1989. *Educational Researcher*, 18(7), 4-10.
- Gento, S. (2001). *La institución educativa. II Predictores de calidad*. Buenos Aires: Docencia.
- Gento, S. (2002). *Instituciones educativas para la calidad total*. Madrid: La Muralla.
- Gento, S. (2014). *Liderazgo en instituciones educativas*. Presentación en la Universidad Veracruzana, Campus Xalapa, Mexico.
- Gento, S. & Cortés, J. (1995). *Liderazgo educativo* (3. ed.). Madrid: UNED.
- Gento, S. & Huber, G. L. (2013). *Resultados de entrevistas sobre liderazgo educativo*. Unpublished manuscript, Madrid: UNED.
- Gento, S., Huber, G. L., González, R., Palomares, A. & Orden, V. J. (2015). Promoting the quality of educational institutions by enhancing educational leadership. *US-China Education Review B*, 5(4), 215-232.
- Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics*, 33, 587-606.
- Guerrero, F. M., & Ramírez, J. M. (2012). El análisis de escalamiento multidimensional: una alternativa y un complemento a otras técnicas multivariantes. *La Sociología en sus Es-*

- cenarios, *Revista electrónica de la Universidad de Antioquía (Colombia)*, 25. Descargar de <https://www.uv.es/asepuma/X/K11C.pdf>
- Gürtler, L. & Huber, G. L. (2006). The ambiguous use of language in the paradigms of QUAN and QUAL. *Qualitative Research in Psychology*, 3, 313-328.
- Gürtler, L. & Huber, G. L. (2015a). *AQUAD 7 Manual - R Integration*. Tübingen: Softwarevertrieb Günter Huber.
- Gürtler, L., & Huber, G. L. (2015b). Combining qualitative and quantitative analyses. En G. L. Huber (Ed.), *Qualitative Psychology Nexus Vol. 13: New perspectives on qualitative research* (p. 89-106). Tübingen: Center for Qualitative Psychology.
- Huber, G. L. (2004). *Materials for a workshop on computer-assisted qualitative analysis*. University of Riga, Latvia, July 23-24, 2004.
- Huber, G. L. (2015). AQUAD Siete. *El software para el análisis de datos cualitativos* (software libre, open source). Recuperado de <http://www.aquad.de>
- Inzunza, S. (2014). Razonamiento estadístico de estudiantes universitarios sobre el análisis de datos en un ambiente computacional. *Bolema: Boletim de Educação Matemática*, 28(50), 1262-1286. doi: 10.1590/1980-4415v28n50a13
- Jones, L. V. (Ed.). (1986). *The collected works of John W. Tukey: Philosophy and principles of data analysis 1949-1964 (Vol. IV)*. London: Chapman & Hall.
- Kracauer, S. (1952). The challenge of qualitative content analysis. *Public Opinion Quarterly*, 16, 631- 642.
- National Institute of Standards and Technology (NIST). *Exploratory Data Analysis*. Recuperado de <http://www.itl.nist.gov/div898/handbook/eda/eda.htm>.
- Patton, M. Q. (2006). Utilization-focused evaluation. *Research in the Schools*, 13(1), i-ii.
- R Core Team (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Recuperado de <http://www.R-project.org>
- Smith, J. K. (1983). Quantitative versus qualitative research: An attempt to clarify the issue. *Educational Researcher*, 12(3), 6-13.
- Smith, J. K., & Heshusius, L. (1986). Closing down the conversation: The end of the quantitative-qualitative debate among educational researchers. *Educational Researcher*, 15(1), 4-12.
- Tashakkori, A. & Teddlie, Ch. (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage.
- Teddlie, Ch., & Tashakkori, A. (2006). A general typology of research designs featuring mixed methods. *Research in the Schools*, 13(1), 12-28.
- Tristl, Ch., Müller, M., & Bachmann, V. (2015). Lexicometric analysis: a methodological prelude. En V. Bachmann & M. Müller (Eds.), *Perceptions of the EU in Eastern Europe and Sub-Saharan Africa: Looking in from the outside* (pp. 69-76). Basingstoke: Palgrave

Macmillan.

Tukey, J. W. (1962). The future of data analysis. *Annals of Mathematical Statistics*, 33(1), 1-67.

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, PA: Addison-Wesley.

Wiggins, B. J. (2011). Confronting the dilemma of mixed methods. *Journal of Theoretical and Philosophical Psychology*, 31(1), 44-60.

Zuber-Skerrit, O. (2011). *Action leadership: Towards a participatory paradigm*. Dordrecht: Springer Science + Business Media.